Kalman Filtering for Attitude Estimation with Quaternions and Concepts from Manifold Theory

Pablo Bernal-Polo * and Humberto Martínez-Barberá

University of Murcia, Department of Information and Communication Engineering, 30100 Murcia, Spain; humberto@um.es

* Correspondence: pablo.bernal.polo@gmail.com

Abstract: The problem of attitude estimation is broadly addressed using the Kalman filter formalism and unit quaternions to represent attitudes. This paper is also included in this framework, but introduces a new viewpoint from which the notions of "multiplicative update" and "covariance correction step" are conceived in a natural way. Concepts from manifold theory are used to define the moments of a distribution in a manifold. In particular, the mean and the covariance matrix of a distribution of unit quaternions are defined. Non-linear versions of the Kalman filter are developed applying these definitions. A simulation is designed to test the accuracy of the developed algorithms. The results of the simulation are analyzed and the best attitude estimator is selected according to the adopted performance metric.

Keywords: attitude; orientation; estimation; Kalman filter; quaternion; manifold

1. Introduction

Mechanical state estimation of a vehicle is a field of interest. A vehicle is considered a rigid body, and its state of motion is represented by 4 mathematical objects: two of them represent its position and velocity, and the other two represent its orientation, and angular velocity. This paper is focused on the estimation of the angular state, composed of orientation, and angular velocity.

Although there are other mathematical tools used for estimation [1], the Kalman Filter [2] has become the algorithm par excellence in this area. Because of its simplicity, the rigor and elegance in its mathematical derivation, and its recursive nature it is very attractive for many practical applications. Its non-linear versions have been widely used in orientation estimation: the Extended Kalman Filter (EKF), and the Unscented Kalman Filter (UKF) [3]. However, there are problems arising from the used parametrization to represent the orientation.

The orientation of a system is represented by the rotation transformation that relates two reference frames: a reference frame anchored to that system, and an external reference frame. A thorough survey of attitude representations is provided in Reference [4]. The parametrization used to represent the rotation transformation could be singular, or present discontinuities among others. Table 1 summarizes the main characteristics of the most used parametrizations.

Having in mind that the *special orthogonal group* SO(3) has dimension three, ideally we would seek for a continuous and non-singular representation expressed by 3 parameters. However, since 1964 we know that "...it is topologically impossible to have a global 3-dimensional parametrization without singular points for the rotation group" [5]. Knowing this, we would not be wrong to say that unit quaternions are the most convenient representation we have, and that we will have for orientations. In Reference [6] the literature on attitude estimation is reviewed until 1982, when other

parametrizations like Euler angles were common, and founds the basis of modern quaternion-based attitude estimation, in which this paper is supported. After that work, many others have explored this viewpoint, and have demonstrated its superiority [7–12].

Table 1. Main characteristics of the most used parametrizations to represent an orientation.

Representation	Parameters	Continuous	Non-Singular	Linear Evolution Equation
Euler angles	3	✗	✗	✗
Axis-angle	3 − 4	✗	✗	✗
Rotation matrix	9	✓	✓	✓
Unit quaternion	4	✓	✓	✓

Quaternions are 4-dimensional entities, but only those having unit norm represent a rotation transformation. This fact implies a problem in applying the ordinary Kalman Filter, so different approaches have emerged. Since a quaternion is of dimension 4, one tends to think at first on a 4×4 covariance matrix, and in the direct application of the Kalman Filter [13]. Given that all predictions are contained in the surface defined by the unit constraint, the covariance matrix shrinks in the orthogonal direction to this surface, which leads to a singular covariance matrix after several updates. A second perspective was firstly approached in Reference [6] and was after named as "Multiplicative Extended Kalman Filter" [8,11,12]. In this second approach we define an "error-quaternion" that is transformed to a 3-vector. We use this vector to build the covariance matrix, and we talk about a "3×3 representation of the quaternion covariance matrix". However, there are still details in this adaptation that are currently being developed. Namely, the "covariance correction step" [14].

This paper presents a new viewpoint on the problem of attitude estimation using Kalman filters when the orientation is represented by unit quaternions. Noticing that unit quaternions live in a manifold (the unit sphere in $\mathbb{R}^4$), we use basic concepts from manifold theory to define the mean and covariance matrix of a distribution of unit quaternions. With these definitions we develop two estimators based on the Kalman filter (one EKF-based and another UKF-based) arriving at the concepts of "multiplicative update" and "covariance correction step" in a natural and satisfying way. The inartificial emergence of these ideas establishes a solid foundation for the development of general navigation algorithms. Lastly, we also analyze the accuracy in the estimations of these two estimators using simulations.

The organization of this paper is as follows. In Section 2 we review quaternion basics. We also expose the new viewpoint on the definition of the quaternion mean and covariance matrix. In Section 3 we present the developed estimation algorithms. In Section 4 we define the performance metric, describe the simulation scheme, and present the results of the simulations. We also discuss the results. Finally, Section 5 concludes the paper.

2. Quaternions Describing Orientations

2.1. Quaternions

Quaternions are hypercomplex numbers composed of a real part and an imaginary part. The imaginary part is expressed using three different imaginary units $\{i, j, k\}$ satisfying the Hamilton axiom:

$$i^2 = j^2 = k^2 = i * j * k = -1. \tag{1}$$

A quaternion q can be represented with 4 real numbers, and using several notations:

$$q = q_0 + q_1 i + q_2 j + q_3 k \equiv \tag{2a}$$

$$\equiv (q_0, q_1, q_2, q_3)^T \equiv \tag{2b}$$

$$\equiv (q_0, \mathbf{q})^T. \tag{2c}$$

We will denote quaternions with bold italic symbols (q), while vectors will be denoted with bold upright symbols ($\mathbf{q}$). Vectors will be written in matrix form, and the transposed of a matrix $\mathbf{M}$ will be denoted as $\mathbf{M}^T$.

Quaternion product is defined by Equation (1) which produces the multiplication rule

$$p * q = \begin{pmatrix} p_0 \, q_0 - \mathbf{p} \cdot \mathbf{q} \\ p_0 \, \mathbf{q} + q_0 \, \mathbf{p} + \mathbf{p} \times \mathbf{q} \end{pmatrix}, \tag{3}$$

where $(\cdot)$ represents the usual *dot product*, and $(\times)$ represents the 3-vector *cross product*. Note that the quaternion product $(*)$ is different from the product denoted by $(\otimes)$ in Ref. [4]. Given this multiplication rule, the inverse of a quaternion q (the one for which $q * q^{-1} = q^{-1} * q = 1$) is given by

$$q^{-1} = \frac{1}{\|q\|^2} \, q^* = \frac{1}{\|q\|^2} \, (q_0, -\mathbf{q})^T, \tag{4}$$

where q^* represents the *complex conjugate quaternion*. Note that if q is a unit quaternion (a quaternion with $\|q\| = 1$), then $q^{-1} = q^*$.

2.2. Quaternions Representing Rotations

Each rotation transformation is mapped with a rotation matrix $\mathbf{R}$, and with two unit quaternions q and $-q$ all of them related through

$$\mathbf{R}(q) = \begin{pmatrix} 1 - 2q_2^2 - 2q_3^2 & 2\,(q_1 \, q_2 - q_3 \, q_0) & 2\,(q_1 \, q_3 + q_2 \, q_0) \\ 2\,(q_1 \, q_2 + q_3 \, q_0) & 1 - 2q_1^2 - 2q_3^2 & 2\,(q_2 \, q_3 - q_1 \, q_0) \\ 2\,(q_1 \, q_3 - q_2 \, q_0) & 2\,(q_2 \, q_3 + q_1 \, q_0) & 1 - 2q_1^2 - 2q_2^2 \end{pmatrix}. \tag{5}$$

Note that $\mathbf{R}(q) = \mathbf{R}(-q)$.

Quaternions representing rotations have the form

$$q = (\cos(\theta/2), \, \hat{\mathbf{q}} \sin(\theta/2))^T, \tag{6}$$

where $\hat{\mathbf{q}}$ denotes the unit vector that defines the rotation axis, and θ the angle of rotation. Having this form, they satisfy the restriction

$$\|q\|^2 = q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1. \tag{7}$$

This means that quaternions describing rotations live in the unit sphere of $\mathbb{R}^4$, S^3. This space is a manifold, and some concepts regarding these mathematical objects are useful in our context. In particular, the concept of chart is of special interest.

2.3. Distributions of Unit Quaternions

When dealing with the Kalman filter, the distribution of a random variable $\mathbf{x}$ is encoded by its mean $\bar{\mathbf{x}} = \mathrm{E}[\mathbf{x}]$ and its covariance matrix $\mathbf{P}$ defined as

$$\mathbf{P} = \mathrm{E}\left[(\mathbf{x} - \bar{\mathbf{x}}) \, (\mathbf{x} - \bar{\mathbf{x}})^T \right]. \tag{8}$$

This definition makes sense when our random variables are defined in the Euclidean space. But how do we define the covariance matrix of a random variable living in a manifold like ours? How can we define the covariance for unit quaternions if $q - \bar{q}$ does not represent a rotation? (Unit quaternions form a group under multiplication, but not under addition. This means that the addition of two unit quaternions may not result in another unit quaternion. Therefore, the addition of two unit quaternions may not represent a rotation.) What would be the covariance matrix if each quaternion

was equiprobable in the unit sphere? We cannot redefine the covariance matrix, because the Kalman filter uses this precise form in its derivations, but we can take advantage of the properties of a manifold. Let us retrieve some important definitions:

Definition 1 (**Homeomorphism**). *A homeomorphism is a function $f : X \rightarrow Y$ between two topological spaces X and Y satisfying the next properties:*

- *f is a bijection,*
- *f is continuous,*
- *its inverse function f^{-1} is continuous.*

If such a function exists, we say that X and Y are homeomorphic.

Definition 2 (**Manifold**). *A n-manifold M^n is a topological space in which each point is locally homeomorphic to the Euclidean space $\mathbb{R}^n$. This is, each point $x \in M^n$ has a neighborhood $N \subset M^n$ for which we can define a homeomorphism $f : N \rightarrow B_n$, with B_n the unit ball of $\mathbb{R}^n$.*

Definition 3 (**Chart**). *A chart for a manifold M^n, is a homeomorphism φ from an open subset $U \subset M^n$, to an open subset of the Euclidean space $V \subset \mathbb{R}^n$. This is, a chart is a function*

$$\varphi : U \subset M^n \rightarrow V \subset \mathbb{R}^n, \tag{9}$$

with φ a homeomorphism. Traditionally, a chart is expressed as the pair (U, φ).

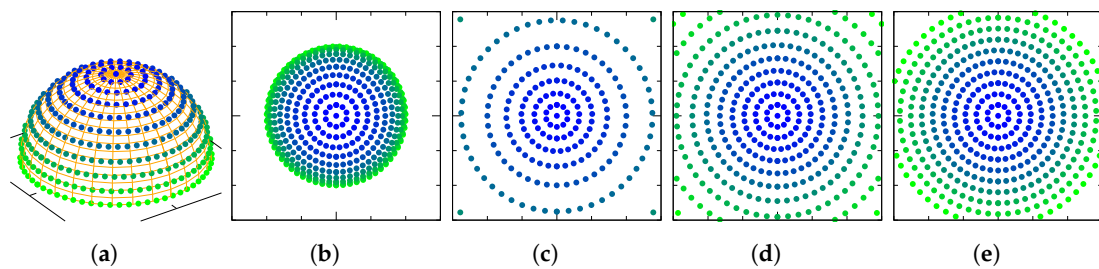Given these definitions we can continue our reasoning.

In Reference [8] it talks about four "attitude error representations". Namely, the one we will call *Orthographic* (O), the *Rodrigues Parameters* (RP), the *Modified Rodrigues Parameters* (MRP), and the *Rotation Vector* (RV). The first three are what we know as *stereographic projections* (and are called *Orthographic*, *Gnomonic*, and *Stereographic* respectively). The last one is a projection called *Equidistant*. But all four are charts defining a homeomorphism from the manifold S^3 to the Euclidean space $\mathbb{R}^3$. This is, they map a point q in the manifold with a point e in $\mathbb{R}^3$. Table 2 arranges these chart definitions, together with their domain and image. We must ensure the charts to be bijections so that they properly define a homeomorphism, and that they do not map q and $-q$ with different points of $\mathbb{R}^3$ since they represent the same rotation. We achieve this by the given definitions of the domain and image for each chart.

Figure 1 shows how points in the sphere S^2 (subspace of the sphere S^3, where quaternions live) are mapped to points in $\mathbb{R}^2$ (subspace of $\mathbb{R}^3$, where the images of the charts are contained) through each one of the named charts. Since our charts are homeomorphisms, it is possible to invert the functions. Figure 2 shows how points from $\mathbb{R}^2$ are mapped to points in the manifold through the inverted charts. As pointed in Reference[8], all four charts provide the same second-order approximation for a point $e \in \mathbb{R}^3$ near the origin, to a quaternion $q \in S^3$:

$$\varphi^{-1}(e) \quad \approx \quad \left(1 - \frac{\|e\|^2}{8} \ , \ \frac{e}{2} \right)^T . \tag{10}$$

Table 2. Main characteristics of the charts studied.

Chart	Domain	Image	$\mathbf{e} = \varphi(\mathbf{q})$	$\mathbf{q} = \varphi^{-1}(\mathbf{e})$
O	$\{\mathbf{q} \in S^3 : q_0 \geq 0\}$	$\{\mathbf{e} \in \mathbb{R}^3 : \|\mathbf{e}\| \leq 2\}$	$2\,\mathbf{q}$	$\begin{pmatrix} \sqrt{1 - \frac{\|\mathbf{e}\|^2}{4}} \\ \mathbf{e}/2 \end{pmatrix}$
RP	$\{\mathbf{q} \in S^3 : q_0 > 0\}$	$\mathbb{R}^3$	$2\,\dfrac{\mathbf{q}}{q_0}$	$\dfrac{1}{\sqrt{4 + \|\mathbf{e}\|^2}} \begin{pmatrix} 2 \\ \mathbf{e} \end{pmatrix}$
MRP	$\{\mathbf{q} \in S^3 : q_0 \geq 0\}$	$\{\mathbf{e} \in \mathbb{R}^3 : \|\mathbf{e}\| \leq 4\}$	$4\,\dfrac{\mathbf{q}}{1 + q_0}$	$\dfrac{1}{16 + \|\mathbf{e}\|^2} \begin{pmatrix} 16 - \|\mathbf{e}\|^2 \\ 8\,\mathbf{e} \end{pmatrix}$
RV	$\{\mathbf{q} \in S^3 : q_0 \geq 0\}$	$\{\mathbf{e} \in \mathbb{R}^3 : \|\mathbf{e}\| \leq \pi\}$	$2\,\widehat{\mathbf{q}}\,\arcsin(\|\mathbf{q}\|)$	$\begin{pmatrix} \cos\left(\frac{\|\mathbf{e}\|}{2}\right) \\ \widehat{\mathbf{e}}\,\sin\left(\frac{\|\mathbf{e}\|}{2}\right) \end{pmatrix}$

| (a) | (b) | (c) | (d) | (e) |

Figure 1. Points in the manifold with $q_3 = 0$ are mapped with points in the Euclidean space through each chart φ. (a) S^2; (b) O; (c) RP; (d) MRP; (e) RV.

| (a) | (b) | (c) | (d) | (e) |
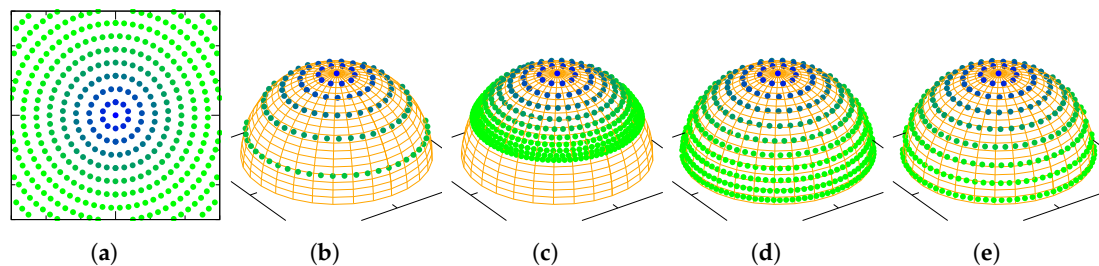
Figure 2. Points in the Euclidean space with $e_3 = 0$ are mapped with points in the manifold through each chart inverse φ^{-1}. (a) $\mathbb{R}^2$; (b) O; (c) RP; (d) MRP; (e) RV.

We should notice that having $\mathbb{R}^3$ and S^3 different metrics, a chart φ will inevitably produce a deformation of the space. However, for quaternions in the neighborhood of the identity quaternion (top of the sphere), our charts behave like the identity transformation between the imaginary part of these quaternions, and the points near the origin in $\mathbb{R}^3$, as suggested by (10). This is a desirable property, as this means that the space around the identity quaternion closely resembles the Euclidean space, which is the space for which the Kalman filter is designed. But this just happens in the neighborhood of the identity quaternion. However, we can extend this property for any quaternion $\overline{q} \in S^3$ noting that any quaternion $q \in S^3$ can be expressed as a "deviation" from the first one through the quaternion product:

$$q = \overline{q} * \delta^{\overline{q}}, \tag{11}$$

where $\delta^{\overline{q}}$ represents such a deviation. (This definition is arbitrary: we could have chosen to relate the quaternions through $q = \delta^{\overline{q}} * \overline{q}$, but it is important to establish one of these definitions, and then be consequent with it. However, (11) entails a computational advantage for the computation of (37).) Then, we define a chart $\varphi_{\overline{q}}$ for each quaternion $\overline{q} \in S^3$ as

$$\mathbf{e}^{\overline{q}} = \varphi_{\overline{q}}(q) = \varphi(\delta^{\overline{q}}), \tag{12}$$

where $\delta^{\overline{q}} = \overline{q}^* * q$, and where we have denoted the point of the Euclidean space mapped with the quaternion $q \in S^3$ through the chart $\varphi_{\overline{q}}$ as $\mathbf{e}^{\overline{q}}$. Then, we will have a set of charts $\{\varphi_{\overline{q}}\}_{\overline{q}}$, each one resembling the Euclidean space around a quaternion $\overline{q} \in S^3$, and mapping this last quaternion to the origin of $\mathbb{R}^3$. We will refer to the Euclidean space associated with the chart $\varphi_{\overline{q}}$ as the $\overline{q}$-centered chart. Thus, the homeomorphism $\varphi_{\overline{q}}^{-1}$ takes a point $\mathbf{e}^{\overline{q}}$ in the $\overline{q}$-centered chart and maps it to a point q in the manifold through

$$q \;=\; \varphi_{\overline{q}}^{-1}\left(\mathbf{e}^{\overline{q}}\right) \;=\; \overline{q} * \varphi^{-1}\left(\mathbf{e}^{\overline{q}}\right). \tag{13}$$

After reviewing these concepts, we can define the covariance matrix of a distribution of unit quaternions.

Given a unit quaternion $\overline{q}$ and a chart φ, we will define the expected value of a distribution of unit quaternions in the $\overline{q}$-centered chart as

$$\overline{\mathbf{e}}^{\overline{q}} \;=\; \mathrm{E}\left[\mathbf{e}^{\overline{q}}\right], \tag{14}$$

and its covariance matrix as

$$\mathbf{P}^{\overline{q}} \;=\; \mathrm{E}\left[\left(\mathbf{e}^{\overline{q}} - \overline{\mathbf{e}}^{\overline{q}}\right)\left(\mathbf{e}^{\overline{q}} - \overline{\mathbf{e}}^{\overline{q}}\right)^T\right], \tag{15}$$

and the probability density of each unit quaternion q would be defined through the homeomorphism $q = \varphi_{\overline{q}}^{-1}(\mathbf{e}^{\overline{q}})$. Then, a distribution of unit quaternions needs of four mathematical objects to be encoded: $(\varphi, \overline{q}, \overline{\mathbf{e}}^{\overline{q}}, \mathbf{P}^{\overline{q}})$. Although a distribution of unit quaternions is unique, given this definition, its expected value $\overline{\mathbf{e}}^{\overline{q}}$ and its covariance matrix $\mathbf{P}^{\overline{q}}$ may take different values depending on the chosen quaternion $\overline{q}$ and chart φ. However, knowing that the Kalman filter is designed for the Euclidean space, it will be convenient to choose a unit quaternion $\overline{q}$ central in the distribution, in order that the manifold space around it closely resembles the most significant region for the covariance matrix in the $\overline{q}$-centered chart. It is particularly convenient to choose a quaternion $\overline{q}$ such that $\overline{\mathbf{e}}^{\overline{q}} = \mathbf{0}$, so that the covariance matrix is centered in the origin of the $\overline{q}$-centered chart.

2.4. Transition Maps

At some step of the Kalman filter, we will have a distribution of unit quaternions defined in a $\overline{q}$-centered chart, and we will be interested in expressing our distribution in another $\overline{p}$-centered chart. The concept of transition map is relevant for this purpose.

Definition 4 (Transition map). *Given two charts $(U_\alpha, \varphi_\alpha)$ and (U_β, φ_β) for a manifold M, with $U_{\alpha\beta} = U_\alpha \cap U_\beta \neq \varnothing$, we can define a function $\varphi_{\alpha\beta} : \varphi_\alpha(U_{\alpha\beta}) \to \varphi_\beta(U_{\alpha\beta})$ as*

$$\varphi_{\alpha\beta}(x) \;=\; \varphi_\beta\left(\varphi_\alpha^{-1}(x)\right), \tag{16}$$

with $x \in \varphi_\alpha(U_{\alpha\beta})$. The function $\varphi_{\alpha\beta}$ is called a transition map. Being that φ_α and φ_β are homeomorphisms, so is $\varphi_{\alpha\beta}$.

For the present case, let us consider two unit quaternions $\overline{p}$ and $\overline{q}$ both related through

$$\overline{p} \;=\; \overline{q} * \overline{\delta}. \tag{17}$$

These two quaternions define the charts $\varphi_{\overline{p}}$ and $\varphi_{\overline{q}}$. We build the transition map that relates a point $\mathbf{e}^{\overline{q}}$ expressed in the $\overline{q}$-centered chart with a point $\mathbf{e}^{\overline{p}}$ expressed in the $\overline{p}$-centered chart doing

$$\mathbf{e}^{\overline{p}} \;=\; \varphi_{\overline{p}}\left(\varphi_{\overline{q}}^{-1}\left(\mathbf{e}^{\overline{q}}\right)\right) \;=\; \tag{18a}$$

$$=\; \varphi\left(\overline{p}^* * \overline{q} * \varphi^{-1}\left(\mathbf{e}^{\overline{q}}\right)\right) \;=\; \tag{18b}$$

$$=\; \varphi\left(\overline{\delta}^* * \varphi^{-1}\left(\mathbf{e}^{\overline{q}}\right)\right). \tag{18c}$$

That is to say, first we take the point $\mathbf{e}^{\overline{q}}$ in the $\overline{q}$-centered chart, and we obtain its associated quaternion q in the manifold using $\varphi_{\overline{q}}^{-1}$. Then, we transform this quaternion q to a point $\mathbf{e}^{\overline{p}}$ in the $\overline{p}$-centered chart. Nevertheless, knowing the quaternion $\overline{\delta}$ we do not need to explicitly compute q. In fact, being able to express the same quaternion q as two different deviations,
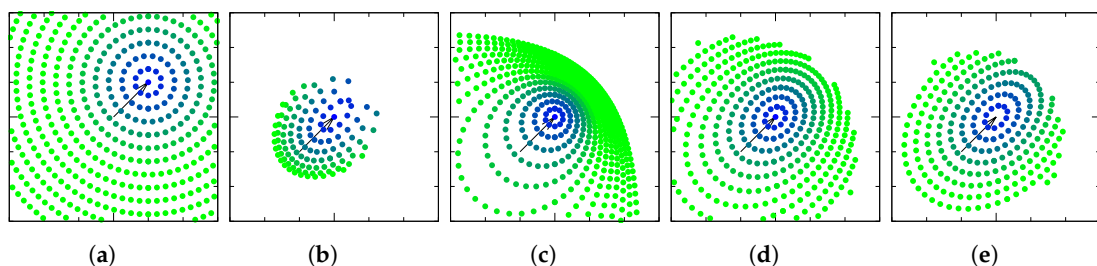
$$\left.\begin{array}{l} q = \overline{q} * \delta^{\overline{q}} \\ q = \overline{p} * \delta^{\overline{p}} \end{array}\right\} \quad\Longrightarrow\quad \delta^{\overline{p}} = \underbrace{\overline{p}^* * \overline{q}}_{\overline{\delta}^*} * \delta^{\overline{q}}. \tag{19}$$

Note the equivalence of expressions (18c) and (19).

Table 3 displays the transition maps for the charts studied. The detailed derivations of these transition maps can be found in Appendix A. Figure 3 attempts to provide some insight into how points are transformed through the transition map of each chart.

Table 3. Transition maps for the charts studied.

Chart	Transition Map $\mathbf{e}^{\overline{p}}\left(\mathbf{e}^{\overline{q}}\right)$
O	$\overline{\delta}_0\,\mathbf{e}^{\overline{q}} - \sqrt{4 - \|\mathbf{e}^{\overline{q}}\|^2}\,\overline{\delta} - \overline{\delta} \times \mathbf{e}^{\overline{q}}$
RP	$2\,\dfrac{\overline{\delta}_0\,\mathbf{e}^{\overline{q}} - 2\,\overline{\delta} - \overline{\delta} \times \mathbf{e}^{\overline{q}}}{2\,\overline{\delta}_0 + \overline{\delta} \cdot \mathbf{e}^{\overline{q}}}$
MRP	$4\,\dfrac{8\,\overline{\delta}_0\,\mathbf{e}^{\overline{q}} - \left(16 - \|\mathbf{e}^{\overline{q}}\|^2\right)\overline{\delta} - 8\,\overline{\delta} \times \mathbf{e}^{\overline{q}}}{16 + \|\mathbf{e}^{\overline{q}}\|^2 + \overline{\delta}_0\left(16 - \|\mathbf{e}^{\overline{q}}\|^2\right) + 8\,\overline{\delta} \cdot \mathbf{e}^{\overline{q}}}$
RV	$2\,\dfrac{\delta^{\overline{p}}}{\|\delta^{\overline{p}}\|}\,\arcsin\|\delta^{\overline{p}}\|\,,\quad\text{with}\quad \delta^{\overline{p}} = \overline{\delta}_0\,\widehat{\mathbf{e}}^{\overline{q}}\sin\left(\tfrac{\|\mathbf{e}^{\overline{q}}\|}{2}\right) - \cos\left(\tfrac{\|\mathbf{e}^{\overline{q}}\|}{2}\right)\overline{\delta} - \overline{\delta} \times \widehat{\mathbf{e}}^{\overline{q}}\sin\left(\tfrac{\|\mathbf{e}^{\overline{q}}\|}{2}\right)$

(a)   (b)   (c)   (d)   (e)

Figure 3. Points in a $\overline{q}$-centered chart are transformed using the transition map defined by each chart, and travel to the chart centered in the quaternion mapped with $\mathbf{e}^{\overline{q}} = (1, 1, 0)^T$ in the previous chart. (a) $\mathbb{R}^2$; (b) O; (c) RP; (d) MRP; (e) RV.

3. Manifold Kalman Filters

In this section we present the models adopted for the Manifold Kalman Filters (MKF), and we display the resulting algorithms.

The state of the system at a time t is defined by an orientation, encoded with a unit quaternion q_t, and by an angular velocity ω'_t. We will consider them to be random variables, and we will try to estimate their value using a Kalman filter.

Our unit quaternions $\{q_t \in \mathbb{H} : \|q_t\| = 1\}$ will define the rotation transformation that relates a vector $\mathbf{v}'_t$ expressed in a reference frame $\mathcal{S}'$ attached to the solid whose state we want to describe, with the same vector $\mathbf{v}_t$ expressed in an external reference frame $\mathcal{S}$:

$$\mathbf{v}_t = \mathbf{R}(q_t)\,\mathbf{v}'_t \equiv v_t = q_t * v'_t * q^*_t. \tag{20}$$

For example, if we measure an acceleration $\mathbf{a}'_t$ in reference frame $\mathcal{S}'$, the acceleration in the inertial reference frame $\mathcal{S}$ would be given by $\mathbf{a}_t = \mathbf{R}(q_t)\,\mathbf{a}'_t$. This acceleration would be the one that we would have to integrate to obtain the position estimated by an accelerometer.

The vector ω'_t will define the angular velocity of the solid measured in $\mathcal{S}'$. Note that we do not include the bias of the sensors in the state of our system. We will assume that our sensors are calibrated, so the biases are zero.

We can predict the value of the random variables that describe the state of our system through the following motion equations:

$$\frac{d\,\omega'(t)}{dt} = \mathbf{q}^\omega(t), \tag{21}$$

$$\frac{d\,q(t)}{dt} = \frac{1}{2}\,q(t) * \omega'(t) = \frac{1}{2}\,q(t) * \begin{pmatrix} 0 \\ \omega'(t) \end{pmatrix}, \tag{22}$$

where $\mathbf{q}^\omega(t)$ is a random variable that represents the process noise, and is associated with the torque acting on the system, and with its inertia tensor. Its expected value at a given time t will be denoted as $\overline{\mathbf{q}}^\omega_t$, and its covariance matrix will be denoted as $\mathbf{Q}^\omega_t$.

We will assume that we have sensors giving measurements of angular velocity ω^m_t (which provide information about the relative change in orientation), and of a vector $\mathbf{v}^m_t$ whose value $\mathbf{v}_t$, expressed in the external reference frame $\mathcal{S}$, is known (this provides information about absolute orientation). Examples of such sensors could be a gyroscope giving angular velocity measurements, an accelerometer measuring the gravity vector near the Earth surface ($\mathbf{v}_t := -\mathbf{g}$), or a magnetometer measuring the Earth magnetic field ($\mathbf{v}_t := \mathbf{B}$). The measurement model relates these measurements with the variables that describe the state of the system:

$$\mathbf{v}^m_t = \mathbf{R}^T(q_t)\,(\mathbf{q}^\mathbf{v}_t + \mathbf{v}_t) + \mathbf{r}^\mathbf{v}_t, \tag{23}$$

$$\omega^m_t = \omega'_t + \mathbf{r}^\omega_t, \tag{24}$$

where $\mathbf{r}^\omega_t$ and $\mathbf{r}^\mathbf{v}_t$ are random variables with zero mean and covariance matrices $\mathbf{R}^\omega_t$ and $\mathbf{R}^\mathbf{v}_t$ respectively that represent the measurement noises, and $\mathbf{q}^\mathbf{v}_t$ is another random variable with mean $\overline{\mathbf{q}}^\mathbf{v}_t$ and covariance matrix $\mathbf{Q}^\mathbf{v}_t$ representing external disturbances in the measurement of the vector $\mathbf{v}_t$. For example, it could represent accelerations others than gravity for an accelerometer, or magnetic disturbances produced by moving irons for a magnetometer.

We will assume that the measurements arrive at discrete times $\{t_n\}_n$. The format $x_{t|t_n}$ will be used to denote a variable x at a time t, having included measurements up to a time t_n with $t > t_n$. For the n-th time stamp, in which a measurement arrives, we will write $x_{t|n}$ for the sake of simplicity. Then, our knowledge about the state at a time t, having included measurements up to a time t_n with $t > t_n$, is described by a distribution encoded in the collection of mathematical

objects $\left(\varphi, \overline{p}, \mathbf{x}^{\overline{p}}_{t|n}, \mathbf{P}^{\overline{p}}_{t|n} \right)$ as described in Section 2.3. For the present case, $\mathbf{x}^{\overline{p}}_{t|n} = \left(\overline{\mathbf{e}}^{\overline{p}}_{t|n}, \overline{\omega}'_{t|n} \right)^T$ is the expected value of the distribution, and $\mathbf{P}^{\overline{p}}_{t|n}$ is its 6×6 covariance matrix, both expressing the quaternion distribution in the $\overline{p}$-centered chart. Preferably, $\overline{p}$ will be a unit quaternion central in the distribution, so that the mapping of points from the $\overline{p}$-centered chart to the manifold causes minimal deformation in such distribution. The unit quaternion $\overline{q}_{t|n} = \varphi^{-1}_{\overline{p}} \left(\overline{\mathbf{e}}^{\overline{p}}_{t|n} \right)$ will be our best estimation of the real quaternion q_t that defines the orientation of the system with respect to the external reference frame $\mathcal{S}$ at time t.

The following subsections present the developed Kalman filters: one version based on the EKF and another version based on the UKF. The EKF is based on the linearization of the non-linear models to calculate the predicted covariance matrices. That is, the EKF approximates non-linear functions using their Jacobian matrices. To apply the EKF, our functions must be differentiable. On the other hand, the UKF is based on a deterministic sampling to approximate the distribution of our random variables. We select a minimal set of samples whose mean and covariance matrix are those of the state distribution. Then, they are transformed by the non-linear models, and the resulting set of points is used to compute the means and covariance matrices necessary to perform the Kalman update. This second approach does not need the functions to be differentiable.

3.1. Manifold Extended Kalman Filter

In this section we present the EKF-based estimator: the Manifold Extended Kalman Filter (MEKF). We offer here the main results of the more detailed derivation given in Appendix B.

A measurement

$$\mathbf{z}_n = \begin{pmatrix} \mathbf{v}^m_n \\ \omega^m_n \end{pmatrix} \tag{25}$$

arrives at time t_n. Our knowledge about the orientation at a previous time t_{n-1} is described by a distribution expressed in the $\overline{q}_{n-1|n-1}$-centered chart. We assume that this distribution has mean

$$\overline{\mathbf{x}}^{\overline{q}_{n-1|n-1}}_{n-1|n-1} = \begin{pmatrix} \overline{\mathbf{e}}^{\overline{q}_{n-1|n-1}}_{n-1|n-1} = 0 \\ \overline{\omega}'_{n-1|n-1} \end{pmatrix}, \tag{26}$$

and covariance matrix $\mathbf{P}^{\overline{q}_{n-1|n-1}}_{n-1|n-1}$. This is, we have an initial four

$$\left(\varphi, \overline{q}_{n-1|n-1}, \overline{\omega}'_{n-1|n-1}, \mathbf{P}^{\overline{q}_{n-1|n-1}}_{n-1|n-1} \right). \tag{27}$$

The state prediction at time t_n given all the information up to t_{n-1} is computed through

$$\overline{\omega}'_{n|n-1} = \overline{\omega}'_{n-1|n-1}, \tag{28}$$

$$\delta^\omega_n = \begin{pmatrix} \cos\left(\dfrac{\|\overline{\omega}'_{n|n-1}\| \, \Delta t_n}{2} \right) \\ \dfrac{\overline{\omega}'_{n|n-1}}{\|\overline{\omega}'_{n|n-1}\|} \sin\left(\dfrac{\|\overline{\omega}'_{n|n-1}\| \, \Delta t_n}{2} \right) \end{pmatrix}, \tag{29}$$

$$\overline{q}_{n|n-1} = \overline{q}_{n-1|n-1} * \delta^\omega_n, \tag{30}$$

$$\mathbf{F}_n = \begin{pmatrix} \mathbf{R}^T(\delta^\omega_n) & \mathbf{I}\,\Delta t_n \\ \mathbf{0} & \mathbf{I} \end{pmatrix}, \tag{31}$$

$$\mathbf{P}^{\overline{q}_{n|n-1}}_{n|n-1} = \mathbf{F}_n \left[\mathbf{P}^{\overline{q}_{n-1|n-1}}_{n-1|n-1} + \mathbf{Q}_n \right] \mathbf{F}^T_n, \tag{32}$$

with

$$\mathbf{Q}_n \quad = \quad \begin{pmatrix} \mathbf{Q}_n^\omega \frac{(\Delta t_n)^3}{3} & -\mathbf{Q}_n^\omega \frac{(\Delta t_n)^2}{2} \\ -\mathbf{Q}_n^\omega \frac{(\Delta t_n)^2}{2} & \mathbf{Q}_n^\omega \Delta t_n \end{pmatrix}. \tag{33}$$

The measurement prediction at the same time is given by

$$\overline{\mathbf{v}}_{n|n\text{-}1}^m \quad = \quad \mathbf{R}^T\left(\overline{q}_{n|n\text{-}1}\right)\left(\overline{\mathbf{q}}_n^{\mathbf{v}} + \mathbf{v}_n\right), \tag{34}$$

$$\overline{\omega}_{n|n\text{-}1}^m \quad = \quad \overline{\omega}'_{n|n\text{-}1}, \tag{35}$$

$$\overline{\mathbf{z}}_{n|n\text{-}1} \quad = \quad \begin{pmatrix} \overline{\mathbf{v}}_{n|n\text{-}1}^m \\ \overline{\omega}_{n|n\text{-}1}^m \end{pmatrix}, \tag{36}$$

$$\mathbf{H}_n \quad = \quad \begin{pmatrix} \left[\overline{\mathbf{v}}_{n|n\text{-}1}^m\right]_\times & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}, \tag{37}$$

$$\mathbf{S}_{n|n\text{-}1} \quad = \quad \mathbf{H}_n \, \mathbf{P}_{n|n\text{-}1}^{\overline{q}_{n|n\text{-}1}} \, \mathbf{H}_n^T + \begin{pmatrix} \mathbf{R}^T\left(\overline{q}_{n|n\text{-}1}\right)\mathbf{Q}_n^{\mathbf{v}}\,\mathbf{R}\left(\overline{q}_{n|n\text{-}1}\right) + \mathbf{R}_n^{\mathbf{v}} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_n^\omega \end{pmatrix}, \tag{38}$$

where $[\mathbf{v}]_\times$ stands for

$$[\mathbf{v}]_\times \quad = \quad \begin{pmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{pmatrix}. \tag{39}$$

At this point, we compute the Kalman gain $\mathbf{K}_n$ and use it to obtain the optimal estimation of the state:

$$\mathbf{K}_n \quad = \quad \mathbf{P}_{n|n\text{-}1}^{\overline{q}_{n|n\text{-}1}} \, \mathbf{H}_n^T \, \mathbf{S}_{n|n\text{-}1}^{-1}, \tag{40}$$

$$\overline{\mathbf{x}}_{n|n}^{\overline{q}_{n|n\text{-}1}} \quad = \quad \overline{\mathbf{x}}_{n|n\text{-}1}^{\overline{q}_{n|n\text{-}1}} + \mathbf{K}_n\left(\mathbf{z}_n - \overline{\mathbf{z}}_{n|n\text{-}1}\right), \tag{41}$$

$$\mathbf{P}_{n|n}^{\overline{q}_{n|n\text{-}1}} \quad = \quad \left(\mathbf{I} - \mathbf{K}_n\,\mathbf{H}_n\right)\mathbf{P}_{n|n\text{-}1}^{\overline{q}_{n|n\text{-}1}}, \tag{42}$$

where $\overline{\mathbf{x}}_{n|n\text{-}1}^{\overline{q}_{n|n\text{-}1}} = \left(\overline{\mathbf{e}}_{n|n\text{-}1}^{\overline{q}_{n|n\text{-}1}} = \mathbf{0}, \overline{\omega}'_{n|n\text{-}1}\right)^T$. Finally, we need to obtain the updated unit quaternion, $\overline{q}_{n|n}$, and compute the mean and the covariance matrix in the $\overline{q}_{n|n}$-centered chart, so that the distribution is expressed in the same conditions as at the beginning of the iteration. The point $\overline{\mathbf{e}}_{n|n}^{\overline{q}_{n|n\text{-}1}}$ that results from (41), and that is defined in the $\overline{q}_{n|n\text{-}1}$-centered chart, correspond to a unit quaternion in the manifold. This is the updated unit quaternion $\overline{q}_{n|n}$ which we are looking for:

$$\overline{q}_{n|n} \quad = \quad \varphi_{\overline{q}_{n|n\text{-}1}}^{-1}\left(\overline{\mathbf{e}}_{n|n}^{\overline{q}_{n|n\text{-}1}}\right) \quad = \tag{43a}$$

$$= \quad \overline{q}_{n|n\text{-}1} * \varphi^{-1}\left(\overline{\mathbf{e}}_{n|n}^{\overline{q}_{n|n\text{-}1}}\right) \quad = \tag{43b}$$

$$= \quad \overline{q}_{n|n\text{-}1} * \overline{\delta}_n. \tag{43c}$$

Knowing that the Kalman update (41) could produce any point in the $\overline{q}_{n|n\text{-}1}$-centered chart we will need to "saturate" to the closest point contained in the image of each chart. The point $\overline{\mathbf{e}}_{n|n}^{\overline{q}_{n|n\text{-}1}}$ in the $\overline{q}_{n|n\text{-}1}$-centered chart is the origin in the $\overline{q}_{n|n}$-centered chart. Then, the expected value of the state in this new chart will be given by $\overline{\mathbf{x}}_{n|n}^{\overline{q}_{n|n}} = \left(\overline{\mathbf{e}}_{n|n}^{\overline{q}_{n|n}} = \mathbf{0}, \overline{\omega}'_{n|n}\right)^T$, as at the beginning of the iteration.

To update the covariance matrix we need to consider its definition (15). We want to compute $\mathbf{P}^{\overline{q}_{n|n}}$ having $\mathbf{P}^{\overline{q}_{n|n\text{-}1}}$ and knowing the relation $\mathbf{e}^{\overline{p}}\left(\mathbf{e}^{\overline{q}}\right)$ provided by the transition maps in Table 3. Continuing with the EKF philosophy, the update for the covariance matrix will be found by linearizing

$\mathbf{e}^{\overline{p}}(\mathbf{e}^{\overline{q}})$ around the point where the majority of information is comprised (in our case, the point $\overline{\mathbf{e}^{\overline{q}}} = \overline{\mathbf{e}}_{n|n}^{\overline{q}n|n-1}$):

$$e_i^{\overline{p}}(\mathbf{e}^{\overline{q}}) = e_i^{\overline{p}}(\overline{\mathbf{e}^{\overline{q}}}) + \sum_j \frac{\partial e_i^{\overline{p}}(\mathbf{e}^{\overline{q}})}{\partial e_j^{\overline{q}}}\bigg|_{\mathbf{e}^{\overline{q}}=\overline{\mathbf{e}}^{\overline{q}}} (e_j^{\overline{q}} - \overline{e}_j^{\overline{q}}) + \mathcal{O}\left(\|\mathbf{e}^{\overline{q}} - \overline{\mathbf{e}}^{\overline{q}}\|^2\right), \qquad (44)$$

where we have used the *big O notation* to describe the limiting behavior of the error term of the approximation as $\mathbf{e}^{\overline{q}} \to \overline{\mathbf{e}}^{\overline{q}}$. In particular, if we define

$$(\mathbf{T})_{ij} = \frac{\partial e_i^{\overline{p}}(\mathbf{e}^{\overline{q}})}{\partial e_j^{\overline{q}}}\bigg|_{\mathbf{e}^{\overline{q}}=\overline{\mathbf{e}}^{\overline{q}}}, \qquad (45)$$

then,

$$\mathbf{e}^{\overline{p}} - \overline{\mathbf{e}}^{\overline{p}} \approx \mathbf{e}^{\overline{p}}(\mathbf{e}^{\overline{q}}) - \mathbf{e}^{\overline{p}}(\overline{\mathbf{e}}^{\overline{q}}) \approx \mathbf{T}\left(\mathbf{e}^{\overline{q}} - \overline{\mathbf{e}}^{\overline{q}}\right), \qquad (46)$$

and the final update for the covariance matrix will be computed through

$$\mathbf{P}_{n|n}^{\overline{q}n|n} = \mathrm{E}\left[(\mathbf{x}_{n|n}^{\overline{q}n|n} - \overline{\mathbf{x}}_{n|n}^{\overline{q}n|n})(\mathbf{x}_{n|n}^{\overline{q}n|n} - \overline{\mathbf{x}}_{n|n}^{\overline{q}n|n})^T \right] \approx \qquad (47a)$$

$$\approx \begin{pmatrix} \mathbf{T}(\overline{\delta}_n) & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \mathbf{P}_{n|n}^{\overline{q}n|n-1} \begin{pmatrix} \mathbf{T}(\overline{\delta}_n) & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}^T. \qquad (47b)$$

Table 4 summarizes the resulting **T**-matrix for each chart, along with their application domain. A detailed derivation of these **T**-matrices can be found in Appendix C.

Table 4. **T**-matrices for the transition maps of the charts studied.

Chart	$\mathbf{T}(\overline{\delta})$ Matrix	Domain
O	$\overline{\delta}_0 \mathbf{I} - [\overline{\delta}]_\times + \dfrac{\overline{\delta}\,\overline{\delta}^T}{\overline{\delta}_0}$	$\{\, \overline{\delta} \in S^3 \;:\; \overline{\delta}_0 > 0 \,\}$
RP	$\overline{\delta}_0 \left(\overline{\delta}_0 \mathbf{I} - [\overline{\delta}]_\times \right)$	$\{\, \overline{\delta} \in S^3 \;:\; \overline{\delta}_0 \neq 0 \,\}$
MRP	$\dfrac{1}{2}\left[(1+\overline{\delta}_0)\left(\overline{\delta}_0 \mathbf{I} - [\overline{\delta}]_\times \right) + \overline{\delta}\,\overline{\delta}^T \right]$	$\{\, \overline{\delta} \in S^3 \;:\; \overline{\delta}_0 \geq 0 \,\}$
RV	$\left[\overline{\delta}_0 \left(\mathbf{I} - \widehat{\overline{\delta}}\,\widehat{\overline{\delta}}^T \right) - [\overline{\delta}]_\times \right] \dfrac{\|\overline{\delta}\|}{\arcsin \|\overline{\delta}\|} + \widehat{\overline{\delta}}\,\widehat{\overline{\delta}}^T$	$\{\, \overline{\delta} \in S^3 \;:\; \overline{\delta}_0 \geq 0 \,,\; \|\overline{\delta}\| \neq 0 \,\}$

After the final computation we obtain the four

$$\left(\varphi, \overline{q}_{n|n}, \overline{\omega}'_{n|n}, \mathbf{P}_{n|n}^{\overline{q}n|n} \right), \qquad (48)$$

that is a condition equivalent to (27) in which we started the iteration.

3.2. Manifold Unscented Kalman Filter

In this section we present the UKF-based estimator: the Manifold Unscented Kalman Filter (MUKF).

A measurement $\mathbf{z}_n$ arrives at time t_n. Our knowledge about the orientation at a previous time $t_{n\text{-}1}$ is described by a distribution expressed in the $\bar{q}_{n\text{-}1|n\text{-}2}$-centered chart. This distribution is encoded in the four

$$\left(\varphi \, , \, \bar{q}_{n\text{-}1|n\text{-}2} \, , \, \bar{\mathbf{x}}_{n\text{-}1|n\text{-}1}^{\bar{q}_{n\text{-}1|n\text{-}2}} \, , \, \mathbf{P}_{n\text{-}1|n\text{-}1}^{\bar{q}_{n\text{-}1|n\text{-}2}} \right) . \tag{49}$$

The first step in the UKF is to create the augmented $N \times 1$ mean $\widetilde{\mathbf{x}}_n$ and $N \times N$ covariance matrix $\widetilde{\mathbf{P}}_n$. Since the measurement equations are linear for the random variables $\mathbf{r}_t^{\omega}$ and $\mathbf{r}_t^{\mathbf{v}}$, we can leave their covariance matrices out of the augmented one and add them later:

$$\widetilde{\mathbf{x}}_n \;\;=\;\; \begin{pmatrix} \bar{\mathbf{x}}_{n\text{-}1|n\text{-}1}^{\bar{q}_{n\text{-}1|n\text{-}2}} \\ \bar{\mathbf{q}}_n^{\omega} \\ \bar{\mathbf{q}}_n^{\mathbf{v}} \end{pmatrix} , \tag{50}$$

$$\widetilde{\mathbf{P}}_n \;\;=\;\; \begin{pmatrix} \mathbf{P}_{n\text{-}1|n\text{-}1}^{\bar{q}_{n\text{-}1|n\text{-}2}} & 0 & 0 \\ 0 & \mathbf{Q}_n^{\omega} & 0 \\ 0 & 0 & \mathbf{Q}_n^{\mathbf{v}} \end{pmatrix} . \tag{51}$$

Then, we obtain the matrix $\mathbf{L}_n$ which satisfies $\mathbf{L}_n \mathbf{L}_n^T = \widetilde{\mathbf{P}}_n$ and we use it to generate the $2N+1$ sigma points $\{\mathcal{X}_j\}_{j=0}^{2N}$ as described in Ref. [15]:

$$\mathcal{X}_{i,0} \;\;=\;\; (\widetilde{\mathbf{x}}_n)_i, \tag{52a}$$

$$\mathcal{X}_{i,j} \;\;=\;\; (\widetilde{\mathbf{x}}_n)_i \;+\; \frac{(\mathbf{L}_n)_{ij}}{\sqrt{2W_j}} \qquad \text{for} \quad j = 1, \dots, N, \tag{52b}$$

$$\mathcal{X}_{i,j+N} \;\;=\;\; (\widetilde{\mathbf{x}}_n)_i \;-\; \frac{(\mathbf{L}_n)_{ij}}{\sqrt{2W_j}} \qquad \text{for} \quad j = 1, \dots, N, \tag{52c}$$

being $W_j = (1 - W_0)/(2N)$ for $j \neq 0$ where W_0 regulates the importance given to the sigma point $\mathcal{X}_0$ in the computation of the mean. These sigma points $\{\mathcal{X}_j\}_j$ are expressed in the $\bar{q}_{n\text{-}1|n\text{-}2}$-centered chart. We need to express them in the manifold before applying the evolution equations and the measurement equations:

$$\mathcal{X}_j^q \;\;=\;\; \varphi_{\bar{q}_{n\text{-}1|n\text{-}2}}^{-1}\left(\mathcal{X}_j^{\mathbf{e}} \right) \;\;=\;\; \bar{q}_{n\text{-}1|n\text{-}2} * \varphi^{-1}\left(\mathcal{X}_j^{\mathbf{e}} \right), \tag{53}$$

$$\mathcal{Y}_j^{\omega} \;\;=\;\; \mathcal{X}_j^{\omega} \;+\; \mathcal{X}_j^{q^{\omega}} \Delta t_n, \tag{54}$$

$$\mathcal{Y}_j^q \;\;=\;\; \mathcal{X}_j^q * \begin{pmatrix} \cos\left(\dfrac{\|\mathcal{Y}_j^{\omega}\| \, \Delta t_n}{2} \right) \\ \hat{\mathcal{Y}}_j^{\omega} \, \sin\left(\dfrac{\|\mathcal{Y}_j^{\omega}\| \, \Delta t_n}{2} \right) \end{pmatrix} , \tag{55}$$

$$\mathcal{Z}_j^{\mathbf{v}} \;\;=\;\; \mathbf{R}^T\left(\mathcal{X}_j^q \right) \left(\mathcal{X}_j^{\mathbf{v}} \;+\; \mathbf{v}_t \right), \tag{56}$$

$$\mathcal{Z}_j^{\omega} \;\;=\;\; \mathcal{Y}_j^{\omega}, \tag{57}$$

where for the j-th sigma point, $\mathcal{X}_j^{\mathbf{e}}$ is its chart point part and $\mathcal{X}_j^q$ is the quaternion with which it is mapped, $\mathcal{X}_j^{\omega}$ is its angular velocity part, $\mathcal{X}_j^{q^{\omega}}$ is its angular velocity noise part, $\mathcal{Y}_j^{\omega}$ is its angular velocity prediction, $\mathcal{Y}_j^q$ is the quaternion part of its prediction (we have assumed that the angular velocity $\mathcal{Y}_j^{\omega}$ is constant in the time interval $[t_{n\text{-}1}, t_n)$ so that we can use (A20)), $\mathcal{X}_j^{\mathbf{v}}$ is the vector process noise part, $\mathcal{Z}_j^{\mathbf{v}}$ is its vector measurement prediction, $\mathcal{Z}_j^{\omega}$ is its angular velocity measurement prediction, and $\Delta t_n = t_n - t_{n\text{-}1}$. Note that when applying the inverse chart φ^{-1} we will need to "saturate" $\mathcal{X}_j^{\mathbf{e}}$ to the closest point in the image of φ. Having these new sigma points, we can

obtain the means and covariance matrices of the distributions present in the UKF. First, defining $\mathcal{Z}_j := \left(\mathcal{Z}_j^{\text{v}}, \mathcal{Z}_j^{\omega} \right)^T$, the means are computed through

$$\bar{q}_{n|n\text{-}1} \quad = \quad \frac{\sum_j W_j \, \mathcal{Y}_j^q}{\| \sum_j W_j \, \mathcal{Y}_j^q \|}, \tag{58}$$

$$\bar{\omega}'_{n|n\text{-}1} \quad = \quad \sum_j W_j \, \mathcal{Y}_j^{\omega}, \tag{59}$$

$$\bar{\mathbf{x}}^{\bar{q}_{n|n\text{-}1}}_{n|n\text{-}1} \quad = \quad \begin{pmatrix} \varphi_{\bar{q}_{n|n\text{-}1}}\left(\bar{q}_{n|n\text{-}1}\right) = \mathbf{0} \\ \bar{\omega}'_{n|n\text{-}1} \end{pmatrix}, \tag{60}$$

$$\bar{\mathbf{z}}_{n|n\text{-}1} \quad = \quad \sum_j W_j \, \mathcal{Z}_j. \tag{61}$$

where we have used a variation of the result provided in Ref. [16]. Namely,

$$\bar{q} \quad \approx \quad \frac{\sum_j q_j}{\| \sum_j q_j \|}, \tag{62}$$

with $q_j \cdot q_k > 0$ for $j, k = 0, \ldots, 2N$. This result is shown to minimize the fourth order approximation of the distance defined as the sum of squared angles between the rotation transformation represented by each quaternion q_j, and the one represented by $\bar{q}$. This approach to compute the mean quaternion is extremely efficient, and its derivation is elegant and simple. In order to ensure that $q_j \cdot q_k > 0$, it is useful to remember the property that both q and $-q$ represent the same rotation. This property is also useful for introducing the quaternions in the domain of φ to execute the next step of the filter.

After this, we use the obtained mean quaternion $\bar{q}_{n|n\text{-}1}$ to express each sigma point in the $\bar{q}_{n|n\text{-}1}$-centered chart, and compute the covariance matrices:

$$\mathcal{Y}_j^{\text{e}} \quad = \quad \varphi_{\bar{q}_{n|n\text{-}1}}\left(\mathcal{Y}_j^q \right) \quad = \quad \varphi\left(\bar{q}^*_{n|n\text{-}1} * \mathcal{Y}_j^q \right), \tag{63}$$

$$\mathbf{P}^{\bar{q}_{n|n\text{-}1}}_{n|n\text{-}1} \quad = \quad \sum_j W_j \, \mathcal{Y}_j \, \mathcal{Y}_j^T, \tag{64}$$

$$\mathbf{P}^{yz}_{n|n\text{-}1} \quad = \quad \sum_j W_j \, \mathcal{Y}_j \left(\mathcal{Z}_j - \bar{\mathbf{z}}_{n|n\text{-}1} \right)^T, \tag{65}$$

$$\mathbf{S}_{n|n\text{-}1} \quad = \quad \sum_j W_j \left(\mathcal{Z}_j - \bar{\mathbf{z}}_{n|n\text{-}1} \right) \left(\mathcal{Z}_j - \bar{\mathbf{z}}_{n|n\text{-}1} \right)^T + \begin{pmatrix} \mathbf{R}_n^{\text{v}} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_n^{\omega} \end{pmatrix}, \tag{66}$$

where we have denoted $\mathcal{Y}_j := \left(\mathcal{Y}_j^{\text{e}}, \mathcal{Y}_j^{\omega} - \bar{\omega}'_{n|n\text{-}1} \right)^T$. Finally, we compute the UKF version of the Kalman gain $\mathbf{K}_n$, and we use it to obtain the optimal estimation of the state:

$$\mathbf{K}_n \quad = \quad \mathbf{P}^{yz}_{n|n\text{-}1} \, \mathbf{S}^{-1}_{n|n\text{-}1}, \tag{67}$$

$$\bar{\mathbf{x}}^{\bar{q}_{n|n\text{-}1}}_{n|n} \quad = \quad \bar{\mathbf{x}}^{\bar{q}_{n|n\text{-}1}}_{n|n\text{-}1} + \mathbf{K}_n \left(\mathbf{z}_n - \bar{\mathbf{z}}_{n|n\text{-}1} \right), \tag{68}$$

$$\mathbf{P}^{\bar{q}_{n|n\text{-}1}}_{n|n} \quad = \quad \mathbf{P}^{\bar{q}_{n|n\text{-}1}}_{n|n\text{-}1} - \mathbf{K}_n \, \mathbf{S}_{n|n\text{-}1} \, \mathbf{K}_n^T, \tag{69}$$

arriving at the same conditions in which we began the iteration, with a distribution expressed in the $\bar{q}_{n|n\text{-}1}$-centered chart, and encoded by the four

$$\left(\varphi, \bar{q}_{n|n\text{-}1}, \bar{\mathbf{x}}^{\bar{q}_{n|n\text{-}1}}_{n|n}, \mathbf{P}^{\bar{q}_{n|n\text{-}1}}_{n|n} \right). \tag{70}$$

Our best estimation for the orientation at this time is

$$\bar{q}_{n|n} \quad = \quad \varphi^{-1}_{\bar{q}_{n|n-1}}\left(\bar{\mathbf{e}}^{\bar{q}_{n|n-1}}_{n|n}\right) \quad = \quad \bar{q}_{n|n-1} * \varphi^{-1}\left(\bar{\mathbf{e}}^{\bar{q}_{n|n-1}}_{n|n}\right), \tag{71}$$

being $\bar{\mathbf{e}}^{\bar{q}_{n|n-1}}_{n|n}$ the part of the mean $\bar{\mathbf{x}}^{\bar{q}_{n|n-1}}_{n|n}$ that represent the quaternion in the $\bar{q}_{n|n-1}$-centered chart.

Note that setting $\bar{q}_{n-1|n-2} := \bar{q}_{n-1|n-1}$ and $\bar{\mathbf{e}}^{\bar{q}_{n-1|n-2}}_{n-1|n-1} := \mathbf{0}$ at the beginning of each iteration yields the traditional version of the algorithm, where a "reset operation" is performed instead of the covariance matrix update.

4. Simulation Results

This section presents the results of the simulations used to measure the accuracy of each estimator. Simulations are chosen instead of real experiments because a real system entails an uncertainty in the measurement of the true attitude: the attitude that is used to compare with that estimated by the algorithms. There are sources of error ranging from a miscalibration of the measurement system to a possible bias in the "true attitude" produced by another attitude estimator, which makes it problematic to define an adequate metric to measure the accuracy of the algorithms. For this reason, the authors consider that using a simulation is more reliable to avoid possible biases in the results due to said sources of error. Others have performed similar types of tests [7,17]. However, the results do not seem to be statistically conclusive: only the estimations of some orientation trajectories are shown.

We perform our comparison through a simulation in which we do have an absolute knowledge of the attitude of the system: a true oracle exists in a simulation. Therefore, we can compare the real orientation with the attitude estimated by the algorithms having fed them only with simulated measurements that we obtain from such known orientations. We will extract our performance metrics from a wide set of orientation trajectories in order to obtain statistically conclusive results.

We try to answer three questions with the simulation test. The first question is, is there a chart for which we get a greater accuracy in attitude estimation? The second one is, what algorithm produces the most accurate attitude estimation, the MEKF or the MUKF? The last question stems from the fact that previous algorithms on attitude estimation, such as the Multiplicative Extended Kalman Filter, did not contemplate updating the distribution from one chart to another as done at (47b) in the MEKF. However, their estimators performed well [6,7,12]. Then the third question is, does this "chart update" imply an improvement in the accuracy of the attitude estimation?

4.1. Performance Metric

We have already described a quaternion q as a deviation from another quaternion $\bar{q}$ as $q = \bar{q} * \delta$. Now we define the instantaneous error between an estimated attitude, represented by a unit quaternion $\bar{q}$, and the real attitude, represented by the unit quaternion $\overset{\star}{q}$, as the angle we have to rotate one of them to transform it into the other. This is, the angle of the rotation transformation defined by the quaternion δ_e such that $\overset{\star}{q} = \bar{q} * \delta_e$. Recalling (6), this angle can be computed as:

$$\theta_e \quad = \quad 2 \arccos\left[\left(\bar{q}^* * \overset{\star}{q}\right)_0\right] \quad = \tag{72a}$$

$$= \quad 2 \arccos\left(\bar{q} \cdot \overset{\star}{q}\right), \tag{72b}$$

having previously ensured that $\bar{q} \cdot \overset{\star}{q} \geq 0$ using the fact that both q and $-q$ represent the same rotation transformation.

Angle θ_e will vary along an orientation trajectory. Then, we will define the mean error in orientation estimation for a given trajectory starting at time $t = 0$ and ending at time $t = T$ as

$$e_\theta \quad = \quad \frac{1}{T}\int_0^T \theta_e(t)\, dt. \tag{73}$$

Finally, e_θ will depend on the followed trajectory, and on the set of taken measurements. We will need to generate several orientation trajectories to obtain the mean value $\bar{e}_\theta$ and the variance $\sigma^2_{\bar{e}_\theta}$ that characterize the distribution of the error in orientation estimation e_θ for each algorithm. We will define the confidence interval for the computed $\bar{e}_\theta$ as

$$\left[\bar{e}_\theta - 3\,\sigma_{\bar{e}_\theta}/\sqrt{N_s}\,,\, \bar{e}_\theta + 3\,\sigma_{\bar{e}_\theta}/\sqrt{N_s} \right], \tag{74}$$

where N_s is the number of samples taken for the $\bar{e}_\theta$ computation, so that $\sigma^2_{\bar{e}_\theta}/N_s$ is the variance of the sample mean distribution.

Being that the lower the better, the value of $\bar{e}_\theta$ gives us a measure of how well an algorithm estimates the orientation. We will consider that the performance of an algorithm A is better than the performance of other algorithm B if $\bar{e}_\theta(A) < \bar{e}_\theta(B)$ and their confidence intervals do not overlap.

4.2. Simulation Scheme

To compute the performance metrics we will need to generate a large number of simulations. Each independent simulation will consist of three steps: initialization, convergence, and estimation.

In the initialization step we set up the initial conditions accordingly to the chosen simulation parameters. This includes generating the initial unit quaternion $\overset{\star}{\bar{q}}_0$ from a uniform distribution in S^3, setting the initial angular velocity $\overset{\star}{\omega}'_0$ to zero, setting the update frequency f_{update}, generating the variances of the process noises σ^2_ω and σ^2_v from a uniform distribution in the intervals $(0, Q^\omega_{\text{max}}]$ and $(0, Q^v_{\text{max}}]$ respectively, and initializing the estimation algorithm. The initialization of the MEKF includes setting $\bar{q}_{0|0} = 1$, $\overline{\omega}'_{0|0} = 0\,\text{rad/s}$, and $\mathbf{P}^{\bar{q}_{0|0}}_{0|0} = 10^2\,\mathbf{I}$. On the other hand, the initialization of the MUKF includes setting $\bar{q}_{0|-1} = 1$, $\mathbf{e}^{\bar{q}_{0|-1}}_{0|0} = \mathbf{0}$, $\overline{\omega}'_{0|0} = (1,1,1)^T\,\text{rad/s}$, and $\mathbf{P}^{\bar{q}_{0|-1}}_{0|0} = 10^2\,\mathbf{I}$. The angular velocity is not initialized to $\mathbf{0}$ in the MUKF because it has been observed that it is sometimes necessary to "break the symmetry" for the algorithm to converge; especially when we do not apply the chart update (when we perform the "reset operation") for the RV chart. The covariance matrices that appear in both algorithms are initialized as $\mathbf{Q}^\omega_n = \mathbf{I}\,\text{rads}^2/\text{s}^4$, $\mathbf{Q}^v_n = 10^{-2}\,\mathbf{I}$ p.d.u. ("p.d.u." stands for "Procedure Defined Unit". In the present case it depends on the definition of the vector $\mathbf{v}$), $\mathbf{R}^\omega_n = R^\omega\,\mathbf{I}\,\text{rads}^2/\text{s}^2$, $\mathbf{R}^v_n = R^v\,\mathbf{I}$ p.d.u., where R^ω and R^v are the variances of the measurement noise that will be used in the simulation. We give this information about the measurement noise to the algorithms because it can be obtained offline, while the information about the process noise cannot. Given that a priori we cannot know how the system will behave, the values of $\mathbf{Q}^\omega_n$ and $\mathbf{Q}^v_n$ have been chosen according to what we understand could be normal. Choosing these values we are assuming that after a second it is normal for the angular velocity to have changed by $1\,\text{rad/s}$, and also that it is normal to find external noises added to the vector $\mathbf{v}_t$ of magnitude 10^{-1} p.d.u.. For the mean values we set $\overline{\mathbf{q}}^\omega_n = \mathbf{0}\,\text{rads/s}^2$, and $\overline{\mathbf{q}}^v = \mathbf{0}$ p.d.u..

In the convergence step we keep the system in the initial orientation $\overset{\star}{\bar{q}}_0$. Simulated measurements are generated using (23) and (24). For each measurement, a different $\mathbf{v}_t$ is sampled from a uniform distribution in the unit sphere of $\mathbb{R}^3$. The values for each component of $\mathbf{q}^v_t$, $\mathbf{r}^v_t$, and $\mathbf{r}^\omega_t$ are obtained from normal distributions with zero mean and variances σ^2_v, R^v, and R^ω respectively. The term $\mathbf{R}^T(\mathbf{q}_t)$ in (23) is obtained from the true attitude $\overset{\star}{\bar{q}}_t$, which in the convergence step takes the value of $\overset{\star}{\bar{q}}_t = \overset{\star}{\bar{q}}_0$. The term ω'_t in (24) is the true angular velocity, which in the convergence step takes the value $\overset{\star}{\omega}'_t = \mathbf{0}$. The tested algorithm updates its state estimation until the inequality $\theta_e(t) < \theta^0_e$ is satisfied, where $\theta_e(t)$ is the value of the error (72), and θ^0_e is a parameter in the simulation. The convergence step could have been replaced by an initialization of the attitude estimated by the algorithm $\overline{q}_t$ to the real value $\overset{\star}{\bar{q}}_t$, but then it would have also been necessary to fix a certain covariance matrix. Since the metric of the space generated by each chart is different, it is difficult to set a covariance matrix that provides the same information for each chart. It seemed more natural to the authors to allow the

algorithm to find the true attitude by its own means, and for the covariance matrix to converge to a value in each case.

Finally, in the estimation step we generate a random but continuous orientation sequence using a Wiener process for the angular velocity:

$$\overset{\star}{\boldsymbol{\omega}}'_t \;=\; \overset{\star}{\boldsymbol{\omega}}'_{t\text{-}\delta t} \;+\; \mathbf{n}_t \sqrt{\delta t}\,, \tag{75}$$

$$\overset{\star}{\boldsymbol{q}}_t \;=\; \overset{\star}{\boldsymbol{q}}_{t\text{-}\delta t} * \left(\begin{array}{c} \cos\left(\dfrac{\|\overset{\star}{\boldsymbol{\omega}}'_t\|\,\delta t}{2} \right) \\[2ex] \dfrac{\overset{\star}{\boldsymbol{\omega}}'_t}{\|\overset{\star}{\boldsymbol{\omega}}'_t\|} \, \sin\left(\dfrac{\|\overset{\star}{\boldsymbol{\omega}}'_t\|\,\delta t}{2} \right) \end{array} \right), \tag{76}$$

where $\mathbf{n}_t$ is a random vector whose components are sampled from a normal distribution with zero mean and variance σ_{ω}^2, and δt is the simulation time step that is related to the algorithm time step Δt trough `dtdtsim` $\delta t = \Delta t$, being `dtdtsim` an integer parameter that determines the simulation updates per algorithm update. Note that we multiply $\mathbf{n}_t$ by $\sqrt{\delta t}$ and not by δt. We do it this way so that the covariance matrix after k steps does not depend on the simulation time step δt. In fact, after a time $T = k\,\delta t$ the covariance matrix of the angular velocity will have grown by $\Delta \mathbf{P}^{\omega} = k\,\mathbf{I}\,\sigma_{\omega}^2\,\delta t = \mathbf{I}\,\sigma_{\omega}^2\,T$, and not by $(\Delta \mathbf{P}^{\omega})' = k\,\mathbf{I}\,\sigma_{\omega}^2\,(\delta t)^2 = \mathbf{I}\,\sigma_{\omega}^2\,T\,\delta t$. After each `dtdtsim` simulation updates, a simulated measurement is generated in the same way it was done in the convergence step, and the algorithm is updated with it. The simulation will run for a time `Tsim` $= k'\,\Delta t$, where k' is an integer number. This way we will perform the last algorithm update at the end of the simulation. The error (72) will be evaluated after each algorithm update, and it will be added up through the simulation to obtain the averaged error (73). After each simulation, we will obtain a sample for the computation of $\bar{e}_{\theta}$ and $\sigma_{\bar{e}_{\theta}}^2$. We will perform N_s of these simulations to obtain the confidence interval (74).

4.3. Results

In this section we present the results of the simulations. The algorithms are tested for update frequencies $f_{\text{update}} = 1/\Delta t$ in the interval $[2, 1000]\,\text{Hz}$. This range has been chosen thinking about the possible limitations of a real system. For example, the maximum data rate of a low cost IMU is around $1000\,\text{Hz}$. On the other hand, the update frequency may be limited by processing. The computational cost of each estimator has been evaluated in two platforms: an Arduino MEGA 2560, and a Raspberry Pi 3 Model B. The code has been written in c++. The resulting maximum update frequencies are presented in Figure 4, which indicates that the MEKF can be executed approximately 3 times faster than the MUKF.
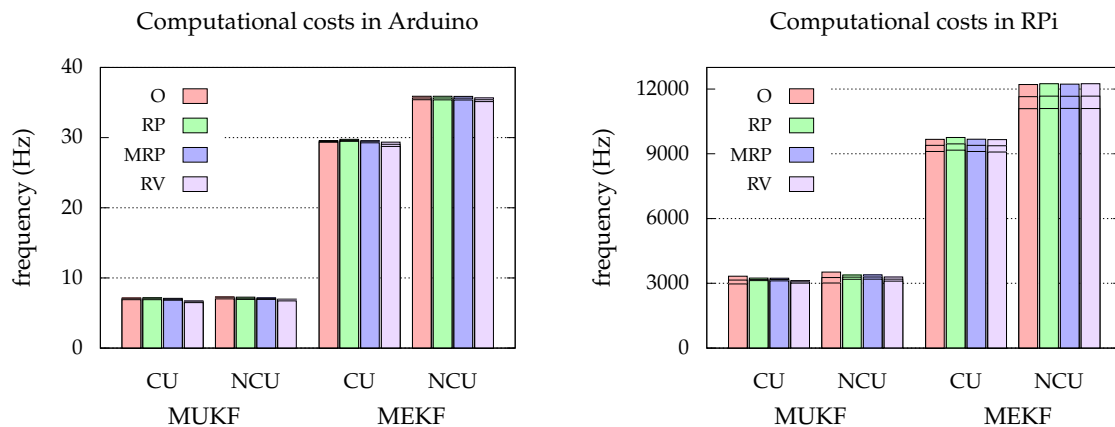
Figure 4. Maximum update frequency for each approach. The lines at the top represent the mean and the deviation (3σ) of the distribution of maximum update frequencies. "CU" stands for Chart Update, while "NCU" stands for No Chart Update.

Although the algorithms have been developed allowing a different Δt_n for each update, the simulations are performed using a constant Δt, and the simulation parameters depicted in Table 5.

Table 5. Parameters used in the simulations.

Parameter	Value
θ_e^0	$1°$
`Tsim`	$10\,\text{s}$
`dtdtsim`	100
N_s	1000
$Q_{\max}^{\omega}$	$10^2\,\text{rads}^2/\text{s}^3$
$Q_{\max}^{\mathbf{v}}$	$1\,\text{p.d.u.}$
R	$\{\,10^{-2}, 10^{-4}, 10^{-6}\,\}$
R^{ω}	$R\,\text{rads}^2/\text{s}^2$
$R^{\mathbf{v}}$	$R\,\text{p.d.u.}$
W_0	$1/25$

The parameters θ_e^0, `Tsim`, `dtdtsim`, and N_s have been chosen trying to reach a compromise between the precision of the results, and the execution time of the simulation. The values for $Q_{\max}^{\omega}$ and $Q_{\max}^{\mathbf{v}}$ have been chosen in such a way that the estimation algorithms face both normal situations ($\mathbf{Q}_n^{\omega} \approx \sigma_{\omega}^2\,\mathbf{I}$ and $\mathbf{Q}_n^{\mathbf{v}} \approx \sigma_{\mathbf{v}}^2\,\mathbf{I}$) and situations that were not foreseen ($\mathbf{Q}_n^{\omega} \neq \sigma_{\omega}^2\,\mathbf{I}$ or $\mathbf{Q}_n^{\mathbf{v}} \neq \sigma_{\mathbf{v}}^2\,\mathbf{I}$). A typical low cost IMU has $R^{\omega} \approx 10^{-4}\,\text{rad}^2/\text{s}^2$ and $R^{\mathbf{v}} \approx 10^{-4}\,g^2$. The values chosen for R represent an imprecise sensor (10^{-2}), a normal sensor (10^{-4}), and a precise sensor (10^{-6}). The value of W_0 has been chosen so that all sigma points have the same importance, but very similar results, if not identical, have been obtained for other selections of W_0.

4.3.1. Chart Choice

The results of the simulation are presented in Figure 5. The average of the performance metric is shown along with its confidence interval for each of the selected update frequencies. The results of the MEKF and the MUKF are shown in different graphs, but drawn in the same one are the results for each chart and for a given MKF. In this way we are able to distinguish if a chart has an advantage over the others.

We observe that there is no chart that is especially advantageous. All things being equal, we would opt for the RP chart. For this chart it is not necessary to worry about the domain since it maps $\boldsymbol{q}$ and $-\boldsymbol{q}$ with the same point of $\mathbb{R}^3$ and with the same $\mathbf{T}$-matrix; or of the image since it is all $\mathbb{R}^3$. In addition, the expressions of φ^{-1} and the $\mathbf{T}$-matrix for the MEKF are simpler for the RP chart. These computational advantages make us prefer the RP chart over the others.

4.3.2. MEKF vs. MUKF

Figure 6 also presents the results of the simulations. This time, we display on the same graph the resulting performance metrics for the MUKF and the MEKF when the RP chart is used. In this way, we can distinguish if one MKF has an advantage over the other.

We note that the MEKF performs the same or better than the MUKF. This differs from the usual experience, in which the UKF outperforms the EKF in traditional non-linear estimation applications. The fact that the charts resemble the Euclidean space near the origin (see Section 2.3) might be favoring the MEKF, since the Jacobian matrices, used to approximate the non-linear functions, are defined at that point. However, the sigma points generated for the MUKF are sampled far from the origin of the chart, where the non-linearities become notorious. We are facing a very particular scenario in which the model is approximately linear for the MEKF, while for the MUKF it is not. In addition, due to the difference in computational cost (see Figure 4), the MUKF update frequencies will generally be lower
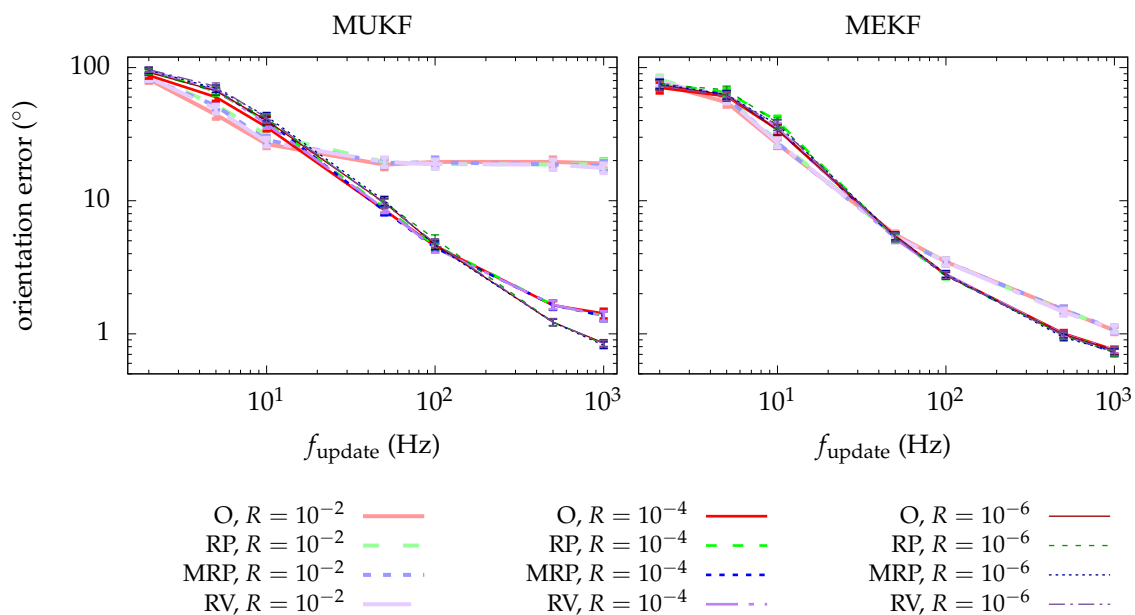
Figure 5. Mean of the performance metric for each approach. Results from different charts are plotted in the same graph. Results from different MKF are plotted in different graphs. Bars represent the confidence interval (3σ) for the mean computation.
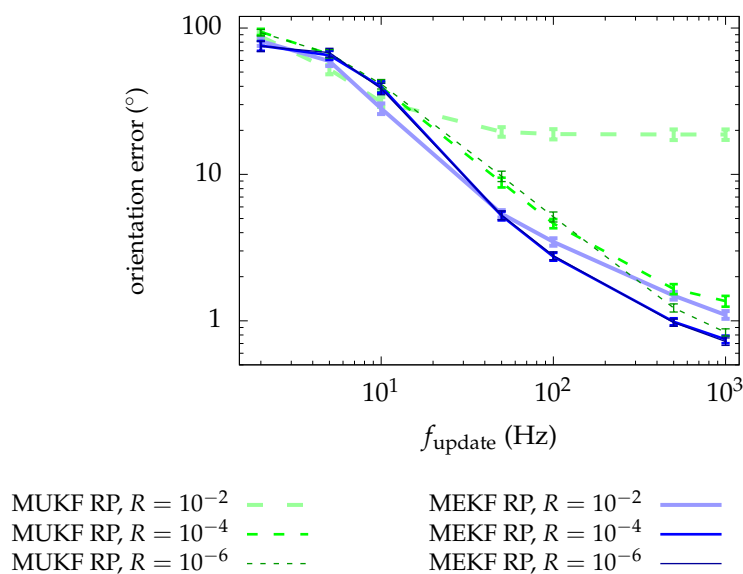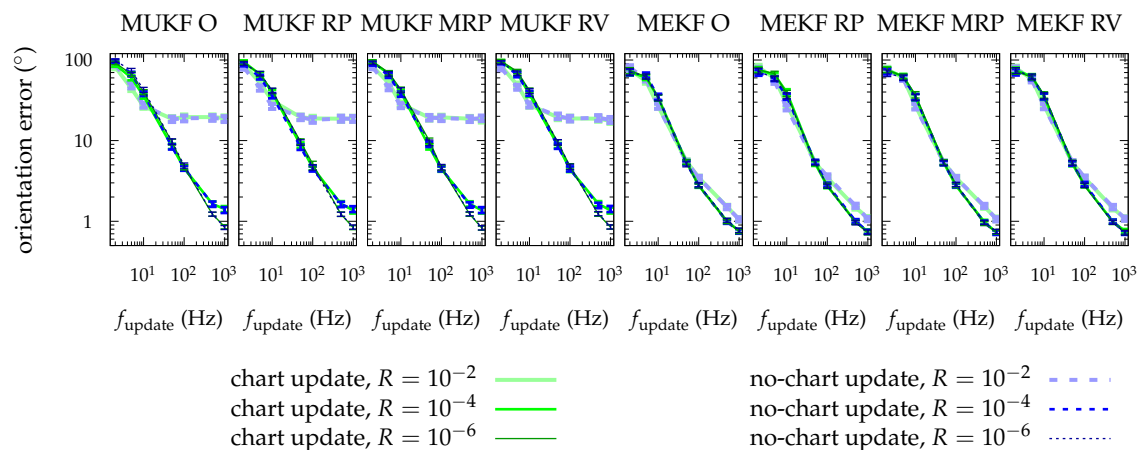

Figure 6. Mean of the performance metric for each MKF. Only results for the RP chart are plotted. Bars represent the confidence interval (3σ) in the mean computation.

than those of the MEKF, which will imply worse accuracy in its estimations. Then, the MEKF with the RP chart seems to be our best option.

4.3.3. Chart Update vs. No Chart Update

Figure 7 presents the results of each MKF with each chart in a different graph, but displayed in the same one are the results using the "chart update" and the results without using it.

We can observe that there is almost no difference between using the "chart update" and not using it. The concepts used in this paper have helped us to understand the mechanisms of the MKF, and

Figure 7. Mean of the performance metric for each approach. Results from the approach in which we apply the chart update, and those of which we do not apply it are plotted together. Bars represent the confidence interval (3σ) in the mean computation.

ultimately to arrive to the concepts of "multiplicative update", and of "covariance correction step" with the **T**-matrix definition. However, it is not necessary to apply the latest update (47b) in practice: we will obtain essentially the same accuracy in our estimations.

5. Conclusions

We have used concepts from manifold theory to define the expected value and the covariance matrix of a distribution in a manifold. In particular, we have defined the expected value and covariance matrix of a distribution of unit quaternions in S^3, the unit sphere in $\mathbb{R}^4$, using the concept of chart. These definitions have helped us to develop Kalman filters for orientation estimation, where the attitude has been represented by a unit quaternion. They have also helped us solve the problem of the "covariance correction step". Two estimators have been developed: one based on the EKF (the MEKF), and another based on the UKF (the MUKF). The MEKF and the MUKF have been tested in simulations, and some results have been obtained. The conclusions of the simulations are:

- There is no chart that presents a clear advantage over the others, but the RP chart has some characteristics that motivate us to prefer it.
- The MEKF is preferable to the MUKF due to its lower computational cost and its greater accuracy in orientation estimation.
- The "chart update" is not necessary for the MKF in practice.

Then, the MEKF with the RP chart and without applying the "chart update" is our best attitude estimator according to the adopted performance metric. This algorithm resembles the conventional "Multiplicative Extended Kalman Filter", but we have obtained the MEKF without having to redefine any aspect of the classic Kalman filter.

Abbreviations

The following abbreviations are used in this manuscript:

EKF Extended Kalman Filter
UKF Unscented Kalman Filter
MKF Manifold Kalman Filter
MEKF Manifold Extended Kalman Filter
MUKF Manifold Unscented Kalman Filter
O   Orthographic
RP   Rodrigues Parameters
MRP Modified Rodrigues Parameters
RV   Rotation Vector

Appendix A. Derivation of Transition Maps

This appendix contains the derivation of the transition map for each chart.

Appendix A.1. Orthographic

Using the inverse of the transformation that defines the chart, φ^{-1},

$$\delta^{\overline{q}} \;=\; \begin{pmatrix} \sqrt{1 - \frac{\|\mathbf{e}^{\overline{q}}\|^2}{4}} \\ \mathbf{e}^{\overline{q}}/2 \end{pmatrix}. \tag{A1}$$

Introducing (A1) into (19),

$$\delta^{\overline{p}} \;=\; \overline{\delta}^* * \delta^{\overline{q}} \;=\; \begin{pmatrix} \overline{\delta}_0 \sqrt{1 - \frac{\|\mathbf{e}^{\overline{q}}\|^2}{4}} + \overline{\boldsymbol{\delta}} \cdot \frac{\mathbf{e}^{\overline{q}}}{2} \\ \overline{\delta}_0 \frac{\mathbf{e}^{\overline{q}}}{2} - \sqrt{1 - \frac{\|\mathbf{e}^{\overline{q}}\|^2}{4}}\, \overline{\boldsymbol{\delta}} - \overline{\boldsymbol{\delta}} \times \frac{\mathbf{e}^{\overline{q}}}{2} \end{pmatrix}. \tag{A2}$$

Finally, applying the chart definition,

$$\mathbf{e}^{\overline{p}} \;=\; 2\,\boldsymbol{\delta}^{\overline{p}} \;=\; \overline{\delta}_0\, \mathbf{e}^{\overline{q}} - \sqrt{4 - \|\mathbf{e}^{\overline{q}}\|^2}\, \overline{\boldsymbol{\delta}} - \overline{\boldsymbol{\delta}} \times \mathbf{e}^{\overline{q}}. \tag{A3}$$

Appendix A.2. Rodrigues Parameters

Using the inverse of the transformation that defines the chart, φ^{-1},

$$\delta^{\overline{q}} \;=\; \frac{1}{\sqrt{4 + \|\mathbf{e}^{\overline{q}}\|^2}} \begin{pmatrix} 2 \\ \mathbf{e}^{\overline{q}} \end{pmatrix}. \tag{A4}$$

Introducing (A4) into (19),

$$\delta^{\overline{p}} \;=\; \overline{\delta}^* * \delta^{\overline{q}} \;=\; \frac{1}{\sqrt{4 + \|\mathbf{e}^{\overline{q}}\|^2}} \begin{pmatrix} 2\,\overline{\delta}_0 + \overline{\boldsymbol{\delta}} \cdot \mathbf{e}^{\overline{q}} \\ \overline{\delta}_0\, \mathbf{e}^{\overline{q}} - 2\,\overline{\boldsymbol{\delta}} - \overline{\boldsymbol{\delta}} \times \mathbf{e}^{\overline{q}} \end{pmatrix}. \tag{A5}$$

Finally, applying the chart definition,

$$\mathbf{e}^{\overline{p}} \;=\; 2\,\frac{\boldsymbol{\delta}^{\overline{p}}}{\delta_0^{\overline{p}}} \;=\; 2\,\frac{\overline{\delta}_0\, \mathbf{e}^{\overline{q}} - 2\,\overline{\boldsymbol{\delta}} - \overline{\boldsymbol{\delta}} \times \mathbf{e}^{\overline{q}}}{2\,\overline{\delta}_0 + \overline{\boldsymbol{\delta}} \cdot \mathbf{e}^{\overline{q}}}. \tag{A6}$$

Appendix A.3. Modified Rodrigues Parameters

Using the inverse of the transformation that defines the chart, φ^{-1},

$$\delta^{\overline{q}} = \frac{1}{16 + \|\mathbf{e}^{\overline{q}}\|^2} \begin{pmatrix} 16 - \|\mathbf{e}^{\overline{q}}\|^2 \\ 8\,\mathbf{e}^{\overline{q}} \end{pmatrix}. \tag{A7}$$

Introducing (A7) into (19),

$$\delta^{\overline{p}} = \overline{\delta}^{*} * \delta^{\overline{q}} = \frac{1}{16 + \|\mathbf{e}^{\overline{q}}\|^2} \begin{pmatrix} \overline{\delta}_0 \left(16 - \|\mathbf{e}^{\overline{q}}\|^2\right) + 8\,\overline{\boldsymbol{\delta}} \cdot \mathbf{e}^{\overline{q}} \\ 8\,\overline{\delta}_0\,\mathbf{e}^{\overline{q}} - \left(16 - \|\mathbf{e}^{\overline{q}}\|^2\right)\overline{\boldsymbol{\delta}} - 8\,\overline{\boldsymbol{\delta}} \times \mathbf{e}^{\overline{q}} \end{pmatrix}. \tag{A8}$$

Finally, applying the chart definition,

$$\mathbf{e}^{\overline{p}} = 4\frac{\boldsymbol{\delta}^{\overline{p}}}{1 + \delta_0^{\overline{p}}} = 4\frac{8\,\overline{\delta}_0\,\mathbf{e}^{\overline{q}} - \left(16 - \|\mathbf{e}^{\overline{q}}\|^2\right)\overline{\boldsymbol{\delta}} - 8\,\overline{\boldsymbol{\delta}} \times \mathbf{e}^{\overline{q}}}{16 + \|\mathbf{e}^{\overline{q}}\|^2 + \overline{\delta}_0\left(16 - \|\mathbf{e}^{\overline{q}}\|^2\right) + 8\,\overline{\boldsymbol{\delta}} \cdot \mathbf{e}^{\overline{q}}}. \tag{A9}$$

Appendix A.4. Rotation Vector

Using the inverse of the transformation that defines the chart, φ^{-1},

$$\delta^{\overline{q}} = \begin{pmatrix} \cos\left(\frac{\|\mathbf{e}^{\overline{q}}\|}{2}\right) \\ \widehat{\mathbf{e}}^{\overline{q}}\,\sin\left(\frac{\|\mathbf{e}^{\overline{q}}\|}{2}\right) \end{pmatrix}. \tag{A10}$$

Introducing (A10) into (19),

$$\delta^{\overline{p}} = \overline{\delta}^{*} * \delta^{\overline{q}} = \begin{pmatrix} \overline{\delta}_0 \cos\left(\frac{\|\mathbf{e}^{\overline{q}}\|}{2}\right) + \overline{\boldsymbol{\delta}} \cdot \widehat{\mathbf{e}}^{\overline{q}}\,\sin\left(\frac{\|\mathbf{e}^{\overline{q}}\|}{2}\right) \\ \overline{\delta}_0\,\widehat{\mathbf{e}}^{\overline{q}}\,\sin\left(\frac{\|\mathbf{e}^{\overline{q}}\|}{2}\right) - \cos\left(\frac{\|\mathbf{e}^{\overline{q}}\|}{2}\right)\overline{\boldsymbol{\delta}} - \overline{\boldsymbol{\delta}} \times \widehat{\mathbf{e}}^{\overline{q}}\,\sin\left(\frac{\|\mathbf{e}^{\overline{q}}\|}{2}\right) \end{pmatrix}. \tag{A11}$$

Finally, applying the chart definition,

$$\mathbf{e}^{\overline{p}} = 2\frac{\boldsymbol{\delta}^{\overline{p}}}{\|\boldsymbol{\delta}^{\overline{p}}\|}\,\arcsin\|\boldsymbol{\delta}^{\overline{p}}\|, \tag{A12}$$

with

$$\boldsymbol{\delta}^{\overline{p}} = \overline{\delta}_0\,\widehat{\mathbf{e}}^{\overline{q}}\,\sin\left(\frac{\|\mathbf{e}^{\overline{q}}\|}{2}\right) - \cos\left(\frac{\|\mathbf{e}^{\overline{q}}\|}{2}\right)\overline{\boldsymbol{\delta}} - \overline{\boldsymbol{\delta}} \times \widehat{\mathbf{e}}^{\overline{q}}\,\sin\left(\frac{\|\mathbf{e}^{\overline{q}}\|}{2}\right). \tag{A13}$$

Note that all transition maps are expressed using the $\overline{\delta}$ quaternion. Given that $\overline{\mathbf{e}} = \varphi(\overline{\delta})$ we could also have expressed them using $\overline{\mathbf{e}}$ which is what we get after applying the Kalman update (41). However, our choice makes transition maps to take a simpler form. In addition, having to compute the quaternion $\overline{\delta}$ to perform (43c), this choice does not imply a computational overhead.

Appendix B. Details in the Derivation of the MEKF

This appendix contains the details in the derivation of the Manifold Extended Kalman Filter used in this study.

Appendix B.1. State Prediction

This subsection contains the derivation of the equations for the state prediction.

Appendix B.1.1. Evolution of the Expected Value of the State

Taking expected values in Equation (21) we obtain

$$\frac{d\,\overline{\omega}'}{dt} \;=\; \overline{\mathbf{q}}^{\omega} \qquad \Longrightarrow \qquad \overline{\omega}'(t) \;=\; \overline{\omega}'_0 + \overline{\mathbf{q}}^{\omega}_t\, t, \tag{A14}$$

with $\overline{\mathbf{q}}^{\omega}_t$ the expected value of the random variable $\mathbf{q}^{\omega}$ at time t. Doing $\overline{\omega}'_0 = \overline{\omega}'_{n\text{-}1|n\text{-}1}$ we arrive at

$$\overline{\omega}'_{t|n\text{-}1} \;=\; \overline{\omega}'_{n\text{-}1|n\text{-}1} + \overline{\mathbf{q}}^{\omega}_t\, t. \tag{A15}$$

On the other hand, approximating (22) with its Taylor series up to first order around the current state $(\overline{q}, \overline{\omega}')$, and taking its expected value we obtain

$$\mathrm{E}\left[\frac{d\,q(t)}{dt}\right] \;\approx\; \frac{1}{2}\,\overline{q}(t) * \overline{\omega}'(t). \tag{A16}$$

This differential equation has no general closed solution. But if we assume that the expected value of the process noise $\overline{\mathbf{q}}^{\omega}(t)$ is zero when $t \in (t_{n\text{-}1}, t_n)$, so that $\overline{\omega}'(t)$ is constant in that interval, then we will have the matrix differential equation

$$\frac{d\,\overline{q}(t)}{dt} \;=\; \check{\Omega}_n\,\overline{q}(t), \tag{A17}$$

with

$$\check{\Omega}_n \;:=\; \frac{1}{2}\left(\begin{array}{cccc} 0 & -\overline{\omega}'_1 & -\overline{\omega}'_2 & -\overline{\omega}'_3 \\ \overline{\omega}'_1 & 0 & \overline{\omega}'_3 & -\overline{\omega}'_2 \\ \overline{\omega}'_2 & -\overline{\omega}'_3 & 0 & \overline{\omega}'_1 \\ \overline{\omega}'_3 & \overline{\omega}'_2 & -\overline{\omega}'_1 & 0 \end{array}\right)_{n|n\text{-}1}. \tag{A18}$$

This differential equation has the solution

$$\overline{q}(t) \;=\; e^{\check{\Omega}\,t}\,\overline{q}_0, \tag{A19}$$

where $\overline{q}_0$ represents the initial conditions. After taking $\overline{q}_0 = \overline{q}_{n\text{-}1|n\text{-}1}$, we obtain the prediction $\overline{q}_{t|n\text{-}1}$, that can be expressed using the quaternion product as

$$\overline{q}_{t|n\text{-}1} \;=\; \overline{q}_{n\text{-}1|n\text{-}1} * \delta^{\omega}_t \;=\; \overline{q}_{n\text{-}1|n\text{-}1} * \left(\begin{array}{c} \cos\left(\dfrac{\|\overline{\omega}'_{t|n\text{-}1}\|\,\Delta t}{2}\right) \\[2ex] \dfrac{\overline{\omega}'_{t|n\text{-}1}}{\|\overline{\omega}'_{t|n\text{-}1}\|}\,\sin\left(\dfrac{\|\overline{\omega}'_{t|n\text{-}1}\|\,\Delta t}{2}\right) \end{array}\right), \tag{A20}$$

with $\Delta t = t - t_{n\text{-}1}$.

Appendix B.1.2. Evolution of the State Covariance Matrix

For a continuous nonlinear system of the form

$$\frac{d\,\mathbf{x}}{dt} \;=\; \mathbf{f}(\mathbf{x},t) + \mathbf{g}(\mathbf{q}), \tag{A21}$$

we know [18] that the covariance matrix satisfies the following differential equation:

$$\frac{d\,\mathbf{P}}{dt} \;=\; \mathbf{F}\mathbf{P} + \mathbf{P}\mathbf{F}^T + \mathbf{G}\mathbf{Q}\mathbf{G}^T, \tag{A22}$$

where $\mathbf{F} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}$, and $\mathbf{G} = \frac{\partial \mathbf{g}}{\partial \mathbf{q}}$. This is so because the evolution equation for $\Delta \mathbf{x} = \mathbf{x} - \bar{\mathbf{x}}$ is approximately given by:

$$\frac{d\mathbf{x}}{dt} \approx \frac{d\bar{\mathbf{x}}}{dt} + \left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\bar{\mathbf{x}}} (\mathbf{x} - \bar{\mathbf{x}}) + \left.\frac{\partial \mathbf{g}}{\partial \mathbf{q}}\right|_{\mathbf{q}=\bar{\mathbf{q}}} (\mathbf{q} - \bar{\mathbf{q}}), \tag{A23}$$

and $\mathbf{P}$ is defined as $\mathbf{P} = \mathrm{E}\left[\Delta\mathbf{x}\,(\Delta\mathbf{x})^T\right]$. However, we have a different definition for $\mathbf{P}$:

$$\mathbf{P} = \mathrm{E}\left[\begin{pmatrix} \mathbf{e}^{\bar{q}} - \bar{\mathbf{e}}^{\bar{q}} \\ \omega' - \bar{\omega}' \end{pmatrix}\begin{pmatrix} \mathbf{e}^{\bar{q}} - \bar{\mathbf{e}}^{\bar{q}} \\ \omega' - \bar{\omega}' \end{pmatrix}^T\right]. \tag{A24}$$

Then we need to find the evolution equation for $\mathbf{e}^{\bar{q}}$. Recall that we are assuming $\bar{\mathbf{e}}^{\bar{q}} = \mathbf{0}$ at the beginning of the iteration. Knowing that any quaternion in the unit sphere can be expressed as a deviation from a central quaternion $\bar{q}$ as $q = \bar{q} * \delta$, and using the differential Equations (22) and (A16), we can find a differential equation for the quaternion δ:

$$q = \bar{q} * \delta \implies \tag{A25a}$$

$$\implies \dot{q} = \dot{\bar{q}} * \delta + \bar{q} * \dot{\delta} \implies \tag{A25b}$$

$$\implies \frac{1}{2} q * \omega' \approx \frac{1}{2} \bar{q} * \bar{\omega}' * \delta + \bar{q} * \dot{\delta}, \tag{A25c}$$

where a dot over a symbol represents time derivative, and we have obviated the time dependence. Isolating the time derivative $\dot{\delta}$,

$$\dot{\delta} \approx \frac{1}{2} \overbrace{\bar{q}^* * q}^{\delta} * \omega' - \frac{1}{2} \overbrace{\bar{q}^* * \bar{q}}^{1} * \bar{\omega}' * \delta = \tag{A26a}$$

$$= \frac{1}{2}\left[\delta * \omega' - \bar{\omega}' * \delta\right] = \tag{A26b}$$

$$= \frac{1}{2}\left[\begin{pmatrix} \delta_0 \\ \delta \end{pmatrix} * \begin{pmatrix} 0 \\ \omega' \end{pmatrix} - \begin{pmatrix} 0 \\ \bar{\omega}' \end{pmatrix} * \begin{pmatrix} \delta_0 \\ \delta \end{pmatrix}\right] = \tag{A26c}$$

$$= \frac{1}{2}\begin{pmatrix} -(\omega' - \bar{\omega}') \cdot \delta \\ \delta_0\,(\omega' - \bar{\omega}') - (\omega' + \bar{\omega}') \times \delta \end{pmatrix} = \tag{A26d}$$

$$= \frac{1}{2}\begin{pmatrix} -\Delta\omega' \cdot \delta \\ \delta_0\,\Delta\omega' - (2\bar{\omega}' + \Delta\omega') \times \delta \end{pmatrix}. \tag{A26e}$$

Knowing that, for each of our charts, the δ quaternion can be approximated by (10) as $\mathbf{e} \to \mathbf{0}$, then we can obtain an approximate differential equation for a point $\mathbf{e}$ expressed in the $\bar{q}$-centered chart. Note that we have not explicitly denoted $\mathbf{e}^{\bar{q}}$ or $\delta^{\bar{q}}$. This will be assumed implicitly, since these quantities will always be expressed in the $\bar{q}$-centered chart in this appendix. Using the chain rule for a time derivative and expression (A26e),

$$\dot{e}_i = \sum_j \overbrace{\frac{\partial e_i}{\partial \delta_j}}^{\approx 2\delta_{ij}} \overbrace{\frac{\partial \delta_j}{\partial t}}^{\dot{\delta}_j} \equiv \tag{A27a}$$

$$\equiv \dot{\mathbf{e}} \approx \delta_0\,\Delta\omega' - (2\bar{\omega}' + \Delta\omega') \times \delta \approx \tag{A27b}$$

$$\approx \left(1 - \frac{\|\mathbf{e}\|^2}{8}\right)\Delta\omega' - (2\bar{\omega}' + \Delta\omega') \times \frac{\mathbf{e}}{2}. \tag{A27c}$$

Then, the first order approximation to differential Equation (A27c) would be

$$\dot{\mathbf{e}} \quad \approx \quad \Delta\omega' - \overline{\omega}' \times \mathbf{e}. \tag{A28}$$

On the other hand, combining Equations (21) and (A14) we obtain

$$\frac{d\,\Delta\omega'}{dt} \quad = \quad \frac{d\,(\omega' - \overline{\omega}')}{dt} \quad = \quad \mathbf{q}^\omega - \overline{\mathbf{q}}^\omega \quad = \quad \Delta\mathbf{q}^\omega. \tag{A29}$$

Summarizing,

$$\frac{d}{dt} \begin{pmatrix} \mathbf{e} \\ \Delta\omega' \end{pmatrix} \quad \approx \quad \begin{pmatrix} -[\overline{\omega}']_\times & \mathbf{I} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{e} \\ \Delta\omega' \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ \Delta\mathbf{q}^\omega \end{pmatrix}, \tag{A30}$$

therefore matrices $\mathbf{F}$, $\mathbf{G}$, and $\mathbf{Q}$ in (A22) are in our case

$$\mathbf{F} \quad = \quad \begin{pmatrix} -[\overline{\omega}']_\times & \mathbf{I} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \tag{A31}$$

$$\mathbf{G} \quad = \quad \mathbf{I}, \tag{A32}$$

$$\mathbf{Q} \quad = \quad \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathrm{E}[\,\Delta\mathbf{q}^\omega\,(\Delta\mathbf{q}^\omega)^T\,] \end{pmatrix}. \tag{A33}$$

We are now in a position to solve the differential Equation (A22). Let us consider its homogeneous version first:

$$\frac{d\,\mathbf{P}_H}{dt} \quad = \quad \mathbf{F}\,\mathbf{P}_H + \mathbf{P}_H\,\mathbf{F}^T, \tag{A34}$$

which has as solution

$$\mathbf{P}_H \quad = \quad e^{\mathbf{F}\,t}\,\mathbf{C}_0\,e^{\mathbf{F}^T\,t}. \tag{A35}$$

Taking into account the definition of matrix exponential, and after computing the powers of $\mathbf{F}$ we obtain

$$e^{\mathbf{F}\,t} \quad = \quad \begin{pmatrix} \sum_{n=0}^{\infty} \frac{(-\Omega)^n\,t^n}{n!} & \sum_{n=1}^{\infty} \frac{(-\Omega)^{n-1}\,t^n}{n!} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \quad \approx \quad \begin{pmatrix} \mathbf{R}^T(\delta^\omega) & \mathbf{I}\,t \\ \mathbf{0} & \mathbf{I} \end{pmatrix}, \tag{A36}$$

where we have denoted $\Omega = [\overline{\omega}']_\times$, and $\delta^\omega = \left(\cos \frac{\|\overline{\omega}'\|\,t}{2}, \frac{\overline{\omega}'}{\|\overline{\omega}'\|} \sin \frac{\|\overline{\omega}'\|\,t}{2} \right)$. We also have assumed that t takes small values so we can approximate the infinite sums truncating in the first term. To find the solution of the non-homogeneous differential equation we use the *variation of constants method*:

$$\mathbf{P} \quad = \quad e^{\mathbf{F}\,t}\,\mathbf{C}(t)\,e^{\mathbf{F}^T\,t} \qquad \Longrightarrow \tag{A37}$$

$$\Longrightarrow \quad \frac{d\,\mathbf{P}}{dt} \quad = \quad \mathbf{F}\,e^{\mathbf{F}\,t}\,\mathbf{C}(t)\,e^{\mathbf{F}^T\,t} + e^{\mathbf{F}\,t}\,\mathbf{C}(t)\,e^{\mathbf{F}^T\,t}\,\mathbf{F}^T + e^{\mathbf{F}\,t}\,\frac{d\,\mathbf{C}(t)}{dt}\,e^{\mathbf{F}^T\,t} \quad = \tag{A38}$$

$$= \quad \mathbf{F}\,\mathbf{P} + \mathbf{P}\,\mathbf{F}^T + e^{\mathbf{F}\,t}\,\frac{d\,\mathbf{C}(t)}{dt}\,e^{\mathbf{F}^T\,t}. \tag{A39}$$

Identifying terms with (A22) we obtain that

$$e^{\mathbf{F}\,t}\,\frac{d\,\mathbf{C}(t)}{dt}\,e^{\mathbf{F}^T\,t} \quad = \quad \mathbf{G}\,\mathbf{Q}\,\mathbf{G}^T \qquad \Longrightarrow \tag{A40}$$

$$\Longrightarrow \qquad \frac{d\,\mathbf{C}(t)}{dt} \quad = \quad e^{-\mathbf{F}\,t}\,\mathbf{Q}\,e^{-\mathbf{F}^T\,t} \quad = \quad \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{(-\mathbf{F})^n\,t^n}{n!}\,\mathbf{Q}\,\frac{(-\mathbf{F}^T)^m\,t^m}{m!} \qquad \Longrightarrow \tag{A41}$$

$$\Longrightarrow \qquad \mathbf{C}(t) \quad = \quad \mathbf{C}_0 \;+\; \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{(-\mathbf{F})^n}{n!}\,\mathbf{Q}\,\frac{(-\mathbf{F}^T)^m}{m!}\,\frac{t^{n+m+1}}{n+m+1}. \tag{A42}$$

Finally, truncating the summation in (A42) at the first non-zero elements, and inserting the result into (A37), we obtain (32) where we have identified $\mathbf{C}_0 = \mathbf{P}(0)$ through the initial conditions.

Appendix B.2. Measurement Prediction

This subsection contains the derivation of the equations for the measurement prediction.

Appendix B.2.1. Expected Value of the Measurement Prediction

Taking expected values on (24), and assuming $\bar{\mathbf{r}}_t^\omega = \mathbf{0}$ we arrive at (35). On the other hand, approximating (23) with its Taylor series up to first order around the current estimation of the state $(\bar{q}, \bar{\omega}')$, taking its expected value, and assuming $\bar{\mathbf{r}}_t^{\mathbf{v}} = \mathbf{0}$ we obtain (34).

Appendix B.2.2. Covariance Matrix of the Measurement Prediction

In order to find the covariance matrix of the measurement prediction we need the linear approximation of the vector measurement around the point $\mathbf{x}_0 := (\,\mathbf{e} = \mathbf{0}\,,\,\mathbf{q}^{\mathbf{v}} = \bar{\mathbf{q}}^{\mathbf{v}}\,,\,\mathbf{r}^{\mathbf{v}} = \bar{\mathbf{r}}^{\mathbf{v}}\,)$:

$$\mathbf{v}^m \quad \approx \quad \bar{\mathbf{v}}^m \;+\; \left.\frac{\partial\,\mathbf{v}^m}{\partial\mathbf{e}}\right|_{\mathbf{x}_0}\mathbf{e} \;+\; \left.\frac{\partial\,\mathbf{v}^m}{\partial\mathbf{q}^{\mathbf{v}}}\right|_{\mathbf{x}_0}(\mathbf{q}^{\mathbf{v}} - \bar{\mathbf{q}}^{\mathbf{v}}) \;+\; \left.\frac{\partial\,\mathbf{v}^m}{\partial\mathbf{r}^{\mathbf{v}}}\right|_{\mathbf{x}_0}(\mathbf{r}^{\mathbf{v}} - \bar{\mathbf{r}}^{\mathbf{v}}). \tag{A43}$$

It is direct to identify

$$\left.\frac{\partial\,\mathbf{v}^m}{\partial\mathbf{q}^{\mathbf{v}}}\right|_{\mathbf{x}_0} \quad = \quad \mathbf{R}^T(\bar{q}), \tag{A44}$$

$$\left.\frac{\partial\,\mathbf{v}^m}{\partial\mathbf{r}^{\mathbf{v}}}\right|_{\mathbf{x}_0} \quad = \quad \mathbf{I}. \tag{A45}$$

On the other hand, rewriting (23) as

$$v^m \quad = \quad \delta^* * \bar{q}^* * (\,q^{\mathbf{v}} + v\,) * \bar{q} * \delta \;+\; r^{\mathbf{v}} \qquad \equiv \tag{A46}$$

$$\equiv \qquad \mathbf{v}^m \quad = \quad \mathbf{R}^T(\delta)\,\mathbf{R}^T(\bar{q})\,(\,\mathbf{q}^{\mathbf{v}} + \mathbf{v}\,) \;+\; \mathbf{r}^{\mathbf{v}}, \tag{A47}$$

and noting that setting $\mathbf{e} = \mathbf{0}$ is equivalent to do $\delta = 1$,

$$\left.\frac{\partial\,v_i^m}{\partial e_j}\right|_{\mathbf{x}_0} \quad = \quad \sum_k \left.\frac{\partial\,v_i^m}{\partial\delta_k}\right|_{\mathbf{x}_0}\left.\frac{\partial\,\delta_k}{\partial e_j}\right|_{\mathbf{e}=0} \quad = \quad \sum_{kl}\left.\frac{\partial\,R_{il}^T(\delta)}{\partial\delta_k}\right|_{\delta=1}\left[\mathbf{R}^T(\bar{q})\,(\,\bar{\mathbf{q}}^{\mathbf{v}} + \mathbf{v}\,)\right]_l\left.\frac{\partial\,\delta_k}{\partial e_j}\right|_{\mathbf{e}=0}. \tag{A48}$$

Now, recalling (5) we have

$$\left.\frac{\partial\,\mathbf{R}^T(\delta)}{\partial\delta_k}\right|_{\delta=1} \quad = \quad 2\begin{pmatrix} 0 & \delta_{3k} & -\delta_{2k} \\ -\delta_{3k} & 0 & \delta_{1k} \\ \delta_{2k} & -\delta_{1k} & 0 \end{pmatrix} \quad \equiv \quad -2\sum_n \varepsilon_{inl}\,\delta_{nk}, \tag{A49}$$

with ε_{inl} the Levi-Civita symbol, and δ_{nk} the Kronecker delta. Recalling (10) we also have

$$\frac{\partial \delta}{\partial e_j} \quad = \quad \begin{pmatrix} -e_j/4 \\ \delta_{1j}/2 \\ \delta_{2j}/2 \\ \delta_{3j}/2 \end{pmatrix}\Bigg|_{\mathbf{e}=\mathbf{0}} \quad \equiv \quad (1-\delta_{0k})\,\delta_{kj}/2. \tag{A50}$$

Then, introducing in (A48) Equations (A49), (34), and (A50),

$$\frac{\partial v_i^m}{\partial e_j}\Bigg|_{\mathbf{x}_0} \quad \approx \quad -\sum_{kln} \varepsilon_{inl}\,\delta_{nk}\,\overline{v}_l^m\,(1-\delta_{0k})\,\delta_{kj} \quad = \quad \sum_l \varepsilon_{ilj}\,\overline{v}_l^m \quad \equiv \quad \left[\,\overline{\mathbf{v}}^m\,\right]_\times, \tag{A51}$$

where we have used $\varepsilon_{ijl} = -\varepsilon_{ilj}$. Finally, assuming the independence of the random variables $\mathbf{x}_t = (\,\mathbf{e}_t\,,\,\omega'_t\,)^T$, $\mathbf{q}_t^{\mathbf{v}}$, $\mathbf{r}_t^{\mathbf{v}}$, and $\mathbf{r}_t^\omega$, and computing the covariance matrix $\mathbf{S}_t = \mathrm{E}\left[\,(\mathbf{z}_t - \overline{\mathbf{z}}_t)(\mathbf{z}_t - \overline{\mathbf{z}}_t)^T\,\right]$ with $\mathbf{z}_t = (\,\mathbf{v}_t^m\,,\,\omega_t^m\,)^T$ and (A43), we arrive at (37) and (38).

Appendix C. Derivation of the T-matrices

This appendix contains the derivation of the **T**-matrix for each chart.

Appendix C.1. Orthographic

Our transition map (A3) can be written as

$$e_i^{\overline{p}}(\mathbf{e}^{\overline{q}}) \quad = \quad \overline{\delta}_0\,e_i^{\overline{q}} - \sqrt{4 - \sum_k (e_k^{\overline{q}})^2}\,\,\overline{\delta}_i - \sum_{lm} \varepsilon_{ilm}\,\overline{\delta}_l\,e_m^{\overline{q}}, \tag{A52}$$

being ε_{ilm} the Levi-Civita symbol. Finding (45) for (A52) we obtain

$$(\mathbf{T})_{ij} \quad = \quad \left[\,\overline{\delta}_0\,\delta_{ij} - \frac{-\sum_k e_k^{\overline{q}}\,\delta_{kj}}{\sqrt{4 - \sum_k (e_k^{\overline{q}})^2}}\,\overline{\delta}_i - \sum_{lm} \varepsilon_{ilm}\,\overline{\delta}_l\,\delta_{mj}\,\right]_{\mathbf{e}^{\overline{q}}=\overline{\mathbf{e}}^{\overline{q}}} \quad = \tag{A53a}$$

$$= \quad \overline{\delta}_0\,\delta_{ij} + \frac{\overline{e}_j^{\overline{q}}}{\sqrt{4 - \sum_k (\overline{e}_k^{\overline{q}})^2}}\,\overline{\delta}_i - \sum_l \varepsilon_{ilj}\,\overline{\delta}_l, \tag{A53b}$$

This expression can be rewritten in matrix form as

$$\mathbf{T} \quad = \quad \overline{\delta}_0\,\mathbf{I} + \frac{\overline{\delta}\,(\overline{\mathbf{e}}^{\overline{q}})^T}{\sqrt{4 - \|\overline{\mathbf{e}}^{\overline{q}}\|^2}} - \left[\,\overline{\delta}\,\right]_\times. \tag{A54}$$

Finally, recalling that for this chart $\overline{\delta}_0 = \sqrt{1 - \|\overline{\mathbf{e}}^{\overline{q}}\|^2/4}$ and $\overline{\delta} = \overline{\mathbf{e}}^{\overline{q}}/2$, we arrive at the final expression

$$\mathbf{T} \quad = \quad \overline{\delta}_0\,\mathbf{I} + \frac{\overline{\delta}\,\overline{\delta}^T}{\overline{\delta}_0} - \left[\,\overline{\delta}\,\right]_\times. \tag{A55}$$

Appendix C.2. Rodrigues Parameters

First, let us denote the numerator of (A6) as $\mathbf{N}(\mathbf{e}^{\overline{q}})$, and its denominator as $D(\mathbf{e}^{\overline{q}})$:

$$\mathbf{N}(\mathbf{e}^{\overline{q}}) \quad := \quad \overline{\delta}_0\,\mathbf{e}^{\overline{q}} - 2\overline{\delta} - \overline{\delta} \times \mathbf{e}^{\overline{q}}, \tag{A56}$$

$$D(\mathbf{e}^{\overline{q}}) \quad := \quad 2\overline{\delta}_0 + \overline{\delta} \cdot \mathbf{e}^{\overline{q}}. \tag{A57}$$

Now let us evaluate (A56) at $\bar{\mathbf{e}}^{\bar{q}}$:

$$\mathbf{N}\left(\bar{\mathbf{e}}^{\bar{q}}\right) \;=\; \overbrace{\bar{\delta}_0\,\underbrace{\bar{\mathbf{e}}^{\bar{q}}}_{2\,\bar{\delta}/\bar{\delta}_0}}^{\mathbf{0}} - 2\,\bar{\delta} \;-\; \overbrace{\bar{\delta}\times\bar{\mathbf{e}}^{\bar{q}}}^{=\,\mathbf{0}\;(\bar{\delta}\,\|\,\bar{\mathbf{e}}^{\bar{q}})} \;=\; \mathbf{0}. \tag{A58}$$

Then, the approximation of $\mathbf{N}\left(\mathbf{e}^{\bar{q}}\right)$ does not have terms of order $\mathcal{O}(1)$. This means that we will only need to approximate $D\left(\mathbf{e}^{\bar{q}}\right)$ to the zeroth order. Any further approximation would produce, after multiplying by the linear approximation of $\mathbf{N}\left(\mathbf{e}^{\bar{q}}\right)$, a higher order term. Let us then calculate each approximation.

We can rewrite (A56) as

$$N_i\left(\mathbf{e}^{\bar{q}}\right) \;=\; \bar{\delta}_0\,e_i^{\bar{q}} - 2\,\bar{\delta}_i - \sum_{kl}\varepsilon_{ikl}\,\bar{\delta}_k\,e_l^{\bar{q}}, \tag{A59}$$

with ε_{ikl} the Levi-Civita symbol. Applying (44) to (A59),

$$N_i\left(\mathbf{e}^{\bar{q}}\right) \;\approx\; \sum_j\left[\bar{\delta}_0\,\delta_{ij} - \sum_{kl}\varepsilon_{ikl}\,\bar{\delta}_k\,\delta_{lj}\right]_{\mathbf{e}^{\bar{q}}=\bar{\mathbf{e}}^{\bar{q}}}\left(e_j^{\bar{q}} - \bar{e}_j^{\bar{q}}\right) \;=\; \tag{A60a}$$

$$=\; \sum_j\left[\bar{\delta}_0\,\delta_{ij} - \sum_k\varepsilon_{ikj}\,\bar{\delta}_k\right]\left(e_j^{\bar{q}} - \bar{e}_j^{\bar{q}}\right), \tag{A60b}$$

being δ_{ij} the Kronecker delta. Returning to matrix notation, the linear approximation of $\mathbf{N}\left(\mathbf{e}^{\bar{q}}\right)$ is

$$\mathbf{N}\left(\mathbf{e}^{\bar{q}}\right) \;=\; \left[\bar{\delta}_0\,\mathbf{I} - \left[\bar{\delta}\right]_\times\right]\left(\mathbf{e}^{\bar{q}} - \bar{\mathbf{e}}^{\bar{q}}\right) + \mathcal{O}\left(\|\mathbf{e}^{\bar{q}} - \bar{\mathbf{e}}^{\bar{q}}\|^2\right). \tag{A61}$$

On the other hand, evaluating (A57) at $\bar{\mathbf{e}}^{\bar{q}}$ we obtain the zeroth order approximation:

$$D\left(\bar{\mathbf{e}}^{\bar{q}}\right) \;=\; 2\,\bar{\delta}_0 + \bar{\delta}\cdot 2\,\frac{\bar{\delta}}{\bar{\delta}_0} + \mathcal{O}\left(\|\mathbf{e}^{\bar{q}} - \bar{\mathbf{e}}^{\bar{q}}\|\right) \;=\; \tag{A62a}$$

$$=\; \frac{2}{\bar{\delta}_0}\,\underbrace{\left(\bar{\delta}_0^2 + \|\bar{\delta}\|^2\right)}_{1} + \mathcal{O}\left(\|\mathbf{e}^{\bar{q}} - \bar{\mathbf{e}}^{\bar{q}}\|\right) \;=\; \tag{A62b}$$

$$=\; \frac{2}{\bar{\delta}_0} + \mathcal{O}\left(\|\mathbf{e}^{\bar{q}} - \bar{\mathbf{e}}^{\bar{q}}\|\right). \tag{A62c}$$

Finally, combining (A61) and (A62c) we can compute the linear approximation of (A6):

$$\mathbf{e}^{\bar{p}}\left(\mathbf{e}^{\bar{q}}\right) \;=\; 2\left\{\left[\bar{\delta}_0\,\mathbf{I} - \left[\bar{\delta}\right]_\times\right](\mathbf{e}^{\bar{q}} - \bar{\mathbf{e}}^{\bar{q}}) + \mathcal{O}\left(\|\mathbf{e}^{\bar{q}} - \bar{\mathbf{e}}^{\bar{q}}\|^2\right)\right\}\left[\frac{\bar{\delta}_0}{2} + \mathcal{O}\left(\|\mathbf{e}^{\bar{q}} - \bar{\mathbf{e}}^{\bar{q}}\|\right)\right] \;=\; \tag{A63a}$$

$$=\; \bar{\delta}_0\left[\bar{\delta}_0\,\mathbf{I} - \left[\bar{\delta}\right]_\times\right](\mathbf{e}^{\bar{q}} - \bar{\mathbf{e}}^{\bar{q}}) + \mathcal{O}\left(\|\mathbf{e}^{\bar{q}} - \bar{\mathbf{e}}^{\bar{q}}\|^2\right). \tag{A63b}$$

Appendix C.3. Modified Rodrigues Parameters

First, let us denote the numerator of (A9) as $\mathbf{N}(\mathbf{e}^{\bar{q}})$, and its denominator as $D(\mathbf{e}^{\bar{q}})$:

$$\mathbf{N}\left(\mathbf{e}^{\bar{q}}\right) \;=\; 8\,\bar{\delta}_0\,\mathbf{e}^{\bar{q}} - \left(16 - \|\mathbf{e}^{\bar{q}}\|^2\right)\bar{\delta} - 8\,\bar{\delta}\times\mathbf{e}^{\bar{q}}, \tag{A64}$$

$$D\left(\mathbf{e}^{\bar{q}}\right) \;=\; 16 + \|\mathbf{e}^{\bar{q}}\|^2 + \bar{\delta}_0\left(16 - \|\mathbf{e}^{\bar{q}}\|^2\right) + 8\,\bar{\delta}\cdot\mathbf{e}^{\bar{q}}. \tag{A65}$$

Now let us evaluate (A64) at $\overline{\mathbf{e}^{\overline{q}}}$:

$$\mathbf{N}(\overline{\mathbf{e}^{\overline{q}}}) = 8\overline{\delta}_0 \overbrace{\overline{\mathbf{e}^{\overline{q}}}}^{4\overline{\delta}/(1+\overline{\delta}_0)} - \left(16 - \overbrace{\|\overline{\mathbf{e}^{\overline{q}}}\|^2}^{16\|\overline{\delta}\|^2/(1+\overline{\delta}_0)^2}\right)\overline{\delta} - 8\overbrace{\overline{\delta} \times \overline{\mathbf{e}^{\overline{q}}}}^{=\,\mathbf{0}\,(\overline{\mathbf{e}^{\overline{q}}}\|\overline{\delta})} = \tag{A66a}$$

$$= \frac{16\overline{\delta}}{1+\overline{\delta}_0}\left[2\overline{\delta}_0 - \left(1 + \overline{\delta}_0 - \overbrace{\frac{\|\overline{\delta}\|^2}{1+\overline{\delta}_0}}^{1-\overline{\delta}_0^2}\right)\right] = \mathbf{0}. \tag{A66b}$$

Then, as with the RP chart, the approximation of $\mathbf{N}\left(\mathbf{e}^{\overline{q}}\right)$ does not have terms of order $\mathcal{O}(1)$, and we will only need to approximate $D\left(\mathbf{e}^{\overline{q}}\right)$ to the zeroth order.

We can write (A64) as

$$N_i\left(\mathbf{e}^{\overline{q}}\right) = 8\overline{\delta}_0 e_i^{\overline{q}} - \left(16 - \sum_k (e_k^{\overline{q}})^2\right)\overline{\delta}_i - 8\sum_{lm}\varepsilon_{ilm}\overline{\delta}_l e_m^{\overline{q}}, \tag{A67}$$

with ε_{ilm} the Levi-Civita symbol. Applying (44) to (A67),

$$N_i\left(\mathbf{e}^{\overline{q}}\right) \approx \sum_j\left[8\overline{\delta}_0\delta_{ij} + \sum_k 2e_k^{\overline{q}}\delta_{kj}\overline{\delta}_i - 8\sum_{lm}\varepsilon_{ilm}\overline{\delta}_l\delta_{mj}\right]_{\mathbf{e}^{\overline{q}}=\overline{\mathbf{e}^{\overline{q}}}}\left(e_j^{\overline{q}} - \overline{e_j^{\overline{q}}}\right) = \tag{A68a}$$

$$= \sum_j\left[8\overline{\delta}_0\delta_{ij} + 2\overline{e_j^{\overline{q}}}\overline{\delta}_i - 8\sum_l\varepsilon_{ilj}\overline{\delta}_l\right]\left(e_j^{\overline{q}} - \overline{e_j^{\overline{q}}}\right), \tag{A68b}$$

being δ_{ij} the Kronecker delta. Returning to matrix notation, the linear approximation of $\mathbf{N}\left(\mathbf{e}^{\overline{q}}\right)$ is

$$\mathbf{N}\left(\mathbf{e}^{\overline{q}}\right) = \left[8\overline{\delta}_0\mathbf{I} + 2\overline{\delta}(\overline{\mathbf{e}^{\overline{q}}})^T - 8\left[\overline{\delta}\right]_\times\right]\left(\mathbf{e}^{\overline{q}} - \overline{\mathbf{e}^{\overline{q}}}\right) + \mathcal{O}\left(\|\mathbf{e}^{\overline{q}} - \overline{\mathbf{e}^{\overline{q}}}\|^2\right) = \tag{A69a}$$

$$= 8\left[\overline{\delta}_0\mathbf{I} + \frac{\overline{\delta}\,\overline{\delta}^T}{1+\overline{\delta}_0} - \left[\overline{\delta}\right]_\times\right]\left(\mathbf{e}^{\overline{q}} - \overline{\mathbf{e}^{\overline{q}}}\right) + \mathcal{O}\left(\|\mathbf{e}^{\overline{q}} - \overline{\mathbf{e}^{\overline{q}}}\|^2\right). \tag{A69b}$$

On the other hand, evaluating (A65) at $\overline{\mathbf{e}^{\overline{q}}}$ we obtain the zeroth order approximation:

$$D\left(\overline{\mathbf{e}^{\overline{q}}}\right) \approx 16 + 16\frac{\|\overline{\delta}\|^2}{(1+\overline{\delta}_0)^2} + \overline{\delta}_0\left(16 - 16\frac{\|\overline{\delta}\|^2}{(1+\overline{\delta}_0)^2}\right) + 8\overline{\delta}\cdot 4\frac{\overline{\delta}}{1+\overline{\delta}_0} = \tag{A70a}$$

$$= \frac{16}{1+\overline{\delta}_0}\left[(1+\overline{\delta}_0) + \frac{\|\overline{\delta}\|^2}{1+\overline{\delta}_0} + \overline{\delta}_0\left((1+\overline{\delta}_0) - \frac{\|\overline{\delta}\|^2}{1+\overline{\delta}_0}\right) + 2\|\overline{\delta}\|^2\right] = \tag{A70b}$$

$$= \frac{16}{1+\overline{\delta}_0}\left[2 + \overline{\delta}_0\left(2\overline{\delta}_0\right) + 2\left(1-\overline{\delta}_0^2\right)\right] = \frac{64}{1+\overline{\delta}_0}, \tag{A70c}$$

where we have used the equality $\|\overline{\delta}\|^2 = 1 - \overline{\delta}_0^2$ for unit quaternions. Finally, combining (A69b) and (A70c) we can compute the linear approximation of (A9):

$$1\mathbf{e}^{\overline{p}}\left(\mathbf{e}^{\overline{q}}\right) = 4\left\{8\left[\overline{\delta}_0\mathbf{I} + \frac{\overline{\delta}\,\overline{\delta}^T}{1+\overline{\delta}_0} - \left[\overline{\delta}\right]_\times\right]\left(\mathbf{e}^{\overline{q}} - \overline{\mathbf{e}^{\overline{q}}}\right) + \mathcal{O}\left(\|\mathbf{e}^{\overline{q}} - \overline{\mathbf{e}^{\overline{q}}}\|^2\right)\right\}\left[\frac{1+\overline{\delta}_0}{64} + \mathcal{O}\left(\|\mathbf{e}^{\overline{q}} - \overline{\mathbf{e}^{\overline{q}}}\|\right)\right] = \tag{A71a}$$

$$= \frac{1+\overline{\delta}_0}{2}\left[\overline{\delta}_0\mathbf{I} + \frac{\overline{\delta}\,\overline{\delta}^T}{1+\overline{\delta}_0} - \left[\overline{\delta}\right]_\times\right]\left(\mathbf{e}^{\overline{q}} - \overline{\mathbf{e}^{\overline{q}}}\right) + \mathcal{O}\left(\|\mathbf{e}^{\overline{q}} - \overline{\mathbf{e}^{\overline{q}}}\|^2\right) = \tag{A71b}$$

$$= \frac{1}{2}\left[(1+\overline{\delta}_0)\left(\overline{\delta}_0\mathbf{I} - \left[\overline{\delta}\right]_\times\right) + \overline{\delta}\,\overline{\delta}^T\right]\left(\mathbf{e}^{\overline{q}} - \overline{\mathbf{e}^{\overline{q}}}\right) + \mathcal{O}\left(\|\mathbf{e}^{\overline{q}} - \overline{\mathbf{e}^{\overline{q}}}\|^2\right). \tag{A71c}$$

Appendix C.4. Rotation Vector

Let us start evaluating the vector $\delta^{\overline{p}}$ in (A12) and (A13) at the point $\overline{\mathbf{e}}^{\overline{q}}$:

$$\delta^{\overline{p}}\left(\overline{\mathbf{e}}^{\overline{q}}\right) = \overbrace{\overline{\delta}_0\,\widehat{\overline{\mathbf{e}}}^{\overline{q}}\sin\left(\frac{\|\overline{\mathbf{e}}^{\overline{q}}\|}{2}\right)}^{=\,\mathbf{0}} - \underbrace{\cos\left(\frac{\|\overline{\mathbf{e}}^{\overline{q}}\|}{2}\right)}_{\overline{\delta}_0}\,\overline{\boldsymbol{\delta}} - \overbrace{\overline{\boldsymbol{\delta}}\times\widehat{\overline{\mathbf{e}}}^{\overline{q}}\sin\left(\frac{\|\overline{\mathbf{e}}^{\overline{q}}\|}{2}\right)}^{=\,\mathbf{0}} = \mathbf{0}. \tag{A72}$$

where the first and third braces are $\overline{\boldsymbol{\delta}}$.

Then, the first order approximation of $\delta^{\overline{p}}$ around $\overline{\mathbf{e}}^{\overline{q}}$ will have the form

$$\delta^{\overline{p}} = \widetilde{\mathbf{T}}\left(\mathbf{e}^{\overline{q}} - \overline{\mathbf{e}}^{\overline{q}}\right) + \mathcal{O}\left(\|\mathbf{e}^{\overline{q}} - \overline{\mathbf{e}}^{\overline{q}}\|^2\right), \tag{A73}$$

and $\|\delta^{\overline{p}}\| \to 0$ as $\mathbf{e}^{\overline{q}} \to \overline{\mathbf{e}}^{\overline{q}}$. Taking the Taylor series of the $\arcsin x$,

$$\frac{\arcsin\|\delta^{\overline{p}}\|}{\|\delta^{\overline{p}}\|} = \frac{\|\delta^{\overline{p}}\| + \mathcal{O}\left(\|\delta^{\overline{p}}\|^3\right)}{\|\delta^{\overline{p}}\|} = 1 + \mathcal{O}\left(\|\delta^{\overline{p}}\|^2\right) = 1 + \mathcal{O}\left(\|\mathbf{e}^{\overline{q}} - \overline{\mathbf{e}}^{\overline{q}}\|^2\right), \tag{A74}$$

so that (A12) is linearized as

$$2\frac{\delta^{\overline{p}}}{\|\delta^{\overline{p}}\|}\arcsin\|\delta^{\overline{p}}\| = 2\,\widetilde{\mathbf{T}}\left(\mathbf{e}^{\overline{q}} - \overline{\mathbf{e}}^{\overline{q}}\right) + \mathcal{O}\left(\|\mathbf{e}^{\overline{q}} - \overline{\mathbf{e}}^{\overline{q}}\|^2\right). \tag{A75}$$

We only lack the $\widetilde{\mathbf{T}}$ matrix. We will need the linear approximations of $\cos\left(\|\mathbf{e}^{\overline{q}}\|/2\right)$ and $\widehat{\mathbf{e}}^{\overline{q}}\sin\left(\|\mathbf{e}^{\overline{q}}\|/2\right)$ around $\overline{\mathbf{e}}^{\overline{q}}$. To this end we will first obtain the linear approximation of $\|\mathbf{x}\|$:

$$\|\mathbf{x}\| = \sqrt{\sum_k x_k^2} = \tag{A76a}$$

$$= \|\overline{\mathbf{x}}\| + \sum_j\left[\frac{\sum_k x_k\,\delta_{kj}}{\sqrt{\sum_k x_k^2}}\right]_{\mathbf{x}=\overline{\mathbf{x}}}\left(x_j - \overline{x}_j\right) + \mathcal{O}\left(\|\mathbf{x} - \overline{\mathbf{x}}\|^2\right) = \tag{A76b}$$

$$= \|\overline{\mathbf{x}}\| + \sum_j\left[\frac{\overline{x}_j}{\sqrt{\sum_k \overline{x}_k^2}}\right]\left(x_j - \overline{x}_j\right) + \mathcal{O}\left(\|\mathbf{x} - \overline{\mathbf{x}}\|^2\right) = \tag{A76c}$$

$$= \|\overline{\mathbf{x}}\| + \widehat{\overline{\mathbf{x}}}^T\left(\mathbf{x} - \overline{\mathbf{x}}\right) + \mathcal{O}\left(\|\mathbf{x} - \overline{\mathbf{x}}\|^2\right). \tag{A76d}$$

Noticing that

$$\frac{\partial\|\mathbf{x}\|}{\partial\mathbf{x}} = \widehat{\overline{\mathbf{x}}}^T + \mathcal{O}\left(\|\mathbf{x} - \overline{\mathbf{x}}\|\right), \tag{A77}$$

our computations are straightforward:

$$\cos\left(\frac{\|\mathbf{x}\|}{2}\right) = \cos\left(\frac{\|\overline{\mathbf{x}}\|}{2}\right) - \left[\sin\left(\frac{\|\mathbf{x}\|}{2}\right)\frac{1}{2}\left[\widehat{\overline{\mathbf{x}}}^T + \mathcal{O}\left(\|\mathbf{x} - \overline{\mathbf{x}}\|\right)\right]\right]_{\mathbf{x}=\overline{\mathbf{x}}}\left(\mathbf{x} - \overline{\mathbf{x}}\right) + \mathcal{O}\left(\|\mathbf{x} - \overline{\mathbf{x}}\|^2\right) = \tag{A78a}$$

$$= \cos\left(\frac{\|\overline{\mathbf{x}}\|}{2}\right) - \frac{1}{2}\sin\left(\frac{\|\overline{\mathbf{x}}\|}{2}\right)\widehat{\overline{\mathbf{x}}}^T\left(\mathbf{x} - \overline{\mathbf{x}}\right) + \mathcal{O}\left(\|\mathbf{x} - \overline{\mathbf{x}}\|^2\right). \tag{A78b}$$

For our particular case,

$$\cos\left(\frac{\|\mathbf{e}^{\overline{q}}\|}{2}\right) = \overline{\delta}_0 - \frac{1}{2}\overline{\boldsymbol{\delta}}^T\left(\mathbf{e}^{\overline{q}} - \overline{\mathbf{e}}^{\overline{q}}\right) + \mathcal{O}\left(\|\mathbf{e}^{\overline{q}} - \overline{\mathbf{e}}^{\overline{q}}\|^2\right). \tag{A79}$$

On the other hand,

$$1\frac{\sin\left(\frac{\|\mathbf{x}\|}{2}\right)}{\|\mathbf{x}\|} \quad = \tag{A80a}$$

$$= \quad \frac{\sin\left(\frac{\|\bar{\mathbf{x}}\|}{2}\right)}{\|\bar{\mathbf{x}}\|} + \left[\left(\frac{\cos\left(\frac{\|\mathbf{x}\|}{2}\right)}{\|\mathbf{x}\|}\frac{1}{2} - \frac{\sin\left(\frac{\|\mathbf{x}\|}{2}\right)}{\|\mathbf{x}\|^2}\right)\left[\widehat{\widehat{\mathbf{x}}}^T + \mathcal{O}\left(\|\mathbf{x}-\bar{\mathbf{x}}\|\right)\right]\right]_{\mathbf{x}=\bar{\mathbf{x}}} (\mathbf{x}-\bar{\mathbf{x}}) + \mathcal{O}\left(\|\mathbf{x}-\bar{\mathbf{x}}\|^2\right) \quad = \tag{A80b}$$

$$= \quad \frac{\sin\left(\frac{\|\bar{\mathbf{x}}\|}{2}\right)}{\|\bar{\mathbf{x}}\|} + \left[\left(\frac{\cos\left(\frac{\|\bar{\mathbf{x}}\|}{2}\right)}{\|\bar{\mathbf{x}}\|}\frac{1}{2} - \frac{\sin\left(\frac{\|\bar{\mathbf{x}}\|}{2}\right)}{\|\bar{\mathbf{x}}\|^2}\right)\widehat{\widehat{\mathbf{x}}}^T\right](\mathbf{x}-\bar{\mathbf{x}}) + \mathcal{O}\left(\|\mathbf{x}-\bar{\mathbf{x}}\|^2\right). \tag{A80c}$$

Now, taking $\mathbf{x} = \bar{\mathbf{x}} + (\mathbf{x}-\bar{\mathbf{x}})$ we arrive at

$$\frac{\mathbf{x}}{\|\mathbf{x}\|}\sin\left(\frac{\|\mathbf{x}\|}{2}\right) \quad = \quad [\bar{\mathbf{x}}+(\mathbf{x}-\bar{\mathbf{x}})]\frac{\sin\left(\frac{\|\mathbf{x}\|}{2}\right)}{\|\mathbf{x}\|} \quad = \tag{A81a}$$

$$= \quad \widehat{\bar{\mathbf{x}}}\sin\left(\frac{\|\bar{\mathbf{x}}\|}{2}\right) + \left[\frac{\sin\left(\frac{\|\bar{\mathbf{x}}\|}{2}\right)}{\|\bar{\mathbf{x}}\|} + \frac{1}{2}\cos\left(\frac{\|\bar{\mathbf{x}}\|}{2}\right)\widehat{\bar{\mathbf{x}}}\widehat{\bar{\mathbf{x}}}^T - \frac{\sin\left(\frac{\|\bar{\mathbf{x}}\|}{2}\right)}{\|\bar{\mathbf{x}}\|}\widehat{\bar{\mathbf{x}}}\widehat{\bar{\mathbf{x}}}^T\right](\mathbf{x}-\bar{\mathbf{x}}) + \mathcal{O}\left(\|\mathbf{x}-\bar{\mathbf{x}}\|^2\right). \tag{A81b}$$

For our particular case,

$$\frac{\mathbf{e}^{\bar{q}}}{\|\mathbf{e}^{\bar{q}}\|}\sin\left(\frac{\|\mathbf{e}^{\bar{q}}\|}{2}\right) \quad = \quad \bar{\delta} + \left[\frac{\|\bar{\delta}\|}{2\arcsin\|\bar{\delta}\|} + \frac{\bar{\delta}_0}{2}\widehat{\bar{\delta}}\widehat{\bar{\delta}}^T - \frac{\|\bar{\delta}\|}{2\arcsin\|\bar{\delta}\|}\widehat{\bar{\delta}}\widehat{\bar{\delta}}^T\right](\mathbf{e}^{\bar{q}}-\bar{\mathbf{e}}^{\bar{q}}) + \mathcal{O}\left(\|\mathbf{e}^{\bar{q}}-\bar{\mathbf{e}}^{\bar{q}}\|^2\right) \quad = \tag{A82a}$$

$$= \quad \bar{\delta} + \frac{1}{2}\left[\left(\mathbf{I}-\widehat{\bar{\delta}}\widehat{\bar{\delta}}^T\right)\frac{\|\bar{\delta}\|}{\arcsin\|\bar{\delta}\|} + \bar{\delta}_0\widehat{\bar{\delta}}\widehat{\bar{\delta}}^T\right](\mathbf{e}^{\bar{q}}-\bar{\mathbf{e}}^{\bar{q}}) + \mathcal{O}\left(\|\mathbf{e}^{\bar{q}}-\bar{\mathbf{e}}^{\bar{q}}\|^2\right). \tag{A82b}$$

Finally, we just have to replace (A79) and (A82b) in (A13) to obtain the required linear approximation. Returning to the original notation we have

$$2\frac{\delta^{\bar{p}}}{\|\delta^{\bar{p}}\|}\arcsin\|\delta^{\bar{p}}\| \quad = \quad 2\delta^{\bar{p}} + \mathcal{O}\left(\|\mathbf{e}^{\bar{q}}-\bar{\mathbf{e}}^{\bar{q}}\|^2\right) \quad = \tag{A83a}$$

$$= \quad 2\bar{\delta}_0\widehat{\mathbf{e}}^{\bar{q}}\sin\left(\frac{\|\mathbf{e}^{\bar{q}}\|}{2}\right) - 2\cos\left(\frac{\|\mathbf{e}^{\bar{q}}\|}{2}\right)\bar{\delta} - 2\bar{\delta}\times\widehat{\mathbf{e}}^{\bar{q}}\sin\left(\frac{\|\mathbf{e}^{\bar{q}}\|}{2}\right) + \mathcal{O}\left(\|\mathbf{e}^{\bar{q}}-\bar{\mathbf{e}}^{\bar{q}}\|^2\right) \quad = \tag{A83b}$$

$$= \quad 2\bar{\delta}_0\bar{\delta} + \bar{\delta}_0\left[\left(\mathbf{I}-\widehat{\bar{\delta}}\widehat{\bar{\delta}}^T\right)\frac{\|\bar{\delta}\|}{\arcsin\|\bar{\delta}\|} + \bar{\delta}_0\widehat{\bar{\delta}}\widehat{\bar{\delta}}^T\right](\mathbf{e}^{\bar{q}}-\bar{\mathbf{e}}^{\bar{q}}) - \left(2\bar{\delta}_0 - \bar{\delta}^T\left(\mathbf{e}^{\bar{q}}-\bar{\mathbf{e}}^{\bar{q}}\right)\right)\bar{\delta} +$$
$$- \bar{\delta}\times\left\{2\bar{\delta} + \left[\left(\mathbf{I}-\widehat{\bar{\delta}}\widehat{\bar{\delta}}^T\right)\frac{\|\bar{\delta}\|}{\arcsin\|\bar{\delta}\|} + \bar{\delta}_0\widehat{\bar{\delta}}\widehat{\bar{\delta}}^T\right](\mathbf{e}^{\bar{q}}-\bar{\mathbf{e}}^{\bar{q}})\right\} + \mathcal{O}\left(\|\mathbf{e}^{\bar{q}}-\bar{\mathbf{e}}^{\bar{q}}\|^2\right) \quad = \tag{A83c}$$

$$= \quad \left[\left(\bar{\delta}_0\left(\mathbf{I}-\widehat{\bar{\delta}}\widehat{\bar{\delta}}^T\right) - [\bar{\delta}]_\times\right)\frac{\|\bar{\delta}\|}{\arcsin\|\bar{\delta}\|} + \bar{\delta}_0^2\widehat{\bar{\delta}}\widehat{\bar{\delta}}^T + \bar{\delta}\bar{\delta}^T\right](\mathbf{e}^{\bar{q}}-\bar{\mathbf{e}}^{\bar{q}}) + \mathcal{O}\left(\|\mathbf{e}^{\bar{q}}-\bar{\mathbf{e}}^{\bar{q}}\|^2\right) \quad = \tag{A83d}$$

$$= \quad \left[\left(\bar{\delta}_0\left(\mathbf{I}-\widehat{\bar{\delta}}\widehat{\bar{\delta}}^T\right) - [\bar{\delta}]_\times\right)\frac{\|\bar{\delta}\|}{\arcsin\|\bar{\delta}\|} + \widehat{\bar{\delta}}\widehat{\bar{\delta}}^T\right](\mathbf{e}^{\bar{q}}-\bar{\mathbf{e}}^{\bar{q}}) + \mathcal{O}\left(\|\mathbf{e}^{\bar{q}}-\bar{\mathbf{e}}^{\bar{q}}\|^2\right). \tag{A83e}$$

Note that the linear approximations of our transition maps are valid for $\mathbf{e}^{\bar{q}}$ near of $\bar{\mathbf{e}}^{\bar{q}}$. However, we have not made any assumption about the $\bar{\delta}$ quaternion. This means that our linear approximations are exact for any $\bar{\delta} = \varphi^{-1}(\bar{\mathbf{e}}^{\bar{q}})$ in the domain of each **T**-matrix, provided that $\mathbf{e}^{\bar{q}}$ is close enough to $\bar{\mathbf{e}}^{\bar{q}}$.

References

1. Crassidis, J.L.; Markley, F.L.; Cheng, Y. Survey of nonlinear attitude estimation methods. *J. Guid. Control Dyn.* **2007**, *30*, 12–28.
2. Kalman, R.E. A new approach to linear filtering and prediction problems. *J. Basic Eng.* **1960**, *82*, 35–45.
3. Julier, S.J.; Uhlmann, J.K. *New Extension of the Kalman Filter to Nonlinear Systems*; AeroSense'97; International Society for Optics and Photonics: Bellingham, WA, USA 1997, pp. 182–193.

4. Shuster, M.D. A survey of attitude representations. *Navigation* **1993**, *8*, 439–517.

5. Stuelpnagel, J. On the parametrization of the three-dimensional rotation group. *SIAM Rev.* **1964**, *6*, 422–430.

6. Lefferts, E.J.; Markley, F.L.; Shuster, M.D. Kalman filtering for spacecraft attitude estimation. *J. Guid. Control Dyn.* **1982**, *5*, 417–429.

7. Crassidis, J.L.; Markley, F.L. Unscented filtering for spacecraft attitude estimation. *J. Guid. Control Dyn.* **2003**, *26*, 536–542.

8. Markley, F.L. Attitude error representations for Kalman filtering. *J. Guid. Control Dyn.* **2003**, *26*, 311–317.

9. Hall, J.K.; Knoebel, N.B.; McLain, T.W. Quaternion attitude estimation for miniature air vehicles using a multiplicative extended Kalman filter. In Proceedings of 2008 IEEE/ION Position, Location and Navigation Symposium, Monterey, CA, USA, 5–8 May 2008; IEEE: Piscataway, NJ, USA; pp. 1230–1237.

10. VanDyke, M.C.; Schwartz, J.L.; Hall, C.D. Unscented Kalman filtering for spacecraft attitude state and parameter estimation. *Adv. Astronaut. Sci.* **2004**, *118*, 217–228.

11. Markley, F.L. Multiplicative vs. additive filtering for spacecraft attitude determination. *Dyn. Control Syst. Struct. Space* **2004**, *6*, 311–317.

12. Crassidis, J.L.; Markley, F.L. Attitude Estimation Using Modified Rodrigues Parameters. Avaliable online: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19960035754.pdf (accessed on 1 January 2019).

13. Bar-Itzhack, I.; Oshman, Y. Attitude determination from vector observations: Quaternion estimation. *IEEE Trans. Aerosp. Electr. Syst.* **1985**, *AES-21*, 128–136.

14. Mueller, M.W.; Hehn, M.; D'Andrea, R. Covariance correction step for kalman filtering with an attitude. *J. Guid. Control Dyn.* **2016**, *40*, 2301–2306.

15. Julier, S.J.; Uhlmann, J.K. Unscented filtering and nonlinear estimation. *Proc. IEEE* **2004**, *92*, 401–422.

16. Gramkow, C. On averaging rotations. *J. Math. Imag. Vision* **2001**, *15*, 7–16.

17. LaViola, J.J. A comparison of unscented and extended Kalman filtering for estimating quaternion motion. In Proceedings of the American Control Conference, 4–6 June 2003, Denver, CO, USA; IEEE: Piscataway, NJ, USA, 2003, Volume 3; pp. 2435–2440.

18. Xie, L.; Popa, D.; Lewis, F.L. *Optimal and Robust Estimation: With an Introduction to Stochastic Control Theory*; CRC Press: Boca Raton, FL, USA, 2007.