Automatica 128 (2021) 109494

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

A geometric framework for rigid body attitude estimation[☆]

Yujendra Bharathi Mitikiri[a], Kamran Mohseni[a,b,*]

[a] Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL 32611, USA
[b] Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611, USA

ARTICLE INFO

Article history:
Received 27 April 2019
Received in revised form 30 July 2020
Accepted 28 December 2020
Available online 8 March 2021

Keywords:
Attitude estimation
Geometric methods
Nonlinear filters
Sensor fusion
Quaternions

ABSTRACT

In this paper, we present a geometric framework to solve two common attitude estimation problems: (i) a geometric problem using measurement of two reference directions, and (ii) a geometro-kinematic problem using measurement of a single reference direction and rate measurement. Both the aforementioned problems may be formulated as angle optimization problems, which can then be solved to obtain exact closed-form solutions. Since the proposed framework preserves the special nonlinear geometry associated with the space of attitudes, and since we present analytic solutions, the proposed framework yields faster and more accurate solutions than those that are based upon linearization techniques. Furthermore, the framework may be extended beyond traditional output error least-squares, to accommodate other practical, but unconventional, optimality metrics. Of special note, we may generalize the classic vector Triad solution, which uses a primary and a secondary measurement, to one with multiple secondary measurements. Lastly, the presented method can be used to derive previous solutions under a single unifying framework, and thus establishes how they are related to each other in a fundamental way. The geometric framework has been verified in simulations as well as experiments.

1. Introduction

The *attitude* of a rigid body is the orientation of a frame fixed in the body, with respect to a second reference frame. The second frame provides a reference for the attitude specification; it is typically chosen to be one fixed to the local ground, the Sun, stationary with respect to fixed stars *etc*; the choice is usually obvious from the context. We shall use the term *ground-frame* to denote the second reference frame in a general sense, while using the term *body-frame* for the frame fixed in the rigid body.

The problem of rigid body attitude estimation assumes two prominent and distinct flavours based upon the nature of the available measurements: (1) a purely geometric problem that utilizes body-frame measurements of directions fixed in the ground-frame, and (2) a geometro-kinematic problem that additionally incorporates information about the attitude dynamics in the form of an angular velocity measurement, again with respect to the ground-frame. Within each flavour, we have different error formulations that lead to different solution methodologies. Two popular formulations for each flavour are shown in the two columns of Table 1.

The first problem of estimating the attitude by measuring reference directions in a body-fixed frame, has been treated abundantly in literature. One of the earliest, and arguably most intuitive, solution was Black's vector Triad attitude estimator (Black, 1964). A least squares formulation of the problem in terms of measurement errors was posed by Wahba in Wahba (1965). Multiple solutions have been reported for Wahba's problem: using polar decomposition (Farrell & Stuelpnagel, 1966), an SVD method, Davenport's q-method (Keat, 1977), the Quaternion estimator (QUEST) (Markley & Mortari, 2000), a factored-quaternion algorithm (FQA) (Yun, Bachmann, & McGhee, 2008), *etc.* Although both Davenport's q-method and QUEST use the quaternion representation of attitude, the quadratic error formulation in terms of measured quantities leads the problem to ultimately reduce to an eigenvalue–eigenvector problem. In this form, given the vast array of tools available for linear problems, the estimation problem is readily solved. This advantage is, however, associated with the accompanying weakness that it is not straightforward to incorporate non-quadratic-output-error costs in the problem.

Relatedly, the advent of small unmanned vehicles has motivated the development of solutions that depend upon minimal measurement resources in order to reduce the weight and cost of

☆ The authors would like to gratefully acknowledge partial support from the Office of Naval Research, USA and the National Science Foundation, USA. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Erik Weyer under the direction of Editor Torsten Söderström.

* Corresponding author at: Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL 32611, USA.
E-mail addresses: yujendra@gmail.com (Y.B. Mitikiri), mohseni@ufl.edu (K. Mohseni).

https://doi.org/10.1016/j.automatica.2021.109494
0005-1098/© 2021 Published by Elsevier Ltd.

Table 1
Solutions for the problem of rigid body attitude estimation. The left column considers a purely geometric problem; the right column additionally considers kinematics.

Using direction measurements	Using angular velocity and direction measurements
(a) Vector Triad estimation (Black, 1964): Uses orthogonal vector triad formed out of two non-collinear vectors. Utilizes complete information from one measurement, and partial information from the other.	(a) Extended Kalman filter (Lefferts, Markley, & Shuster, 1982): Linearizes attitude dynamics and assumes small attitude corrections. Filter gains derived analytically for quadratic output error optimality.
(b) Output error least squares (Keat, 1977; Shuster & Oh, 1981): Minimizes squared error of direction measurements *vs.* predictions from estimated attitude.	(b) Nonlinear complementary filter (Mahony, Hamel, & Pflimlin, 2008): Nonlinear observer that is (almost) globally asymptotically stable on $SO(3)$. Involves ad-hoc gain tuning.

the sensor payload. This interest is partly fuelled by the availability of cheap commercial-off-the-shelf inertial measurement units (IMUs) that contain MEMS-based gyroscopes and accelerometers. In particular, it is of considerable interest to estimate the attitude using a single direction measurement, possibly supplemented by a rate measurement, thus leading us to the second of the stated problems. The second problem is most frequently solved using an extended Kalman filter (EKF) (Lefferts et al., 1982) or one of its several variants (Andrle & Crassidis, 2015; Markley, 2003; Stovner, Johansen, Fossen, & Schjølberg, 2018). The EKF provides a point-wise attitude estimate and does not require gain tuning. However, resulting from linearization of an intrinsically nonlinear problem, this solution is not robust to large changes in the attitude state (Bar-Itzhack, 1996).

More recently, some solutions have been reported in literature which use nonlinear observers or filters to solve the geometro-kinematic problem (Bar-Itzhack, 1996; Batista, Silvestre, & Oliveira, 2011; Choukroun, Bar-Itzhack, & Oshman, 2004; Grip, Fossen, Johansen, & Saberi, 2012; Mahony et al., 2008; Martin & Sarras, 2018; Trumpf, Mahony, Hamel, & Lageman, 2012). These solutions have typically used an appropriate error signal in negative feedback to estimate the attitude. The solutions in Bar-Itzhack (1996) and Mahony et al. (2008) are quite general, and while having been developed for multiple direction measurements, they extend smoothly to the case of a single direction measurement. The solutions presented in Batista, Silvestre, and Oliveira (2012) and Grip et al. (2012) are more specific to the availability of single direction measurements. A common characteristic in this group of solutions is the use of negative feedback from an error signal to estimate the attitude and an (a-priori) unknown gain, that needs to be tuned in order to achieve satisfactory estimator performance. The final word in rigid body attitude estimation has not been spoken yet, as evidenced by recent articles such as Andrle and Crassidis (2015), Berkane and Tayebi (2019), Izadi and Sanyal (2016) and Stovner et al. (2018).

In contrast to the reported approaches available in literature, this paper analyzes the attitude estimation problems from a geometric perspective. Despite the abundance of work on attitude estimation, very few provide the geometric insight behind the solution. The works in Black (1964) and Mortari (1998) are among the few that do. This paper endeavours to follow their lead and derive the attitude estimator with a strong grounding in the fundamental geometry associated with attitude estimation. In the process, we obtain a unifying geometric framework that yields solutions to both the geometric and the geometro-kinematic problems, while overcoming some of the shortcomings in previously reported solutions. The proposed geometric framework extends the estimation method beyond the traditional output error least squares (Section 4.1.1), to geometric least squares (Section 4.1.2), hard geometric constraints (Section 4.1.3), and also a completely novel generalization of the classic Triad estimator with multiple secondary measurements (Section 4.1.4). The analytic solutions do not resort to gain-tuning to yield the optimal attitude estimates at every time step. Besides the mathematical elegance of having an analytic solution, this also has several applications in autonomous guidance, navigation, and control systems: it enables the deployment of frugal single-vector-measurement sensor-suites, and the zero-latency accuracy of the solution is useful in multiple-vector-measurement suites in overcoming sudden failures or intermittent losses in some of the components without leading to large transient errors that could potentially cause system breakdown.

A brief outline of the paper is as follows. We begin by introducing the geometric approach and formulating the stated problems in the language of mathematics in Section 2. Section 3.1 presents the solution to the geometric problem, and relates it to the existing solutions from literature, while the subsequent Section 3.2 does the same for the geometro-kinematic problem. A filtering method is introduced in Section 4 to address the issue of measurement noise and bias. This is followed by verification of the theory using simulations and experiment in Sections 5 and 6.

2. Notation and problem statement

In this section, we introduce our notation, describe the geometry associated with direction measurements and formulate the attitude estimation problems as well-posed mathematical problems.

The attitude of the rigid body with respect to a reference ground-frame shall be represented using a unit quaternion, indicated using a circle accent denoting unit magnitude, *e.g.* $\mathring{p} = [p_0\ p_1\ p_2\ p_3]^T$, $\mathring{q} = [q_0\ q_1\ q_2\ q_3]^T \ldots$, such that $\mathring{p}^T\mathring{p} = \mathring{q}^T\mathring{q} = \ldots = 1$, so $\mathring{p}, \mathring{q} \in \mathbb{S}^3$, the unit 3-sphere. The first component is the *scalar* component, while the remaining three components are the *vector* components of the quaternion. The product of two quaternions $\mathring{p}$ and $\mathring{q}$ shall be denoted as $\mathring{p}\otimes\mathring{q}$. Let $p = [p_1\ p_2\ p_3]^T$, and $q = [q_1\ q_2\ q_3]^T$ be the vector components of $\mathring{p}$ and $\mathring{q}$. The product $\mathring{p}\otimes\mathring{q}$ is defined as

$$\mathring{p}\otimes\mathring{q} = \begin{bmatrix} p_0 \\ p \end{bmatrix} \otimes \begin{bmatrix} q_0 \\ q \end{bmatrix} = \begin{bmatrix} p_0 q_0 - p^T q \\ p_0 q + p q_0 + p \times q \end{bmatrix}. \tag{1}$$

The above multiplication rule (1) is in fact valid for any arbitrary (possibly non-unit) quaternions. In the case of 3-vectors, they are expanded to a 4-vector by prefixing with a scalar component of zero, when invoked as quaternion operands.

A *reference direction*, denoted in bold as $\mathring{\mathbf{h}}$, $\mathring{\mathbf{k}}$, ..., shall be defined as a unit magnitude 3-vector that points in a specified direction in Euclidean space. Examples include the direction of fixed stars relative to the body, the Earth's magnetic field, gravitational field *etc*. The components of any such direction may be measured in any three-dimensional orthogonal coordinate system. In the context of our problems, two obvious choices for the coordinate system are the *ground-frame* coordinate system (relative to which the rigid body's attitude is to be determined), and the *body-frame* coordinate system. We assume the availability of measurement apparatus to obtain the direction's components, $\mathring{g}$, $\mathring{h}$, $\mathring{k}$, ..., $\mathring{u}$, $\mathring{v}$, $\mathring{w}$, ... $\in \mathbb{S}^2 \subset \mathbb{R}^3$, in the ground- and body-frames. Since directions have unit magnitude, we use a circle-accent similar to unit quaternions to denote them. Additionally, as noted under (1), the components of a direction may be prefixed with a zero and treated as a 4-component unit quaternion, when the context demands so.

A rotation quaternion $\mathring{q}$ transforms the components of a vector between the ground-frame components $\mathring{h}$ and the body-frame components $\mathring{v}$ as:

$$\mathring{h} = \mathring{q} \otimes \mathring{v} \otimes \mathring{q}^{-1} \qquad \text{or} \qquad \mathring{q} \otimes \mathring{v} = \mathring{h} \otimes \mathring{q}\,, \tag{2}$$

where the $\mathring{h}$ and $\mathring{v}$ are zero-prefixed 4-vectors. Eq. (2) expresses the direction measurement constraint as a linear equation in $\mathring{q}$ subject to an implicit nonlinear normalization constraint $\mathring{q}^T\mathring{q} = 1$. An attitude quaternion $\mathring{q} = [q_0 \; q^T]^T$ is thus related to the rotation matrix C as:

$$C = (q_0^2 - q^T q)1_{3\times 3} + 2qq^T + 2q_0[q\times], \tag{3}$$

where $1_{m\times n}$ is the $m \times n$ identity matrix, and $[q\times]p$ is the vector product $q \times p$. The axis–angle formalism is related to $\mathring{q}$ as:

$$\mathring{q} = \begin{bmatrix} \cos(\Phi/2) \\ \sin(\Phi/2)\mathring{n} \end{bmatrix}, \tag{4}$$

for a rotation through Φ about the direction $\mathring{n}$. Subsequently in this document, we shall use the notation $c_{(\cdot)}$ and $s_{(\cdot)}$ to denote the cosine and sine functions for brevity. The attitude quaternion evolves according to the kinematic equation

$$\dot{\mathring{q}} = \frac{1}{2}\mathring{q} \otimes \omega = \frac{1}{2}[\mathring{q}\otimes]\omega = \frac{1}{2}[\otimes\omega]\mathring{q}, \tag{5}$$

where the angular velocity $\omega \in \mathbb{R}^3$ is prefixed with a zero, as described under (1). The symbols $[\mathring{q}\otimes]$ and $[\otimes\mathring{\omega}]$ denote the left and right quaternion multiplication matrices.

The quaternion formalism leads to an elegant division algebra for rotations by furnishing simple algebraic operations for inversion as conjugation ($\mathring{q} = [q_0 \;\; q^T]^T \Rightarrow \mathring{q}^{-1} = [q_0 \;\; -q^T]^T$), the composition of sequential rotations as quaternion multiplication, and interpolation between rotations as geometric interpolation.

2.1. Geometry of direction measurement

A rotation quaternion (or, for that matter, any rotation representation) has three scalar degrees of freedom. A body-frame measurement $\mathring{v}$ of a reference direction has 3 scalar components, that are related to the ground-frame measurement $\mathring{h}$, in terms of the rotation quaternion. However, we also know that the measurement would preserve the magnitude of the direction, *i.e.*, $\mathring{h}^T\mathring{h} = \mathring{v}^T\mathring{v} = 1$, so there is one scalar degree of redundancy in our measurement $\mathring{v}$ and only two scalar degrees of information. Reconciling with this redundancy, we can therefore isolate the attitude from a three-dimensional set of possibilities to a single-dimensional set.

The redundancy can be visualized as shown in Fig. 1. Consider a minimal rigid body formed by a triangular patch with one point on the axis of rotation. The measurement of a single reference direction $\mathring{\mathbf{h}}$ in the body-frame confines the body's attitude to form a conical solid of revolution about $\mathring{\mathbf{h}}$: those and only those attitudes on the cone would yield the same components $\mathring{v}$ (if none of the three points were on the axis, it would be a cylinder of revolution, but the cone is easier for subsequent visualizations). We shall refer to the set of attitude quaternions consistent with a measurement as the *feasibility cone* V_h corresponding to that measurement $\mathring{v}$, *i.e.*, the measurement confines the attitude quaternion $\mathring{q}$ to lie in V_h. From the previous discussion, V_h is one-dimensional and $\mathring{q}$ has effectively a single degree of freedom. We shall repeatedly draw intuition from the geometry in Fig. 1 to guide us in the solutions to the stated problems.

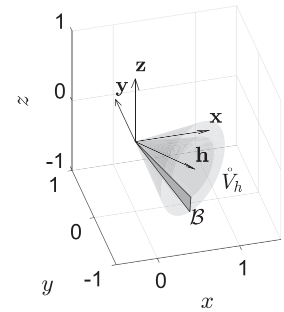

Fig. 1. Possible attitudes of a minimal rigid body $\mathcal{B}$ formed out of three non collinear points (represented by the triangular patch) consistent with a measurement of a single reference direction $\mathring{\mathbf{h}}$. The subspace is a cone of revolution V_h about the direction being measured.

2.2. Problem 1. Estimation from measurements of two reference directions

Let the components of two reference directions $\mathring{\mathbf{h}}$ and $\mathring{\mathbf{k}}$ be $\mathring{v} = [v_1 \;\; v_2 \;\; v_3]^T$ and $\mathring{w} = [w_1 \;\; w_2 \;\; w_3]^T$ in the body-frame, and $\mathring{h} = [h_1 \;\; h_2 \;\; h_3]^T$ and $\mathring{k} = [k_1 \;\; k_2 \;\; k_3]^T$ in the ground-frame respectively. As described above, each reference direction measurement provides two scalar degrees of information regarding the attitude of the rigid body. It is immediately clear that the problem is overspecified, and we have more equations than unknowns. Geometrically, we have two feasibility cones V_h and W_k, with the plane of the body intersecting along a line, but with different roll angles for the body about this line. Thus there is no exact solution to this problem in general, unless some of the measurement information is redundant or discarded.

The approach in this paper is to first determine two solutions $\mathring{q}$ and $\mathring{p}$, one each lying on each of the feasibility cones V_h and W_k corresponding to the measurements $\mathring{v}$ and $\mathring{w}$, and closest to the other cone in a *geometric sense* of presenting the smallest angle of deviation. The two estimates $\mathring{q}$ and $\mathring{p}$ form the Triad solutions, and can subsequently be filtered appropriately to obtain the final attitude estimate.

The first problem can therefore be stated as: *given the measurements $\mathring{v}$ and $\mathring{w}$ in a rotated body-frame, of two reference directions with ground-frame measurements $\mathring{h}$ and $\mathring{k}$, we would like to determine two estimates of the rotated system's attitude quaternion $\mathring{q} \in V_h$ closest (in a geometric sense) to W_k and $\mathring{p} \in W_k$ closest (in a geometric sense) to V_h, where V_h and W_k are the respective feasibility cones.*

2.3. Problem 2. Estimation from rate measurement and measurement of single direction

Suppose we have a body-frame measurement $\omega = [\omega_1 \; \omega_2 \; \omega_3]^T$ of the angular velocity $\boldsymbol{\omega}$ of the rigid body, and that we also have a body-frame measurement $\mathring{v} = [v_1 \;\; v_2 \;\; v_3]^T$ of a reference direction $\mathring{\mathbf{h}}$. The components of $\mathring{\mathbf{h}}$ in the ground-frame are also known, say $\mathring{h} = [h_1 \;\; h_2 \;\; h_3]^T$. The problem is to derive an optimal estimate of the body's attitude $\mathring{q}(t)$ on the basis of the pair of measurements $\omega(t)$ and $\mathring{v}(t)$, and knowing $\mathring{h}$. In order to determine this estimate, we shall first integrate the kinematics to find a dead-reckoned estimate $\mathring{p}$ for the attitude, and then find the attitude $\mathring{q}$ on the feasibility cone V_h closest (in a geometric sense) to the dead-reckoned attitude $\mathring{p}$. The sequence of corrections from $\mathring{p}$ to $\mathring{q}$ may also be used to determine any constant bias in the angular velocity measurement under certain excitation conditions as described later.

The second problem can therefore be stated as: *given the body-frame measurement $\mathring{v}$ of a single reference direction with* *ground-frame measurement $\mathring{h}$, and an attitude estimate $\mathring{p}$ obtained by integrating the body's angular velocity ω relative to the ground-frame, from a known initial condition, we would like to estimate the rotated system's attitude quaternion $\mathring{q} \in V_h$, the feasibility cone corresponding to the body-frame measurement $\mathring{v}$, which is closest (in a geometric sense) to $\mathring{p}$.*

2.4. Nature of measurements

The reference direction measurements are assumed to have noise in each of the components, but that they are subsequently normalized for unit magnitude before being passed on to the attitude estimator. This is the most common situation in practice. The convention of using directions rather than vectors seems to have originated from the initial work on spacecraft attitude estimation using star trackers. Any deterministic errors in the measurement are also assumed to be compensated for, *e.g.* acceleration compensation in gravity sense, local field compensation in magnetic field sense. The issue of acceleration compensation is addressed, for example, by the authors in Mitikiri and Mohseni (2019b). However, that treatment is outside the scope of this work, since it uses approximate dynamical relations, and is specific to a certain class of robots. In contrast, the attitude estimation presented in this work is based upon geometry and kinematics, and applicable to a wider variety of problems.

The angular velocity is not of unit magnitude, in general. Its measurement is also assumed to have noise in each of the components. Besides the noise, the angular velocity may also have a constant (or slowly time-varying) bias error. Compensation of a quasistatic gyroscopic bias using the geometric approach was originally addressed by the authors in Mitikiri and Mohseni (2019a) and is incorporated in this paper.

The kind of noise in the measurements has typically been assumed to be zero-mean, Gaussian noise (Choukroun, Bar-Itzhack, & Oshman, 2006; Crassidis, Markley, & Cheng, 2007), and we shall primarily follow the same convention in this paper.

3. Geometric framework for attitude estimation

Before diving into the solutions, we shall state a few simple results in preparation of the main results.

Lemma 1. *The Euclidean distance $\|\mathring{q}-\mathring{1}\|$ of an attitude quaternion, $\mathring{q} = [c_{\Phi/2} \;\; s_{\Phi/2}\mathring{n}^T]^T$, from the identity element, $\mathring{1}$, is a positive definite and monotonic function of the magnitude of the principal angle of rotation Φ.*

Proof. This is a simple consequence of the trigonometric half-angle identities.

$$\|\mathring{q} - \mathring{1}\|^2 = (c_{\Phi/2} - 1)^2 + s_{\Phi/2}^2 = 4\sin^2(\Phi/4),$$

which is a positive definite monotonic function of $|\Phi|$ for $\Phi \in [-2\pi, 2\pi]$. A corollary is that the distance $\|\mathring{q} - \mathring{p}\|$ between two attitude quaternions is a positive-definite, monotonic function of the angle corresponding to the quaternion $\mathring{q}^{-1} \otimes \mathring{p}$ that takes $\mathring{q}$ to $\mathring{p}$. □

We next provide two particular solutions for the simpler problem of estimating the attitude quaternion using a single reference direction measurement, in Lemma 2. Recall the algebraic constraint $\mathring{q} \otimes \mathring{v} = \mathring{h} \otimes \mathring{q}$ imposed by a direction measurement on the attitude quaternion $\mathring{q}$, where $\mathring{h}$ and $\mathring{v}$ are the components of a direction in the ground and body-frames respectively.

Lemma 2. *Suppose the components of a reference direction are given by $\mathring{h}$ and $\mathring{v}$ in the ground- and body-frame respectively. Let $\Phi = \operatorname{acos} \mathring{v}^T\mathring{h}$, $c = \cos\Phi/2 = \sqrt{(1+\mathring{v}^T\mathring{h})/2}$ and $s = \sin\Phi/2 = \sqrt{(1-\mathring{v}^T\mathring{h})/2}$. Then, two particular solutions for the body's attitude are given by (Davenport, 1968):*

$$\mathring{r}_n = \begin{bmatrix} c \\ s\dfrac{\mathring{v} \times \mathring{h}}{\|\mathring{v} \times \mathring{h}\|} \end{bmatrix}, \quad \mathring{r}_x = \begin{bmatrix} 0 \\ \dfrac{\mathring{v} + \mathring{h}}{\|\mathring{v} + \mathring{h}\|} \end{bmatrix}. \tag{6}$$

Proof. These two solutions are orthogonal in quaternion space, and correspond to the smallest and largest single axis rotations in $[0, \pi]$ that are consistent with the direction measurement in three-dimensional Euclidean space. Geometrically, the first is a rotation through $\operatorname{acos}(\mathring{v}^T\mathring{h})$ about $(\mathring{v} \times \mathring{h})/\|\mathring{v} \times \mathring{h}\|$, the second is a rotation through π about $(\mathring{v} + \mathring{h})/\|\mathring{v} + \mathring{h}\|$. Noting that $\|\mathring{v} \times \mathring{h}\| = \sin\Phi = 2sc$, and $\|\mathring{v} + \mathring{h}\| = 2c$, direct substitution of $\mathring{r}_n$ and $\mathring{r}_x$ for $\mathring{q}$ may be seen to satisfy the constraint of (2). As a clarification, when $\mathring{v} \to \mathring{h}$, $\mathring{r}_n$ and $\mathring{r}_x$ are assumed to take the obvious limits, $\mathring{1}$ and $\mathring{h}$, and when $\mathring{v} \to -\mathring{h}$, they are assumed to take the obvious limits, $\mathring{i}$ and $\mathring{j}$, where $[\mathring{h}\ \mathring{i}\ \mathring{j}]$ is an orthogonal vector triplet. In the latter case ($\mathring{v} + \mathring{h} \to 0$), the orthogonal triad is non-unique, but certain to exist. □

The two special solutions can be rotated by any arbitrary angle about the reference direction $\mathring{h}$ and we would still lie within the feasibility cone, as shown in the next lemma.

Lemma 3. *If $\mathring{q}$ lies in the feasibility cone V_h of the measurement $\mathring{v}$ for the reference direction $\mathring{h}$, then so does any attitude quaternion obtained by rotating $\mathring{q}$ through an arbitrary angle about $\mathring{h}$. Conversely, all attitude quaternions lying on the feasibility cone are related to each other by rotations about $\mathring{h}$.*

Proof. Let Φ be any angle, and let $\mathring{p}$ be $\mathring{q}$ rotated through Φ about $\mathring{h}$, *i.e.*,

$$\mathring{p} = \begin{bmatrix} c \\ s\mathring{h} \end{bmatrix} \otimes \mathring{q},$$

where $c = \cos\Phi/2$ and $s = \sin\Phi/2$. Again, substituting $\mathring{p}$ for $\mathring{q}$ in (2) shows that $\mathring{p}$ also satisfies the measurement constraint if $\mathring{q}$ does.

Conversely, $\mathring{q}^{-1}\otimes\mathring{h}\otimes\mathring{q} = \mathring{v} = \mathring{p}^{-1}\otimes\mathring{h}\otimes\mathring{p}$ implies that $\mathring{p}\otimes\mathring{q}^{-1}\otimes\mathring{h}$ equals $\mathring{h} \otimes \mathring{p} \otimes \mathring{q}^{-1}$, *i.e.*, the rotation $\mathring{p} \otimes \mathring{q}^{-1}$ commutes with the axis $\mathring{h}$. This implies that $\mathring{p} \otimes \mathring{q}^{-1}$ is a rotation about $\mathring{h}$ itself, which proves the claim. □

We already see from Lemma 3 that we have a one dimensional infinity of possible solutions for the attitude quaternion if we have a single reference direction measurement. In order to obtain a unique solution, we could add either another direction measurement (geometric problem), or include an angular velocity measurement (geometro-kinematic problem).

3.1. Attitude estimation from two direction measurements

We now derive a unique solution for the attitude quaternion when we have measurements of two reference directions and would like to incorporate both of them in deriving the attitude estimate. Let $\mathring{v}$ and $\mathring{w}$ be the body-frame components of reference directions $\mathring{h}$ and $\mathring{k}$ ($\mathring{h}, \mathring{k} \in \mathbb{S}^2$ contain the ground-frame components of the same two vectors) respectively. Suppose the rotation quaternion is estimated to be $\mathring{p} = [p_0\ p]^T$ on the basis of $\mathring{v}$, and it is independently estimated to be $\mathring{q} = [q_0\ \ q]^T$ on the basis of $\mathring{w}$, both estimates being obtained by applying, for example, Lemma 2.

The estimates $\mathring{p}$ and $\mathring{q}$ are each indeterminate to one scalar degree of freedom as shown in Lemma 3: a rotation about the corresponding vectors $\mathring{h}$ and $\mathring{k}$ respectively. Let these rotations be parameterized by the quaternions $\mathring{r}_p = [c_1\ s_1\mathring{h}]^T$ and $\mathring{r}_q = [c_2\ s_2\mathring{k}]^T$ respectively where $c_i = \cos\Phi_i/2$ and $s_i = \sin\Phi_i/2$ for $i \in \{1, 2\}$. The problem is to determine the optimal values of Φ_1 and Φ_2 so as to minimize the displacement from the rotated $\mathring{r}_p \otimes \mathring{p}$ to $\mathring{r}_q \otimes \mathring{q}$.

$$\mathring{r}_p \otimes \mathring{p} = \begin{bmatrix} c_1 \\ s_1\mathring{h} \end{bmatrix} \otimes \begin{bmatrix} q_0 \\ q \end{bmatrix} = \begin{bmatrix} c_1 q_0 - s_1 q^T \mathring{h} \\ c_1 q + s_1 q_0 \mathring{h} + s_1 \mathring{h} \times q \end{bmatrix},$$
$$\mathring{r}_q \otimes \mathring{q} = \begin{bmatrix} c_2 \\ s_2\mathring{k} \end{bmatrix} \otimes \begin{bmatrix} p_0 \\ p \end{bmatrix} = \begin{bmatrix} c_2 p_0 - s_2 p^T \mathring{k} \\ c_2 p + s_2 p_0 \mathring{k} + s_2 \mathring{k} \times p \end{bmatrix}. \tag{7}$$

We could either minimize $\|\mathring{r}_p \otimes \mathring{p} - \mathring{r}_q \otimes \mathring{q}\|^2$, or equivalently from Lemma 1, maximize the first component of $(\mathring{r}_p \otimes \mathring{p})^{-1} \otimes \mathring{r}_q \otimes \mathring{q}$. In order to keep the reasoning straightforward, we choose the former. So we need to minimize the cost function

$$\begin{aligned} J(\Phi_1, \Phi_2) &= (c_1 q_0 - s_1 q^T \mathring{h} - c_2 p_0 + s_2 p^T \mathring{k})^2 \\ &\quad + \|c_1 q + s_1(q_0 \mathring{h} + \mathring{h} \times q) - c_2 p - s_2(p_0 \mathring{k} + \mathring{k} \times p)\|^2, \\ &= 2 + 2l_1 c_1 c_2 + 2l_2 s_1 s_2 + 2l_3 c_1 s_2 + 2l_4 s_1 c_2, \end{aligned} \tag{8}$$

where $l_1 = -q_0 p_0 - q^T p$, $l_2 = (-q_0 p^T + p_0 q^T - (q \times p)^T)\mathring{h} \times \mathring{k} - (q_0 p_0 + q^T p)\mathring{h}^T \mathring{k}$, $l_3 = \mathring{k}^T(q_0 p - p_0 q + q \times p)$, and $l_4 = \mathring{h}^T(p_0 q - q_0 p + p \times q)$, are known quantities. Now minimizing the cost function with respect to the independent pair of variables $\Phi_1 + \Phi_2$ and $\Phi_1 - \Phi_2$ yields

$$\begin{bmatrix} \Phi_1 - \Phi_2 \\ \Phi_1 + \Phi_2 \end{bmatrix} = 2 \begin{bmatrix} \operatorname{atan2}(l_3 - l_4, -(l_1 + l_2)) \\ \operatorname{atan2}(-(l_3 + l_4), l_2 - l_1) \end{bmatrix}. \tag{9}$$

Eq. (9) can be solved for Φ_1, and Φ_2, and that completes the solution. The above derivation can be summarized in the form of the following theorem:

Theorem 4. *If $\mathring{p}$ and $\mathring{q}$ are any two special attitude estimates for a rotated system, derived independently using the body-frame measurements $\mathring{v}$ and $\mathring{w}$ of two linearly independent ground-frame directions $\mathring{h}$ and $\mathring{k}$ respectively, then the optimal estimate incorporating the measurement $\mathring{w}$ in $\mathring{p}$ is $\mathring{r}_p \otimes \mathring{p}$, and the optimal estimate incorporating the measurement $\mathring{v}$ in $\mathring{q}$ is given by $\mathring{r}_q \otimes \mathring{q}$, where $\mathring{r}_p = [c_1\ s_1\mathring{h}]^T$ and $\mathring{r}_q = [c_2\ s_2\mathring{k}]^T$, $c_i = \cos\Phi_i$, $s_i = \sin\Phi_i$, and Φ_1 and Φ_2 are given by (9).*

Proof. The proof follows from the construction leading to Eqs. (7), (8), and (9). Refer to Fig. 2. □

Remark 4.1 (*Degeneracy When Measuring the Same Reference Direction*)**.** When $\mathring{h} = \mathring{k}$, we obtain $l_1 = l_2$ and $l_3 + l_4 = 0$. So, we obtain a unique solution for $\Phi_1 - \Phi_2$, but $\Phi_1 + \Phi_2$ is indeterminate, and the solution is again the original pair of degenerate feasibility cones around the direction $\mathring{h} = \mathring{k}$.

Remark 4.2 (*Relation to the Triad Attitude Estimate (Black, 1964)*)**.** The attitude estimates $\mathring{r}_p \otimes \mathring{p}$ and $\mathring{r}_q \otimes \mathring{q}$, where $\mathring{r}_p = [c_1\ s_1\mathring{h}]^T$ and $\mathring{r}_q = [c_2\ s_2\mathring{k}]^T$, are the same as the Triad solution in literature (Shuster & Oh, 1981). Each of them individually yields an estimate that is competely consistent with a primary measurement, but only partially consistent with a secondary measurement.

Corollary 5. *The rotation from the Triad estimate $\mathring{r}_p \otimes \mathring{p}$ to $\mathring{r}_q \otimes \mathring{q}$ in Theorem 4 is about an axis perpendicular to both $\mathring{h}$ and $\mathring{k}$.*

Proof. Let $\mathring{p}' = \mathring{r}_p \otimes \mathring{p}$ and $\mathring{q}' = \mathring{r}_q \otimes \mathring{q}$ be the optimal Triad estimates. Let us now optimize upon these optimal estimates. That should return no required corrections, *i.e.* $\mathring{r}'_p = \mathring{r}'_q = \mathring{1}$. This is equivalent to saying $\Phi'_1 = \Phi'_2 = 0$. This in turn is equivalent to $l'_3 = l'_4 = 0$, or $\mathring{h}^T(p'_0 q' - q'_0 p' + p' \times q') = \mathring{k}^T(q'_0 p' - p'_0 q' + q' \times p') = 0$. But then $q'_0 p' - p'_0 q' - p' \times q'$ is just the vector portion of $\mathring{q}' \otimes \mathring{p}'^{-1}$, the rotation taking the optimal Triad estimate $\mathring{p}'$ to $\mathring{q}'$ in the ground-frame. □

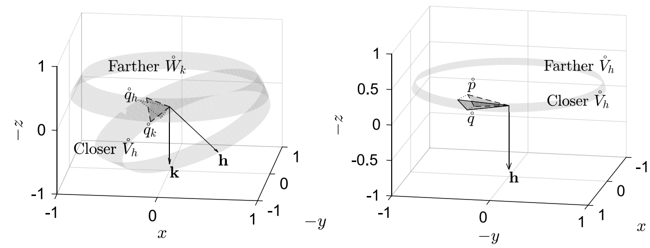

Fig. 2. A visual depiction of the solutions presented in Theorems 4 and 6. The image on the left shows the two solutions $\mathring{r}_p \otimes \mathring{q}$ (dotted triangle) and $\mathring{r}_q \otimes \mathring{p}$ (dashed triangle) of Theorem 4. The figure on the right shows the solution $\mathring{q}$ (solid triangle) of Theorem 6 obtained by projecting the integrated attitude $\mathring{p}$ (dashed triangle) onto the feasibility cone of direction measurement b.

3.2. Attitude estimation from rate measurement and single direction measurement

We first utilize the angular velocity measurement ω that determines the evolution of the attitude in time. The kinematic differential equation for the quaternion is the linear (in the attitude $\mathring{q}$) first order ODE:

$$\dot{\mathring{q}} = \frac{1}{2}\mathring{q} \otimes \omega = \frac{[\otimes\omega]\mathring{q}}{2}. \tag{10}$$

In continuous time, the integration of (10) for a constant ω gives a dead-reckoned attitude estimate

$$\mathring{p}(t+T) = \exp([\otimes\omega]T/2)\mathring{q}(t). \tag{11}$$

Recall that, as stated in Section 2 under (1), 3-vectors are expanded to a 4-vector by prefixing with a scalar component of zero, when invoked as quaternion operands. For a time-varying ω, the state transition matrix replaces the exponential. In discrete time, denoting the integrated estimate as $\mathring{p}(i+1)$, the above equation takes the form

$$\mathring{p}(i+1) = \mathring{q}(i) + \frac{T}{2}\mathring{q}(i) \otimes \omega(i) + \mathcal{O}(\|\omega T\|^2), \tag{12}$$

where T is the time step from the previous estimation of $\mathring{q}(i)$ to the current estimation $\mathring{p}(i+1)$. In the subsequent derivation, we shall omit the time argument of $\mathring{p}$, as there is no ambiguity.

We next write down the constraints imposed by the measurement upon the attitude quaternion $\mathring{q} = [c\ s\mathring{n}^T]^T = [c\ sn_1\ sn_2\ sn_2]^T$, where $c = \cos(\Phi/2)$ and $s = \sin(\Phi/2)$ are functions of the rotation angle Φ, and $\mathring{\mathbf{n}}$ is a unit vector along the rotation axis with components $\mathring{n} = [n_1\ n_2\ n_3]^T$ in the ground-frame. The constraint is given in (2). Converting the quaternion multiplication to matrix notation, (2) can also be written as:

$$\begin{bmatrix} -s\mathring{n}^T\mathring{v} \\ c\mathring{v} + s[\mathring{n}\times]\mathring{v} \end{bmatrix} = \begin{bmatrix} -s\mathring{h}^T\mathring{n} \\ c\mathring{h} + s[\mathring{h}\times]\mathring{n} \end{bmatrix},$$
$$\text{i.e.,} \qquad \begin{bmatrix} -s(\mathring{h} - \mathring{v})^T\mathring{n} \\ c(\mathring{h} - \mathring{v}) + s[(\mathring{h} + \mathring{v})\times]\mathring{n} \end{bmatrix} = 0,$$

where $[\mathring{n}\times]$ denotes the cross product matrix associated with the 3-vector $\mathring{n}$. Expanding the vectors,

$$\begin{bmatrix} 0 & -f_1 & -f_2 & -f_3 \\ f_1 & 0 & -g_3 & g_2 \\ f_2 & g_3 & 0 & -g_1 \\ f_3 & -g_2 & g_1 & 0 \end{bmatrix} \begin{bmatrix} c \\ sn_1 \\ sn_2 \\ sn_3 \end{bmatrix} = 0, \tag{13}$$

$$\text{where} \quad f = \mathring{h} - \mathring{v} \quad \text{and} \quad g = \mathring{h} + \mathring{v}\,,$$
$$\text{so that} \quad f_1 g_1 + f_2 g_2 + f_3 g_3 = f^T g = \mathring{h}^T \mathring{h} - \mathring{v}^T \mathring{v} = 0\,.$$

While it is not obvious, (13) has a double redundancy, so the system of four linear equations actually has rank 2 and nullity 2. This can be seen by considering the solution:

$$\mathring{q} = \begin{bmatrix} c \\ (-cf_2 + sn_3 g_1)/g_3 \\ (cf_1 + sn_3 g_2)/g_3 \\ sn_3 \end{bmatrix}, \tag{14}$$

where sn_1 and sn_2 are solved in terms of c and sn_3 using the inner two row equations in (13). Substituting these in the outer two rows of (13) satisfies them trivially, so these two rows do not yield any additional information. This makes sense as we have not yet imposed the normalization constraint that $n_1^2+n_2^2+n_3^2 = 1$ (c and s, representing $\cos\Phi/2$ and $\sin\Phi/2$, are already assumed to satisfy $c^2+s^2 = 1$). And we are anyway to end up with one degree of freedom in $\mathring{q}$ if using the direction measurement constraint alone, as discussed in Section 2.

The predicted estimate $\mathring{p}$ is then aligned with the direction measurement to yield the attitude estimate $\mathring{q}$. The alignment is realized as a minimum angle rotation. The displacement of the vector-aligned quaternion estimate, $\mathring{q}$ in (14), from the integrated estimate, $\mathring{p}$ in (10), can be expressed as the difference of $\mathring{p}^{-1} \otimes \mathring{q}$ from $\mathring{1}$. But minimizing the distance of a quaternion from the unit quaternion is the same as minimizing the rotation angle Φ (Lemma 1), which is, in turn, the same as maximizing the scalar component of the quaternion, $\cos(\Phi/2)$. Note that, the quaternions $\mathring{p}^{-1} \otimes \mathring{q}$ and $-\mathring{p}^{-1} \otimes \mathring{q}$ affect the same rigid body rotation in 3-dimensional Euclidean space, but minimizing the distance of one from $\mathring{1}$ maximizes the distance of the other in quaternion space. So we just extremize the distance, rather than specifically minimize it. Once we have the solution set, we can check which solution corresponds to a maximum and which to a minimum, and choose the latter.

We therefore need to extremize the scalar component of $\mathring{p}^{-1}\mathring{q}$, where $\mathring{p} = [p_0\ p_1\ p_2\ p_3]^T$ is the attitude estimate obtained by integrating the angular velocity ω as given in (10) and $\mathring{q}$ is expressed in terms of c/s and n_3 as in (14), while enforcing the constraint in (2). This can be accomplished by using the method of Lagrange multipliers to define a cost function that invokes the error norm as well as the constraint. Below, we have multiplied the cost function by the constant g_3 and the constraint by g_3^2, noting that the solution is unaffected by such a scaling:

$$\begin{aligned}
J(\Phi, n_3) &= g_3[\mathring{p}^{-1} \otimes \mathring{q}]_0 + \lambda g_3^2(n_1^2 + n_2^2 + n_3^2 - 1) \\
&= (cp_0 + sn_3 p_3)g_3 + (-cf_2 + sn_3 g_1)p_1 + (cf_1 + sn_3 g_2)p_2 \\
&\quad + \lambda\left(n_3^2 g^T g + 2\frac{cn_3}{s}(f_1 g_2 - f_2 g_1) + \frac{c^2}{s^2}(f_1^2 + f_2^2) - g_3^2\right) \\
&= c(g_3 p_0 + f_1 p_2 - f_2 p_1) + sn_3 g^T p \\
&\quad + \lambda\left(n_3^2 g^T g + 2\frac{cn_3}{s}(f_1 g_2 - f_2 g_1) + \frac{c^2}{s^2}(f_1^2 + f_2^2) - g_3^2\right),
\end{aligned}$$

where p denotes the vector portion of $\mathring{p}$. Now we set the first order partial derivatives of J to 0:

$$0 = \partial_\Phi J = -s(g_3 p_0 + f_1 p_2 - f_2 p_1) + cn_3 g^T p + \left(-\frac{2\lambda}{s^2}\right)\left(\frac{c}{s}(f_1^2 + f_2^2) + n_3(f_1 g_2 - f_2 g_1)\right), \tag{15}$$

$$0 = \partial_{n3} J = sg^T p + 2\lambda g^T g n_3 + 2\lambda \frac{c}{s}(f_1 g_2 - f_2 g_1), \tag{16}$$

$$0 = \partial_\lambda J = n_3^2 g^T g - g_3^2 + \frac{2cn_3}{s}(f_1 g_2 - f_2 g_1) + \frac{c^2}{s^2}(f_1^2 + f_1^2). \tag{17}$$

This yields, for the ratio $\kappa = c/sn_3$:

$$\kappa = \frac{p_0 g^T g + p^T g \times f}{\left(\begin{aligned} p_0(g_1 f_2 - g_2 f_1) + \sum_{1,2} p_i(g_i g_3 - f_i f_3) \\ + p_3(f_1^2 + f_2^2 + g_3^2) \end{aligned}\right)}. \tag{18}$$

Fortuitously, $c/s = \cot(\Phi/2)$ is therefore just proportional to n_3, and upon expressing c/s in terms of n_3 in the normalization constraint ((17)), the resulting equation becomes extremely simple to solve:

$$g_3^2 = g^T g n_3^2 + 2\kappa(f_1 g_2 - f_2 g_1)n_3^2 + \kappa^2 n_3^2(f_1^2 + f_2^2)\,,$$

or

$$n_3 = \frac{g_3}{\sqrt{g^T g + 2\kappa(f_1 g_2 - f_2 g_1) + \kappa^2(f_1^2 + f_2^2)}}\,, \tag{19}$$

$$\frac{c}{s} = \frac{\kappa g_3}{\sqrt{g^T g + 2\kappa(f_1 g_2 - f_2 g_1) + \kappa^2(f_1^2 + f_2^2)}}\,. \tag{20}$$

The other components of the attitude quaternion can be obtained using the inner two rows of (14). Thus we obtain the following theorem.

Theorem 6. *If the angular velocity of a rigid body is integrated to yield an attitude quaternion estimate $\mathring{p}$, then the estimate $\mathring{q} \in V_h$ lying in the feasibility cone of measurement $\mathring{v}$ which is closest to $\mathring{p}$, is given by Eqs.* (14), (18), (19), (20).

Proof. The proof follows from the construction leading to (14), where n_3, c, and s are determined from (18), (19) and (20). Refer to Fig. 2. □

Remark 6.1 (*Solution When Reference Direction is Aligned with z-Axis*)**.** A common application of the presented solution would be to an aerial robot that uses an accelerometer to measure the gravity vector (after acceleration compensation). Since the ground-frame's z-axis is aligned with the reference direction $\mathring{\mathbf{h}}$, we have $f = [(-v_1)\ (-v_2)\ (1 - v_3)]^T$ and $g = [v_1\ v_2\ (1 + v_3)]^T$. Eqs. (18), (20) now simplify to:

$$\kappa = \frac{c}{sn_3} = \frac{(1 + v_3)p_0 - v_1 p_2 + v_2 p_1}{v_1 p_1 + v_2 p_2 + (1 + v_3)p_3}\,, \tag{21}$$

$$\mathring{q} = \begin{bmatrix} c \\ sn_1 \\ sn_2 \\ sn_3 \end{bmatrix} = \frac{1}{\sqrt{2(1+\kappa^2)(1+v_3)}} \begin{bmatrix} \kappa(1 + v_3) \\ \kappa v_2 + v_1 \\ -\kappa v_1 + v_2 \\ (1 + v_3) \end{bmatrix}, \tag{22}$$

where we have used the fact that $(1 + v_3)^2 + v_1^2 + v_2^2 = 2(1 + v_3)$. While the introduction of the auxiliary variable κ in (18), (19), (20) seems adhoc, its role is more clearly visible now. It parameterizes the feasibility cone V_h in terms of the two special solutions provided in Lemma 2:

$$\mathring{q} = \frac{\kappa \mathring{r}_n + \mathring{r}_x}{\sqrt{1 + \kappa^2}} = \frac{(\mathring{r}_n \mathring{r}_n^T + \mathring{r}_x \mathring{r}_x^T)\mathring{p}}{\|(\mathring{r}_n \mathring{r}_n^T + \mathring{r}_x \mathring{r}_x^T)\mathring{p}\|}\,. \tag{23}$$

Remark 6.2 (*Relation to the EKF (Lefferts et al., 1982)*)**.** A filtered attitude estimate $\mathring{q}_f$ can be obtained by projecting the integrated estimate, $\mathring{p}$, onto the feasibility cone corresponding to a filtered direction measurement v_f, to yield the vector aligned estimate $\mathring{q}$ of Theorem 6. The predict-step in Theorem 6 is identical to that in the EKF, especially multiplicative EKF (MEKF): we just integrate the dynamics of the state from the previous time step. Note that the MEKF accommodates nonlinearity in the dynamics in the prediction step, and so it is okay for the attitude dynamics to be bilinear in the state (attitude) and input (angular velocity). The MEKF elegantly accounts for the quaternion normalization by using an auxiliary error state variable ($\delta q(\boldsymbol{a})$ in Markley (2003)) that is almost always close to the identity in practice.

It is the update step where the geometric method shows better accuracy in relation to the MEKF. For the update step, the MEKF assumes that the auxiliary error variable lies in the plane tangent to the identity. The solution in this paper does not make this approximation, and incorporates the geometry of the attitude space into the solution.

The following corollary follows from Theorem 6.

Corollary 7. *The correction that takes the integrated estimate $\mathring{p}$ into the feasibility cone V_h is essentially a rotation about an axis that is orthogonal to the reference direction $\mathring{h}$.*

Proof. With the simplifying choice for the ground-frame z-axis that leads to (22), the proof is simple. The correcting rotation $\mathring{r} = \mathring{q} \otimes \mathring{p}^{-1}$ in the ground-frame coordinate system is:

$$\mathring{r} = \frac{1}{\sqrt{2(1+\kappa^2)(1+v_3)}} \begin{bmatrix} \kappa(1+v_3) \\ \kappa v_2 + v_1 \\ -\kappa v_1 + v_2 \\ (1+v_3) \end{bmatrix} \otimes \begin{bmatrix} p_0 \\ -p_1 \\ -p_2 \\ -p_3 \end{bmatrix}.$$

So, using the expression for κ in (21), and expanding the multiplication rule in (1), we obtain the component $r_3 = 0$. In the general case of an arbitrary h, the proof is more tedious, but still valid (Mitikiri & Mohseni, 2018). The underlying reason for this result is just that a rotation about any other axis would have an unnecessary component about $\mathring{h}$, and that would make the correction to reach V_h suboptimal. □

An elegant expression for the corrected attitude estimate $\mathring{q}$ in terms of the integrated estimate $\mathring{p}$ and direction measurement $\mathring{v}$ of a single vector is:

$$\mathring{q} = \frac{\mathring{p} - \mathring{h} \otimes \mathring{p} \otimes \mathring{v}}{\|\mathring{p} - \mathring{h} \otimes \mathring{p} \otimes \mathring{v}\|}. \tag{24}$$

Eq. (24) is directly consistent with the measurement constraint $\mathring{h} \otimes \mathring{q} = (\mathring{h} \otimes \mathring{p} + \mathring{p} \otimes \mathring{v})/\|\mathring{h} \otimes \mathring{p} + \mathring{p} \otimes \mathring{v}\| = \mathring{q} \otimes \mathring{v}$, so it lies on the feasibility cone by definition. At the same time, the correction $\mathring{q} \otimes \mathring{p}^{-1}$ in the ground-frame coordinate system is about an axis perpendicular to $\mathring{h}$ as required by Corollary 7. If we call $\mathring{p} \otimes \mathring{v} \otimes \mathring{p}^{-1}$ as $\mathring{h}_p$, then

$$\mathring{h}^T(\mathring{q} \otimes \mathring{p}^{-1}) = \mathring{h}^T(\mathring{1} - \mathring{h} \otimes \mathring{h}_p)/\| \dots \| = 0.$$

For a rigorous derivation of (24) from (23), the reader is referred to Mitikiri and Mohseni (2018).

Remark 7.1 (*Relation to the Explicit complementary filter (ECF) (Mahony et al., 2008)*)**.** The ECF in Theorem 5.2 in Mahony et al. (2008) may be realized out of Theorem 6 by noting that the correction quaternion in the body-frame is given by:

$$\mathring{p}^{-1} \otimes \mathring{q} = \frac{\mathring{1} - \mathring{p}^{-1} \otimes \mathring{h} \otimes \mathring{p} \otimes \mathring{v}}{\|\mathring{p} - \mathring{h} \otimes \mathring{p} \otimes \mathring{v}\|} = \frac{\mathring{1} - \mathring{v}_p \otimes \mathring{v}}{\|\mathring{p} - \mathring{h} \otimes \mathring{p} \otimes \mathring{v}\|}, \tag{25}$$

where, $\mathring{v}_p = \mathring{p}^{-1} \otimes \mathring{h} \otimes \mathring{p}$ is the expected measurement of $\mathbf{h}$ in the body-frame, if $\mathring{p}$ was already the correct attitude. On the other hand, the correction from the integrated estimate can be obtained by including a correction term ω_c in the angular velocity such that:

$$\frac{\mathring{q} - \mathring{p}}{T} = \frac{1}{2} \mathring{p} \otimes \omega_c + \mathcal{O}(\|\omega_c T\|^2),$$

where ω_c is the equivalent correction required in the angular velocity over a time-step T. Let $\mathring{v}_p^T \mathring{v} = 2c^2 - 1$. For small corrections, $\mathring{v}_p \approx \mathring{v}$, and so the incremental correction angular velocity is given to first order by:

$$\omega_c \approx \frac{2}{T}\left[\mathring{p}^{-1} \otimes \mathring{q} - \mathring{1}\right] \approx \frac{1}{T}\begin{bmatrix} 0 \\ \mathring{v} \times \mathring{v}_p \end{bmatrix},$$

whose vector portion is exactly the same as that reported in Theorem 5.2 in Mahony et al. (2008), with the gain k_P equal to the time step $1/T$. Note that this also ensures that Theorem 5.2 in Mahony et al. (2008) is dimensionally consistent: k_P must have dimensions of reciprocal time. For values of k_P larger than $1/T$, we obtain a larger correction ω_c, and a larger weight for measurement $\mathring{v}$ in the final filtered estimate. The projection of Eqs. (23) or (24) induces all the error in the direction measurement $\mathring{v}$ onto the attitude estimate $\mathring{q}$. In practice, $\mathring{v}$ itself is determined by optimally filtering between the predicted measurement $\mathring{v}_p$ and the sensor measurement, as described in Section 4.2.

We finally show that the projection from the integrated estimate $\mathring{p}$ to $\mathring{q}$ on the feasibility cone V_h leads to progressively smaller angular deviations from all attitudes on the feasibility cone. This result will be used in establishing the relation between the geometric attitude estimation and Wahba's problem.

Corollary 8. *As the attitude of a rigid body rotates from $\mathring{p}$ outside the feasibility cone corresponding to a direction measurement $\mathring{v}$ of $\mathring{h}$, onto $\mathring{q}$ on the feasibility cone, the angle to any fixed attitude $\mathring{r}$ on the feasibility cone monotonically reduces.*

Proof. Let $\mathring{p}'$ lie on the path of projection from $\mathring{p}$ to $\mathring{q}$, it must be obtained as a rotation through an angle $x\Phi$, where $\Phi < \pi$ is the angle from $\mathring{p}$ to $\mathring{q}$ and $x \in [0, 1]$, about an axis $\mathring{k}$ that is orthogonal to $\mathring{h}$ in the ground-frame coordinate system (Corollary 7):

$$\mathring{p}' = \begin{bmatrix} c_x \\ s_x \mathring{k} \end{bmatrix} \otimes \mathring{p},$$

where $c_x = \cos(x\Phi/2)$ and $s_x = \sin(x\Phi/2)$. The attitude $\mathring{q}$ corresponds to $x = 1$. Since $\mathring{r}$ lies on the feasibility cone, we must also have:

$$\mathring{r} = \begin{bmatrix} c_r \\ s_r \mathring{h} \end{bmatrix} \otimes \mathring{q},$$

where $c_r = \cos(\Phi_r/2)$ and $s_r = \sin(\Phi_r/2)$, for some $\Phi_r \in [-\pi, \pi]$. Then, the scalar component of $\mathring{r} \otimes \mathring{p}'^{-1}$ is $c_r c_{1-x}$, which progressively increases as x goes from 0 to 1, and the angle therefore monotonically reduces from $\mathrm{acos}(c_r c_{1-x})$ to Φ_r (Lemma 1). □

4. Noise filtering and bias compensation

4.1. Interpolation with measurements of two directions

Consider the two attitude estimates obtained using Theorem 4 upon the body-frame measurements $\mathring{v}$ and $\mathring{w}$ of the ground-frame directions $\mathring{h}$ and $\mathring{k}$. We shall use $\mathring{p}$ and $\mathring{q}$ (instead of the longer expressions $\mathring{r}_p \otimes \mathring{p}$ and $\mathring{r}_q \otimes \mathring{q}$ as used in Theorem 4, Remark 4.2, and Corollary 5) to denote the Triad solutions. As described under Remark 4.2, $\mathring{p}$ and $\mathring{q}$ utilize complete information from one of the direction measurements and partial information from the other. We now propose a mechanism to interpolate between the two Triad estimates in order to filter out noise in the individual measurements. The geometrically interpolated quaternion, $\mathring{q}_f$, from $\mathring{p}$ to $\mathring{q}$ is given by any of the following four equivalent expressions (Dam, Koch, & Lillholm, 1998):

$$\begin{aligned} \mathring{q}_f &= \mathring{p} \otimes (\mathring{p}^{-1} \otimes \mathring{q})^x = \mathring{q} \otimes (\mathring{q}^{-1} \otimes \mathring{p})^{1-x} \\ &= (\mathring{p} \otimes \mathring{q}^{-1})^{1-x} \otimes \mathring{q} = (\mathring{q} \otimes \mathring{p}^{-1})^x \otimes \mathring{p}. \end{aligned} \tag{26}$$

The interpolation ratio x is now chosen to perform a desired weighting of the two Triad estimates $\mathring{p}$ and $\mathring{q}$ in the final result.

4.1.1. Output error least-squares a.k.a. Wahba's problem (Keat, 1977)

Let the Triad estimates again be denoted as $\mathring{p}$ and $\mathring{q}$. From Corollary 5, we know that the rotation $\mathring{q} \otimes \mathring{p}^{-1}$ is about an axis orthogonal to both the ground-frame directions $\mathring{h}$ and $\mathring{k}$:

$$\mathring{q} \otimes \mathring{p}^{-1} = \begin{bmatrix} c_{\Phi/2} \\ s_{\Phi/2}(\mathring{h} \times \mathring{k})^T / \|\mathring{h} \times \mathring{k}\| \end{bmatrix},$$

for some Φ. Next, let $\mathring{q}_f$ be the solution to Wahba's problem, that minimizes the loss function

$$J = \alpha \|\mathring{q}_f \otimes \mathring{v} \otimes \mathring{q}_f^{-1} - \mathring{h}\|^2 + \beta \|\mathring{q}_f \otimes \mathring{w} \otimes \mathring{q}_f^{-1} - \mathring{k}\|^2,$$

with weights α, β. Now $\mathring{q}_f$ must lie on the feasibility cone containing $\mathring{p}$ and $\mathring{q}$. Otherwise, we could move it towards the cone so as to reduce both the errors $\|\mathring{q}_f \otimes \mathring{v} \otimes \mathring{q}_f^{-1} - \mathring{h}\|^2$ and $\|\mathring{q}_f \otimes \mathring{w} \otimes \mathring{q}_f^{-1} - \mathring{k}\|^2$ in the loss function (Corollary 8). So, if $\mathring{q}_f \otimes \mathring{p}^{-1}$ and $\mathring{q} \otimes \mathring{q}_f^{-1}$ rotate the body through Φ_p and Φ_q about $\mathring{h} \times \mathring{k}$, then we must have $\Phi_p + \Phi_q = \Phi$. The loss function would be $2\alpha(1 - \cos \Phi_p) + 2\beta(1 - \cos \Phi_q)$. Thus the solution to Wahba's problem maximizes $\alpha \cos \Phi_p + \beta \cos \Phi_q$, subject to $\Phi_p + \Phi_q = \Phi$:

$$-\alpha \sin \Phi_p + \beta \sin(\Phi - \Phi_p) = 0$$
$$\Rightarrow \begin{bmatrix} \tan \Phi_p \\ \tan \Phi_q \end{bmatrix} = \begin{bmatrix} \sin \Phi / (\alpha/\beta + \cos \Phi) \\ \sin \Phi / (\beta/\alpha + \cos \Phi) \end{bmatrix}.$$

The interpolated estimate $\mathring{q}_f$ may be derived as the rotation through Φ_p about $\mathring{h} \times \mathring{k}$ from $\mathring{p}$, or $-\Phi_q$ about $\mathring{h} \times \mathring{k}$ from $\mathring{q}$.

4.1.2. Geometric least-squares

As an alternative to the cosine-maximization in Wahba's problem resulting from an output error least-squares formulation, one could specify a cost function that is quadratic directly in the angular deviations:

$$J = \alpha \Phi_p^2 + \beta \Phi_q^2 \quad \text{subject to } \Phi_p + \Phi_q = \Phi,$$

where Φ_p and Φ_q are the angular deviations of the estimated attitude $\mathring{q}_f$ from the Triad estimates $\mathring{p}$ and $\mathring{q}$, and Φ is the angular deviation between $\mathring{p}$ and $\mathring{q}$. This yields the optimal solution

$$0 = \partial_{\Phi p}(\alpha \Phi_p^2 + \beta(\Phi - \Phi_p)^2) = 2\alpha \Phi_p - 2\beta(\Phi - \Phi_p),$$
$$\Rightarrow \begin{bmatrix} \Phi_p \\ \Phi_q \end{bmatrix} = \frac{\Phi}{\alpha + \beta} \begin{bmatrix} \beta \\ \alpha \end{bmatrix}.$$

An advantage in this formulation is that the optimal angle of rotation Φ_p from $\mathring{p}$ to $\mathring{q}_f$ is monotonic with respect to the deviation Φ from $\mathring{p}$ to $\mathring{q}$. This monotonic response may not hold in an output error least-squares formulation (refer to Fig. 3). The non-monotonic behaviour is more likely if the ratios of the weights $\alpha/\beta \gg 1$. Since we may assume that $\Phi \in [-\pi \ +\pi]$, and the optimal solution $\mathring{q}_f$ is sure to lie in between $\mathring{p}$ and $\mathring{q}$, there is no need to wrap the cost function arguments to lie in between $[-\pi \ +\pi]$.

4.1.3. Incorporating hard inequality constraints

In some applications, it is desirable to impose hard constraints upon the estimated attitude (Kalabic, Gupta, Di Cairano, Bloch, & Kolmanovsky, 2014; Singh, Bortolami, & Page, 2010). Since the presented solution is geometric in nature, it is straightforward to include geometric constraints on the solution using Barrier Lyapunov functions (BLFs) (Tee, Ge, & Tay, 2009) for bounded solutions. Such a strategy can easily be employed in our framework, in contrast with the linear algebraic solutions which are more suitable to handle quadratic forms. The interpolation factor x from $\mathring{p}$ to $\mathring{q}$ is now determined as the argument that minimizes a cost function that contains a BLF:

$$x = \underset{x \in [0,1]}{\operatorname{argmin}}(\alpha f(x) + \beta g(1 - x)), \tag{27}$$

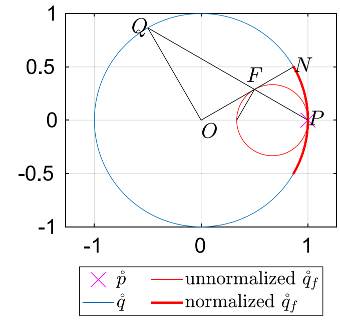

Fig. 3. Spherical linear interpolation between $\mathring{p}$ and $\mathring{q}$ using an output error least-squares could lead to non-monotonic behaviour. Let $\mathring{p}$ be at P. As $\mathring{q}$ moves on the unit sphere $\mathbb{S}^3$ beyond Q, its linear interpolation at F would move farther from P, but its normalization onto the unit sphere, $\mathring{q}_f$ at N, would move closer to P. In particular, when $\mathring{q}$ is opposite to $\mathring{p}$, $\mathring{q}_f$ returns back to $\mathring{p}$!.

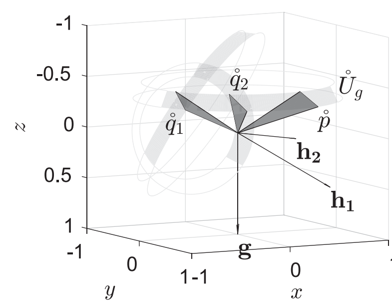

Fig. 4. A generalized $(1 + n)$-measurement Triad attitude estimation using a primary measurement of $\mathbf{g}$ and secondary measurements of $\mathbf{h_i}$ in the body-frame. The primary measurement yields a special attitude estimate $\mathring{p}$ on the feasibility cone U_g, and the secondary measurements yield Triad solutions $\mathring{q}_i$ in conjunction with the primary measurement. The interpolated attitude $\mathring{q}_f$ (not displayed in the figure for clarity) lies on U_g between $\mathring{q}_1$ and $\mathring{q}_2$ so as to minimize an appropriate cost functional.

where $f(x)$ is a BLF (such as, for example, $\sec(\pi x/2)$) with a minimum at the Triad estimate $\mathring{p}$, and $g(1 - x)$ is an appropriate convex function with a minimum at $\mathring{q}$. The optimal x may be obtained using straightforward calculus for a given f and g as the solution of

$$\alpha \left.\frac{df}{dx}\right|_x - \beta \left.\frac{dg}{dx}\right|_{1-x} = 0.$$

4.1.4. A generalized $1 + n$-measurement triad

As a final demonstration of the power of the geometric method, we show how one may generalize the Triad method to the case of multiple secondary direction measurements. Suppose we are given the body-frame measurements $\mathring{u}$, $\mathring{v}_i$ of the ground-frame directions $\mathring{g}$, $\mathring{h}_i$, $i \in \{1, \ldots, n\}$, and the problem is to determine the attitude which is completely consistent with the primary measurement ($\mathring{u}$ of $\mathring{g}$), and partially consistent with a filtered version of the remaining secondary measurements ($\mathring{v}_i$ of $\mathring{h}_i$) (see Fig. 4).

Let $\mathring{p}$ be any attitude (*e.g.* $\mathring{r}_n$ in Lemma 2) lying on the feasibility cone U_g corresponding to the primary measurement. Corresponding to each of the secondary measurements, there exists a unique Triad solution $\mathring{q}_i$ on U_g which may be obtained from $\mathring{p}$ by a rotation through Φ_i about $\mathring{g}$ in the ground-frame (Lemma 3 and Theorem 4). We shall now optimize the attitude in order to have the minimum square output error. As seen previously, this criterion is equivalent to cosine maximization.

Let the optimal attitude be $\mathring{q}$, which is obtained by rotating $\mathring{p}$ through Φ about $\mathring{g}$ in the ground-frame. The optimal attitude would have a residual output error with respect to each of the secondary measurements. If we define the predicted output as

$$\mathring{v}_{pi} = \mathring{p}^{-1} \otimes \mathring{h}_i \otimes \mathring{p},$$

and the reconstructed input as

$$\mathring{h}_{qi} = \mathring{q} \otimes \mathring{v}_i \otimes \mathring{q}^{-1},$$

then the square output error in the ith measurement is (affinely) related to the cosine $\mathring{h}_i^T \mathring{h}_{qi}$ of the angle between $\mathring{h}_i$ and $\mathring{h}_{qi}$. Noting that $\mathring{g}^T \mathring{h}_i = \mathring{u}^T \mathring{v}_{pi}$, $\mathring{g}^T \mathring{h}_{qi} = \mathring{u}^T \mathring{v}_i$, and using the spherical law of cosines, this cosine may be expressed in terms of the scalar parameter $\Phi_i - \Phi$ and other known quantities as

$$\mathring{h}_i^T \mathring{h}_{qi} = \mathring{u}^T \mathring{v}_{pi} \mathring{u}^T \mathring{v}_i + c_{\Phi i - \Phi} \sqrt{1 - (\mathring{u}^T \mathring{v}_{pi})^2} \sqrt{1 - (\mathring{u}^T \mathring{v}_i)^2}.$$

A weight α_i on the ith output square error is therefore transformed as a weight

$$\beta_i = \alpha_i \sqrt{1 - (\mathring{u}^T \mathring{v}_{pi})^2} \sqrt{1 - (\mathring{u}^T \mathring{v}_i)^2},$$

multiplying the cosine $c_{\Phi i - \Phi}$ (the first term is independent of the variable of optimization). Now optimizing upon the parameter Φ yields

$$\tan \Phi = \frac{\sum_i \beta_i s_{\Phi i}}{\sum_i \beta_i c_{\Phi i}}, \tag{28}$$

where

$$c_{\Phi i} = \frac{(\mathring{u} \times \mathring{v}_{pi})^T (\mathring{u} \times \mathring{v}_i)}{\|\mathring{u} \times \mathring{v}_{pi}\| \|\mathring{u} \times \mathring{v}_i\|} = \frac{\mathring{v}_{pi}^T \mathring{v}_i - \mathring{u}^T \mathring{v}_{pi} \mathring{u}^T \mathring{v}_i}{\|\mathring{u} \times \mathring{v}_{pi}\| \|\mathring{u} \times \mathring{v}_i\|},$$

$$s_{\Phi i} = \frac{\mathring{u}^T \mathring{v}_{pi} \times \mathring{v}_i}{\|\mathring{u} \times \mathring{v}_{pi}\| \|\mathring{u} \times \mathring{v}_i\|},$$

are the trigonometric ratios for the residual angle from $\mathring{p}$ to the optimal Triad estimate corresponding to the ith measurement.

With a geometric least-squares formulation, there is no analytic solution, but the solution may be obtained by numerically solving

$$0 = \sum_i \left[\frac{\beta_i \varphi_i \sin(\Phi_i - \Phi)}{\sin \varphi_i} \right].$$

for Φ, where $\cos \varphi_i = \mathring{h}_i^T \mathring{h}_{pi}$.

4.2. Filtering with angular velocity measurement

The updated estimate $\mathring{q}$ in Theorem 6 was derived as a hard projection from the predicted estimate $\mathring{p}$ onto the feasibility cone, V_h, corresponding to the direction measurement $\mathring{v}$ of $\mathring{h}$. This hard projection is the best estimate only if the direction measurement was perfect and noiseless. In a real situation with non-zero noise, we would like to incorporate some kind of filtering that weighs the noise in the direction measurement against the noise in the angular velocity measurement.

In a Kalman filter framework, the filtering during the update-step is typically implemented on the state estimate, as this is the most intuitive interpretation in linear systems. The translation to the space of attitude quaternions is straightforward, but inefficient. Since the direction measurement $\mathring{v}$ confines the state to the corresponding feasibility cone V_h, the interpolation between the integrated estimate $\mathring{p}$ and the corrected estimate $\mathring{q}$ is computationally expensive to account for the yaw-degeneracy in the covariance matrix. An equivalent and more elegant alternative is to consider interpolating between the predicted direction measurement $\mathring{v}_p = \mathring{p}^{-1} \otimes \mathring{h} \otimes \mathring{p}$ and the actual measurement $\mathring{v}_m$ (subscript m denoting the noisy measurement). The two approaches are shown in Fig. 5.

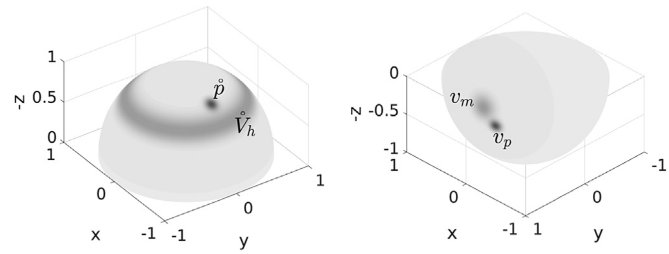

Fig. 5. Left: Traditional Kalman filtering translated to attitude estimation involves interpolating between the integrated attitude estimate $\mathring{p}$ and a feasibility cone V_h of attitudes corresponding to a noisy direction measurement. The 3-sphere attitude space has been projected on a 2-sphere for visualization purposes (by, for example, ignoring the roll component). Right: A computationally efficient alternative for addressing the yaw-degeneracy in the feasibility cone V_h is to filter the direction measurement $\mathring{v}_m$ with respect to its predicted value $\mathring{v}_p$.

The filter for the direction measurement could be implemented geometrically on a unit 2-sphere by assuming spherical Gaussian noise, as done in Section 4.1.2, or 4.1.3. If the noise is relatively small, as is quite common in practice, the filter is implemented by linearizing as

$$\mathring{v} = (V_m + V_p)^{-1}(V_m \mathring{v}_p + V_p \mathring{v}_m), \tag{29}$$

where V_m and V_p are covariance matrices corresponding to the actual direction measurement $\mathring{v}_m$ and the predicted direction measurement $\mathring{v}_p$. The covariance matrix V of the fused measurement in the latter case is:

$$(V_m + V_p)^{-1}(V_m V_p V_m + V_p V_m V_p)(V_m + V_p)^{-1}. \tag{30}$$

The covariance matrix V_p of the predicted measurement may be expressed as

$$V_p = \nabla_p \mathring{v}_p \Pi (\nabla_p \mathring{v}_p)^T. \tag{31}$$

where Π is the covariance matrix of the predicted attitude estimate $\mathring{p}$, and the gradient $\nabla_p \mathring{v}_p$ is given by

$$2\left[(p_0 \mathring{h} + \mathring{h} \times p) \quad (p^T \mathring{h} + p \mathring{h}^T - \mathring{h} p^T + p_0[\mathring{h}\times])\right]. \tag{32}$$

An expression for the covariance matrix Π of the integrated estimate $\mathring{p}$ may be obtained from the kinematic equation (12) for small time-steps.

$$\Pi = \Xi + \frac{T^2}{4} \begin{bmatrix} q_0 & -q^T \\ q & q_0 + [q\times] \end{bmatrix} W \begin{bmatrix} q_0 & q^T \\ -q & q_0 - [q\times] \end{bmatrix}, \tag{33}$$

where Ξ and W are the covariances of the attitude estimate $\mathring{q}$ at the previous time step, and the angular velocity measurement ω.

We now analyse the effect of measurement noise on the updated quaternion $\mathring{q}$. Continuing to assume relatively small, zero-mean, Gaussian noise in the measurements, it can be shown (Mitikiri & Mohseni, 2018) that perturbations $\mathring{p}$ and $\mathring{v}$ in the predicted estimate and direction measurement cause a variation $\delta\mathring{q}$ in the updated estimate given by

$$\delta\mathring{q} = (1 - \mathring{q}\mathring{q}^T)(\mathring{r}_n \mathring{r}_n^T + \mathring{r}_x \mathring{r}_x^T)\delta\mathring{p}\,, \tag{34}$$

and

$$\delta\mathring{q} = -\frac{1}{2}\mathring{q} \otimes \mathring{v} \otimes \delta\mathring{v}\,. \tag{35}$$

Eqs. (12), (34), (35) can be used to derive an equation for the evolution of noise in the integrated and vector-aligned estimates.

4.3. Gyroscopic bias estimation

We shall now consider compensating for the effects of gyroscopic bias on the geometric attitude estimation as done in Mitikiri and Mohseni (2019a). The angular velocity of the body is measured to have components $\hat{\omega}$ in the body-frame. This is the typical scenario in most applications, where the gyroscope is part of an Inertial Measurement Unit (IMU) that is fixed with respect to the body. However, the measured angular velocity $\hat{\omega}$ has an error with respect to the true quantity ω. The angular velocity measurement error is assumed to be an Ornstein–Uhlenbeck process, with mean $\overline{\varpi}$, time-constant τ, and random-walk increments $\tilde{\omega}$:

$$\hat{\omega} = \omega + \overline{\varpi} + \tilde{\omega}. \tag{36}$$

Since the gyroscopic bias is exponentially autocorrelated with a time constant τ that is much larger than the time-step T between measurements, this error manifests as a relatively low frequency source in comparison to the Gaussian noise considered in the previous section. The slow variation enables the design of an observer that could estimate the noise as well as compensate for it.

As described in Mitikiri and Mohseni (2019a), a constant bias $\overline{\varpi}$ may be estimated using the equation

$$\sum_j (\mathring{v}_j \mathring{v}_j^T - 1_{3\times 3})\overline{\varpi} = \sum_j 2r_j/T\,. \tag{37}$$

for a direction measurement $\mathring{v}_j$ at the jth time-step, and where r_j is the vector portion of the correction quaternion $\mathring{r} = \mathring{p}^{-1} \otimes \mathring{q}$. In the absence of any other measurement errors, a fixed bias error may be completely estimated using (37) on two linearly independent direction measurements.

In case of a time varying bias $\overline{\varpi} + \tilde{\omega}$, we would modify (37) to assign different weights to the terms in the summation, so as to form a filter. Such an estimator may be expressed in terms of the matrices A_i and b_i, defined inductively, as shown below:

$$\begin{aligned} A_i &= (1 - T/\tau)A_{i-1} + (T/\tau)(\mathring{v}_i \mathring{v}_i^T - 1_{3\times 3}),\\ b_i &= (1 - T/\tau)b_{i-1} + 2r_i/\tau,\\ A_i&(\overline{\varpi} + \tilde{\omega}_i) = b_i, \end{aligned} \tag{38}$$

with the initial conditions $A_0 = 0$, $b_0 = 0$.

While (38) is sufficient to estimate the bias when the persistency-of-excitation condition is met, it may fail when the body stops rotating if we have the measurements of a single vector. The failure upon loss of excitation occurs as $\mathring{v}_i$ approaches a limit, and the matrix A_i gradually approaches the now constant $\mathring{v}_i \mathring{v}_i^T - 1_{3\times 3}$ over time, thus becoming singular. Failure may be avoided under such circumstances by updating only the components of A_i and b_i that have additional information in the new measurements, as done in the following estimator design:

$$\begin{aligned} A_i &= (\mathring{v}_i \mathring{v}_i^T)A_{i-1}\\ &\quad + (1_{3\times 3} - \mathring{v}_i \mathring{v}_i^T)((1 - T/\tau)A_{i-1} - (T/\tau)),\\ b_i &= (\mathring{v}_i \mathring{v}_i^T)b_{i-1} + (1_{3\times 3} - \mathring{v}_i \mathring{v}_i^T)(1 - T/\tau)b_{i-1} + 2r_i/\tau,\\ A_i&(\overline{\varpi} + \tilde{\omega}_i) = b_i, \end{aligned} \tag{39}$$

5. Simulation results

5.1. Generalized $(1 + n)$-measurement Triad

As shown in Section 4.1.4, it is possible to apply the geometric method to generalize the Triad method to the situations when we have one primary, and n secondary, measurements. The solution in (28) is verified by simulations for the case when $n = 2$. For this experiment, the primary measurement is of the unit vector in the z-direction of the ground-frame, and has an rms noise of 0.02 rad. The secondary measurements are of $[1 \quad 0 \quad 1]^T/\sqrt{2}$ and $[\sqrt{3}/4 \quad 3/4 \quad 1/2]$ in the ground-frame, and both have an rms noise of 0.2 rad. Since the primary measurement has an accuracy that is an order better than the secondary measurements, a least-squares solution would be almost completely consistent with the primary measurement. Thus, the generalized Triad solution is a simple alternative to the least-squares solution in this scenario (see Fig. 6).

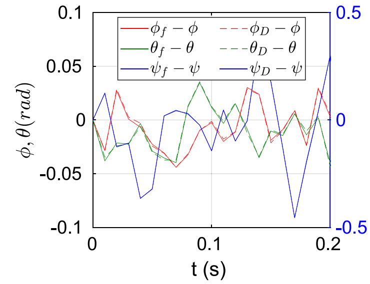

Fig. 6. A generalized Triad solution for several secondary measurements in conjunction with a primary measurement. When the primary measurement is much more accurate than the secondary measurements, the generalized Triad solution of (28) is almost identical to the solution obtained using Davenport's q-method in such situations. Attitude estimation using the generalized Triad is subscripted using an f and plotted using solid lines, while the estimation using the q-method is subscripted by D and plotted using dashed lines.

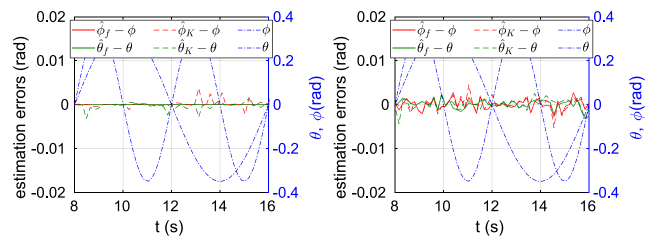

Fig. 7. Filtering using Eqs. (23) or (24) as described in Section 4.2 to obtain a filtered attitude estimate. The roll and pitch angles are prescribed to be sinusoids of amplitude $\pi/9$ rad. Left: The proposed geometric solution yields a perfect roll and pitch estimate ϕ_f and θ_f for a perfect direction measurement using Eqs. (29) and (30). An optimally tuned MEKF develops errors in the update step (ϕ_K and θ_K in the figure) despite the perfect direction measurement. The angular velocity measurement generates 0.04 units rms noise at each time-step. Right: With more realistic noise of 0.001 units in the direction measurement and the same noise through the angular velocity measurement, the inaccuracy in the MEKF becomes less perceptible relative to the other errors. The rms error of the geometric filter is nearly half that of the MEKF (1.1e−6 *vs.* 2e-6 sq-units) even in this case.

5.2. Validation of Theorem 6

The next group of simulations verify the result of Theorem 6, and Remark 6.2.

The attitude estimate $\mathring{q}$ of Theorem 6 can be filtered to reduce the noise, as described in Section 4.2. The improvement with the geometric filter is visually best perceived when there is no noise in the direction measurement (Fig. 7 left). In this situation, the geometric filter determines the required update exactly using (23) or (24). The MEKF suffers from a slight loss of accuracy in the update step when the Kalman gain is determined and the correction linearized (the approximations leading to (158) in Lefferts et al. (1982) or (47) in Markley (2003)). In a more practical situation (Fig. 7 right) with the direction measurement having a noise of 0.001 units, the geometric filter yields half the noise variance (1.1e−6 *vs.* 2e−6 sq-units) as the MEKF, on account of the better algorithmic accuracy in the update step. In both cases, the angular velocity measurement is modelled to contribute a white noise of 0.04 rad each time step. The difference between the MEKF and the proposed filter becomes less significant as the angular velocity measurement becomes less noisy, and the correction quaternion approaches unity. For example, when the noise in the angular velocity measurement is only 0.01 rad each

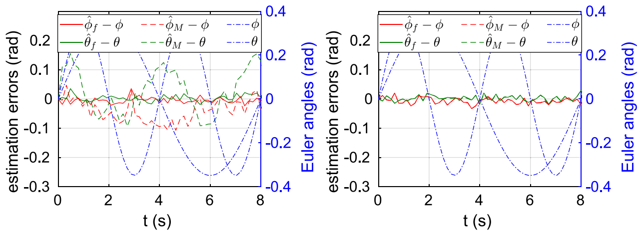

Fig. 8. A comparison of the estimator in Theorem 6 against the ECF in Mahony et al. (2008). The ECF estimate ($\hat{\phi}_M$, $\hat{\theta}_M$) has larger residual errors unless we use the optimal gain suggested in this paper in a two-step estimation. Left: The ECF with gains recommended in Mahony et al. (2008). Right: the ECF using the gain derived in Remark 7.1 in two-step estimation.

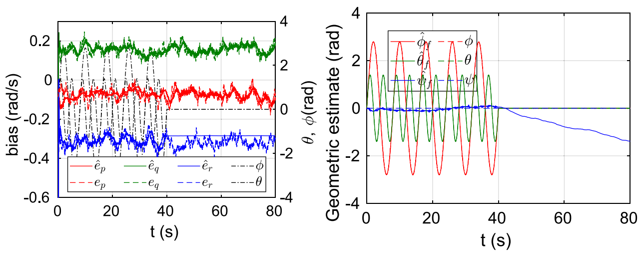

Fig. 9. Left: Failsafe estimation of a time-varying gyroscopic bias using (39). The true time-varying bias is plotted using dashed lines, while the estimates $\hat{e}$ are plotted using solid lines. Excitation is ceased midway through the experiment at $t = 40$ s. The estimator then continues to track the bias orthogonal to the single direction measurement. Right: Failsafe attitude estimation upon loss of excitation. Again, the estimates are shown as solid lines, while the true values are shown as dotted lines.

time step (this case is not shown in above figure), the geometric filter yields almost the same error variance as the MEKF (1.1e−6 *vs.* 1.1e−6 sq-units).

5.3. Optimal gains for the ECF using Remark 7.1

The attitude estimator in Theorem 6 ($\hat{\phi}_f$ and $\hat{\theta}_f$) is compared with the ECF of Mahony et al. (2008) ($\hat{\phi}_M$ and $\hat{\theta}_M$) in Fig. 8. The true attitude angles are denoted ϕ and θ. The angular velocity measurement contributes an rms noise of 0.1 rad/s each time step, while the direction measurement has an rms noise of 0.01 rad. The geometric filter with the optimal gain matrix provides nearly 6-fold superior accuracy compared to the ECF with a gain of 1 (the gain recommended in Mahony et al. (2008)), while the proposed geometric filter uses an optimal gain matrix derived out of the noise parameters and the time step. Equivalent performance may be obtained with both the solutions only upon following a two-step attitude estimation in the ECF, first filtering the direction measurement as described under Section 4.2, and using the gains suggested in Remark 7.1. The two-step estimation is essential so as to ensure that the angular velocity correction ω_c is with respect to the filtered direction measurement v_f obtained from the first step, and that the subsequent vector-measurement based correction is expressed in the body-frame obtained after integrating the angular velocity in the first step.

5.4. Gyroscopic bias estimation

In order to verify the estimation of a time-varying gyroscopic bias as proposed in Section 4.3 we model it as an Ornstein–Uhlenbeck process centred about $[-0.08 \quad 0.16 \quad -0.32]$ rad/s and with an auto-correlation time constant $\tau = 1$ s. The rigid body is rotated in a sinusoidal motion, $\phi = (9\pi/9)\sin(\pi t/4)$ and

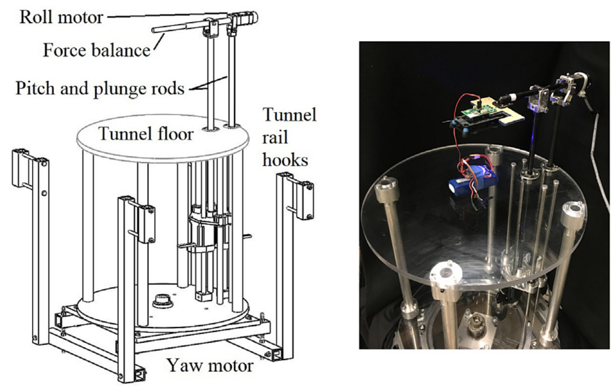

Fig. 10. On the left, a schematic of the 4 Degree of freedom Model Positioning System (MPS) described in Linehan, Shields, and Mohseni (2014). The MPU9250 mounted on the PCB (green in the picture on the right) and being tested on the MPS. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

$\theta = (4\pi/9)\sin(\pi t/2)$ until $t = 40$ s, after which all motion is ceased.

In Fig. 9 on the left, the bias estimates track the low-frequency component in the time-varying bias error. The estimator runs at 100 Hz ($T = 0.01$ s in (39)). In the figure on the right, it can be seen that the estimated attitude does not drift while the body is rotating (persistent excitation). Once rotation ceases, the yaw estimate begins to drift but the roll and pitch estimates continue to track the true values in spite of the bias errors. It may be noted that the estimation algorithm does not break down upon loss of excitation. Instead, it continues to update the bias the best that it can using measurements of a single direction.

6. Experimental validation of geometric attitude estimation using rate and single direction measurement

This section provides experimental verification for the geometric attitude estimator by using a recently developed autopilot in our group, which is equipped with an IMU, the MPU9250, and is described in Bingler and Mohseni (2017). The autopilot is mounted on an inhouse designed model positioning system (MPS) that can independently prescribe roll, pitch, plunge and yaw manoeuvres on a test module (see Fig. 10).

The roll motion has an amplitude of $5\pi/6$ and a period of 4 s. The pitch motion has the same period, and an amplitude of $4\pi/9$. The encoder on the MPS provides the true angles at 1 kHz, while the attitude estimator on the MPU9250 provides estimates at 90 Hz. The estimated roll and pitch angles are plotted along with the true values in Fig. 11. The residual errors in estimating the roll and pitch angles can be attributed to experimental errors. Also shown in the zoomed insets is the high-accuracy, zero latency tracking from the direction measurements to the attitude estimation. This may be compared with the larger errors using the ECF. As shown in Remark 7.1, the ECF is an approximation of the exact geometric estimation that is associated with latency on account of a feedback based correction mechanism. In this experiment, the ECF was used with a gain k_P equal to 1, as suggested in Mahony et al. (2008).

7. Conclusion

We have reported a geometry-based analytic solution for the problem of attitude estimation using two reference direction measurements, and using a rate measurement and a measurement of a single reference direction. The presented approach

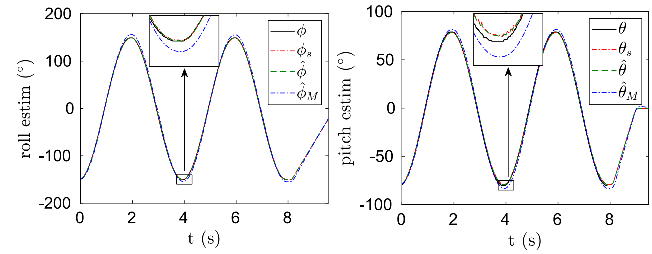

Fig. 11. Top: Attitude estimation for a pure sinusoidal roll manoeuvre on a real system. Bottom: Attitude estimation for a pure sinusoidal pitch manoeuvre on a real system. The solid black lines are the true roll and pitch angles, ϕ and θ, returned by the encoder, the dash–dot red curves are their estimates using the accelerometer directly as $\phi_s = \text{atan2}(g_y/g_z)$ and $\theta_s = \text{asin}(-g_x/\|g\|)$, and the green dashed curves for $\hat{\phi}$ and $\hat{\theta}$ using Theorem 6 presented in this paper after the filtering described in Section 4. The dash–dot blue curve shows the attitude estimate obtained using the ECF with the filter gains suggested in Mahony et al. (2008). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

leads to a unified framework to derive, as special cases,some popular solutions: namely, the Triad solution (Black, 1964), Wahba's formulation (Wahba, 1965), the extended Kalman filter (Lefferts et al., 1982), and the ECF (Mahony et al., 2008). A useful next step would be to extend the geometric solution to consider three direction measurements, and the geometro-kinematic problem to two direction measurements.

References

Andrle, M. S., & Crassidis, J. L. (2015). Attitude estimation employing common frame error representations. *AIAA Journal of Guidance, Control, and Dynamics*, *38*(9), 1614–1624.

Bar-Itzhack, I. Y. (1996). REQUEST: A recursive QUEST algorithm for sequential attitude determination. *AIAA Journal of Guidance, Control, and Dynamics*, *19*(5), 1034–1038.

Batista, P., Silvestre, C., & Oliveira, P. (2011). Partial attitude and rate gyro bias estimation: observability analysis, filter design, and performance evaluation. *Taylor & Francis International Journal of Control*, *84*(5), 895–903.

Batista, P., Silvestre, C., & Oliveira, P. (2012). A GES attitude observer with single vector observations. *Elsevier Automatica*, *48*(2), 388–395.

Berkane, S., & Tayebi, A. (2019). Attitude estimation with intermittent measurements. *Elsevier Automatica*, *105*, 415–421.

Bingler, A., & Mohseni, K. (2017). Dual radio autopilot system for lightweight, swarming micro/miniature aerial vehicles. *AIAA Aerospace Information Systems*, *14*(5), 293–305.

Black, H. D. (1964). A passive system for determining the attitude of a satellite. *American Institute of Aeronautics and Astronautics*, *2*(7), 1350–1351.

Choukroun, D., Bar-Itzhack, I. Y., & Oshman, Y. (2004). Optimal-REQUEST algorithm for attitude determination. *AIAA Journal of Guidance, Control, and Dynamics*, *27*(3), 418–427.

Choukroun, D., Bar-Itzhack, I. Y., & Oshman, Y. (2006). Novel quaternion Kalman filter. *IEEE Transactions on Aerospace and Electronic systems*, *42*(1), 174–190.

Crassidis, J. L., Markley, F. L., & Cheng, Y. (2007). Survey of nonlinear attitude estimation methods. *AIAA Journal of Guidance, Control, and Dynamics*, *30*(1), 12–28.

Dam, E. B., Koch, M., & Lillholm, M. (1998). *Quaternions, interpolation and animation* (2nd ed.). Datalogisk Institut, Københavns Universitet Copenhagen.

Davenport, P. B. (1968). *A vector approach to the algebra of rotations with applications: NASA TN D-4696*.

Farrell, J. L., & Stuelpnagel, J. C. (1966). A least-squares estimate of satellite attitude. *SIAM Review*, *8*(3), 384–386.

Grip, H. F., Fossen, T. I., Johansen, T. A., & Saberi, A. (2012). Attitude estimation using biased gyro and vector measurements with time varying reference vectors. *IEEE Transactions on Automatic Control*, *57*(5), 1332–1338.

Izadi, M., & Sanyal, A. K. (2016). Rigid body pose estimation based on the Lagrange–d'Alembert principle. *Automatica*, *71*, 78–88.

Kalabic, U., Gupta, R., Di Cairano, S., Bloch, A., & Kolmanovsky, I. (2014). *IEEE American control conference* (pp. 5586–5593).

Keat, J. (1977). *Analysis of least-squares attitude determination routine DOAOP: CSC/TM-77/6034*, Comp. Sci. Corp.

Lefferts, E. J., Markley, F. L., & Shuster, M. D. (1982). Kalman filtering for spacecraft attitude estimation. *AIAA Journal of Guidance, Control, and Dynamics*, *5*(5), 417–429.

Linehan, T., Shields, M., & Mohseni, K. (2014). Development, characterization, and validation of a four axis wind tunnel positioning system. In *Proceedings of the AIAA aerospace sciences meeting, 2014-1308*, National Harbor, MD, USA.

Mahony, R., Hamel, T., & Pflimlin, J-M. (2008). Nonlinear complementary filters on the special orthogonal group. *IEEE Transactions on Automatic Control*, *53*(5), 1203–1218.

Markley, F. L. (2003). Attitude error representations for Kalman filtering. *AIAA Journal of Guidance, Dynamics and Control*, *26*(3), 311–317.

Markley, F. L., & Mortari, D. (2000). Quaternion attitude estimation using vector observations. *AIAA Journal of the Astronautical Sciences*, *48*(2), 359–380.

Martin, P., & Sarras, I. (2018). Partial attitude estimation from a single measurement vector. In *IEEE conference on control technology and applications (CCTA)* (pp. 1325–1331).

Mitikiri, Y., & Mohseni, K. (2018). Analytic solutions to two quaternion attitude estimation problems. arXiv, 1901.08905v3.

Mitikiri, Y., & Mohseni, K. (2019a). Compensation of measurement noise and bias in geometric attitude estimation. In *IEEE intl. conference on robotics and automation*, Montreal, Quebec, Canada.

Mitikiri, Y., & Mohseni, K. (2019b). Acceleration compensation for gravity sense using an accelerometer in an aerodynamically stable UAV. In *IEEE intl. conference on decision and control*, Nice, France.

Mortari, D. (1998). Euler *q*-Algorithm for attitude determination using vector observations. *AIAA Journal of Guidance, Control, and Dynamics*, *21*(2) 328–334.

Shuster, M. D., & Oh, S. D. (1981). Three-axis attitude determination from vector observations. *AIAA Journal of Guidance, Control, and Dynamics*, *4*(1), 70–77.

Singh, L., Bortolami, S., & Page, L. (2010). Optimal guidance and thruster control in orbital approach and rendezvous for docking using model predictive control. In *AIAA guidance, navigation, and control conference, Vol. 7754*.

Stovner, B. N., Johansen, T. A., Fossen, T. I., & Schjølberg, I. (2018). Attitude estimation by multiplicative exogenous Kalman filter. *Automatica*, *95*, 347–355.

Tee, K. P., Ge, S. S., & Tay, E. H. (2009). Barrier Lyapunov functions for the control of output-constrained nonlinear systems. *Automatica*, *45*(4), 918–927.

Trumpf, J., Mahony, R., Hamel, T., & Lageman, C. (2012). Analysis of non-linear attitude observers for time-varying reference measurements. *IEEE Transactions on Automatic Control*, *57*(11), 2789–2800.

Wahba, G. 1965. A least-squares estimate of satellite attitude, 7(3) 409–409.

Yun, X., Bachmann, E. R., & McGhee (2008). A simplified quaternion based algorithm for orientation estimation from earth gravity and magnetic field measurements. *IEEE Transactions on Instrumentation and Measurement*, *57*(3), 638–652.

Yujendra Bharathi Mitikiri received his B.Tech. degree in Computer Science and Engineering from the Indian Institute of Technology Madras at Chennai, India, in 2001. He worked as an analog circuit designer at Texas Instruments, Bangalore, India, from 2001 to 2015, and was elected a Senior member technical staff in 2015. He received his M.S. in Aerospace Engineering and Ph.D. in Mechanical Engineering from the University of Florida in 2017 and 2020 respectively.

Kamran Mohseni received his B.S. degree from the University of Science and Technology, Tehran, Iran, his M.S. degree in Aeronautics and Applied Mathematics from the Imperial College of Science, Technology and Medicine, London, U.K., and his Ph.D. degree from the California Institute of Technology (Caltech), Pasadena, CA, USA, in 2000. He was a Postdoctoral Fellow in Control and Dynamical Systems at Caltech for almost a year. In 2001, he joined the Department of Aerospace Engineering Sciences, University of Colorado at Boulder. In 2011, he joined the University of Florida, Gainesville, FL, USA as the W.P. Bushnell Endowed Professor in the Department of Electrical and Computer Engineering and the Department of Mechanical and Aerospace Engineering.