QuateRA: The Quaternion Regression Algorithm*

Marcelino M. de Almeida†
The University of Texas at Austin, Austin, Texas, 78712

Daniele Mortari‡
Texas A&M University, College Station, Texas, 77843

Renato Zanetti§, Maruthi Akella¶
The University of Texas at Austin, Austin, Texas, 78712

This work proposes a batch solution to the problem of estimating fixed angular velocity using orientation measurements. Provided that the angular velocity remains constant with time, we show that the orientation quaternion belongs to a constant plane of rotation as time evolves. Motivated by this fundamental property, we are able to determine the angular velocity's direction by estimating the quaternion plane of rotation. Under the small angle assumption on the attitude measurement noise, the plane of rotation is estimated by minimizing a constrained Total Least Squares cost function, and our algorithm produces a unique optimizing solution through a batch approach (no need for iterations). The angular velocity magnitude is estimated by projecting the measured quaternions onto the estimated plane of rotation, and then computing the least squares evolution of the quaternion angle in the plane. We perform a Monte Carlo analysis of the proposed algorithm, validating our method and comparing it with a Multiplicative Extended Kalman Filter, which is a traditional method in the literature.

I. Introduction

This paper presents a batch solution to the problem of angular velocity estimation using a time-sequence of orientation measurements in terms of the Euler quaternion parameterization. Our approach is motivated by the constant translational velocity estimation problem, whose solution is well known and has well-understood statistical properties [1]. Surprisingly, the rotational counterpart is significantly more challenging and has not yet been solved in a batch estimation sense (to the best of our knowledge). Based on reasonable assumptions for the quaternion noise measurement model, we derive a simple two-step algorithm that establishes a closed-form solution for the constant angular velocity estimation problem without the need to use iterative algorithms.

*A preliminary version of this paper has been presented at the *2019 AAS/AIAA Astrodynamics Specialist Conference*, which was held in August 11-15 in Portland - ME. The previously presented manuscript is numbered as 19-654, and shares the same title as the current manuscript.

†PhD, Dept. of Aerospace Engineering and Engineering Mechanics, The University of Texas at Austin. email: marcelino.malmeidan@utexas.edu.

‡Professor, Dept. of Aerospace Engineering, Texas A&M University, e-mail: mortari@tamu.edu.

§Assistant Professor, Dept. of Aerospace Engineering and Engineering Mechanics, The University of Texas at Austin, email: renato@utexas.edu.

¶E.P. Schoch Professor, Aerospace Engineering and Engineering Mechanics, The University of Texas at Austin, email: makella@mail.utexas.edu.

The problem of estimating the angular velocity under pure spin is a specialized case to the general problem of estimating the angular velocity for a tumbling body. However, the understanding of the pure spin problem aids solving the generalized case assuming that the tumbling motion can be approximated to pure spin throughout a sufficiently short-duration finite sequence of measurements. This kinematic approach can be particularly useful when estimating the angular velocity of a non-cooperative target whose inertia properties and external torques are unknown.

The lack of precise knowledge of the rigid-body's inertia matrix and torque vector also poses a major challenge to standard angular velocity estimation techniques. Many of the existing angular velocity estimators [2, 3] rely on the knowledge of the target's specific inertia and torque parameters. An exception can be made for the *derivative approach* described in Ref. [4], but as the author acknowledges, the angular velocity estimator can produce considerable error due to the presence of measurement noise. In Ref. [5], the authors present the Pseudolinear Kalman Filter (PSELIKA), which does not depend on knowledge of inertia matrix or input torques. However, PSELIKA is developed with the goal of "simplicity rather than accuracy" [5], serving as a relatively coarse angular velocity estimator for control loop damping purposes.

In Ref. [6], generalizations to Wahba's problem are proposed by accepting sequential vector measurements instead of the traditional simultaneous ones (see Ref. [7] and the references therein). These generalizations implicate the need to estimate for initial orientation and angular velocity (not only orientation, as in Wahba's problem). The work of Ref. [6] proposes the following problems:

- First Generalized Wahba's Problem (FGWP) - The system is in pure spin with known spin-axis but unknown spin rate. The author presents a closed-form solution to this problem based on two measurements. The work of Ref. [8] uses semidefinite optimization to solve FGWP for more than two measurements.
- Second Generalized Wahba's Problem (SGWP) - The system is tumbling (torque-free) with known inertia matrix. This system is shown to be observable with at least three vector measurements, but no solution is provided by Ref. [6]. A solution to the three-vector measurement problem is provided in Ref. [9] and a numerical solution is provided in Ref. [10] for four measurements or more.

An alternative solution to the pure spin angular velocity estimation problem is to use methods based on the Multiplicative Extended Kalman Filter (MEKF)[11–13], since these rely on kinematics only. These methods should usually converge if properly initialized and iterated through a backward smoothing process [14]. Iterated nonlinear programming methods present the drawback that these might converge to local minima, unless proven otherwise. Our batch solution in this paper departs from filtering-based ones in that no iterations are necessary for the proposed algorithm.

The primary contribution of this work is the Quaternion Regression Algorithm (QuateRA). Instead of solving the problem through well-stablished filtering approaches[11–13], we solve the problem through a geometrical standpoint. We provide an alternative to an attitude EKF by introducing an attitude regression algorithm. QuateRA builds upon the

work of Ref. [15], and it is a batch solver (does not require iterations), even though the problem is nonlinear. QuateRA uses a sequence of orientation measurements to determine the system's axis of rotation (AOR) through a Singular Value Decomposition (SVD) procedure, and then it uses the AOR to estimate for the angular velocity magnitude (AVM). We develop QuateRA's AOR estimation with use of the Total Least Squares (TLS) cost function, and we are able to provide a solution under mild assumptions on the measurement noise. In fact, the AOR estimation algorithm herein presented shares important similarities with the problem of averaging quaternions [16, 17], but instead of finding an average quaternion, we search for an average quaternion plane. The quaternion average is actually a particular solution to our algorithm. In the current work, we also discuss some asymptotic statistical properties involving QuateRA, apart from validating those results with Monte Carlo simulations.

QuateRA's AOR estimation was first introduced by Ref. [15], and experimental validation was presented in Ref. [18]. Ref. [19] used QuateRA's AOR estimate in conjunction with a modified MEKF to estimate the relative angular velocity of a non-cooperative target. The current work departs from our earlier contributions in the following aspects:

- The previous works used QuateRA's AOR estimation based on heuristics, instead of being a solution that formally minimizes a cost function. In the current work, we start from a constrained version of TLS (the constraints are a direct consequence of the quaternion unit-norm condition), and reach the same solution suggested by Ref. [15] under the assumption of small angle approximation for the quaternion measurement noise.
- None our rpreviously reported results in this field (Refs. [15, 18, 19]) analyzed the statistical properties of QuateRA. In the current work, we explore the strong consistency properties of QuateRA, and we derive covariance matrices for the angular velocity estimation. We also present Monte Carlo analysis to endorse the derived statistical properties.
- When estimating the AVM, Ref. [15] suggested the use of performing "*dirty*" derivatives on the most recently measured quaternions. In contrast, the work of [18] showed that one can often obtain better results by pre-filtering the measured quaternions before employing the derivative. The AVM estimation in Ref. [19] is performed by using a modified MEKF. The AVM estimation suggested by Ref. [15] is actually biased under mild measurement noise, while the solutions presented in Refs. [18] and [19] remedy the bias problem, but introduce tuning parameters. In contrast, this work reprojects the measured quaternions onto the plane of rotation estimated by QuateRA, and calculates the AVM as an average quaternion displacement over time.

The remainder of this paper is organized as follows: Section II introduces the rotational attitude kinematics, describing some notations and parametrizations, as well as the assumed measurement model. Section III presents the estimation problem formulation, introducing the optimization cost function and constraints of the problem. Section IV presents QuateRA, and Section V introduces a Monte Carlo analysis of QuateRA, comparing it with an MEKF formulation, and a solution using a nonlinear solver for the same problem. Finally, Section VI presents conclusions for this work.

II. Attitude Kinematics and Measurement Model

A. Attitude Kinematics

We adopt the notation $\boldsymbol{q}_A^B$ to represent the relative orientation quaternion between frames A and B. A quaternion is written in the form:

$$\boldsymbol{q}_A^B = \begin{bmatrix} q_{As}^B \\ \boldsymbol{q}_{Av}^B \end{bmatrix},$$

where $\boldsymbol{q}_{Av}^B$ and q_{As}^B are the vector and scalar components of the quaternion $\boldsymbol{q}_A^B$, respectively. Also, quaternions satisfy the norm constraint $\left\| \boldsymbol{q}_A^B \right\| = 1$.

We denote the quaternion inverse rotation as $(\boldsymbol{q}_A^B)^{-1} = \boldsymbol{q}_B^A$, which is given by:

$$\boldsymbol{q}_B^A = \begin{bmatrix} q_{As}^B \\ -\boldsymbol{q}_{Av}^B \end{bmatrix}.$$

The quaternion composition rule is denoted as:

$$\boldsymbol{q}_A^C = \boldsymbol{q}_B^C \otimes \boldsymbol{q}_A^B,$$

in which:

$$\boldsymbol{q}_B^C\otimes = \begin{bmatrix} q_{Bs}^C & -(\boldsymbol{q}_{Bv}^C)^T \\ \boldsymbol{q}_{Bv}^C & q_{Bs}^C \boldsymbol{I} - [\boldsymbol{q}_{Bv\times}^C] \end{bmatrix}, \tag{1}$$

where $\boldsymbol{I}$ is the 3×3 identity matrix, and $[\boldsymbol{v}_\times]$ is the skew-symmetric cross product matrix associated with a vector $\boldsymbol{v} \in \mathbb{R}^3$. The matrix $\boldsymbol{q}_B^C\otimes$ is a 4D rotation matrix, implying orthogonality, i.e., it satisfies $\boldsymbol{q}_B^C \otimes (\boldsymbol{q}_B^C\otimes)^T = (\boldsymbol{q}_B^C\otimes)^T \boldsymbol{q}_B^C\otimes = \boldsymbol{I}_4$. Also, we denote the *identity quaternion*:

$$\boldsymbol{q}_I \triangleq (\boldsymbol{q}_A^B)^{-1} \otimes \boldsymbol{q}_A^B = \boldsymbol{q}_A^B \otimes (\boldsymbol{q}_A^B)^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}^T \tag{2}$$

Given a vector $\boldsymbol{v} \in \mathbb{R}^3$, then we define $\boldsymbol{v}\otimes \in \mathbb{R}^{4\times 4}$ as:

$$\boldsymbol{v}\otimes \triangleq \begin{bmatrix} 0 & -\boldsymbol{v}^T \\ \boldsymbol{v} & -[\boldsymbol{v}_\times] \end{bmatrix}.$$

With some slight abuse of notation, we define the composition of a quaternion $\boldsymbol{q} \in \mathbb{S}^3$ with a vector $\boldsymbol{v} \in \mathbb{R}^3$ as:

$$\boldsymbol{q} \otimes \boldsymbol{v} \triangleq \boldsymbol{q} \otimes \begin{bmatrix} 0 \\ \boldsymbol{v} \end{bmatrix}.$$

Alternatively, $\boldsymbol{v}^B$ can be calculated from $\boldsymbol{v}^A$ using the expression $\boldsymbol{v}^B = \boldsymbol{C}_A^B \boldsymbol{v}^A$, where $\boldsymbol{C}_A^B$ is the direction cosine matrix respective to $\boldsymbol{q}_A^B$:

$$\boldsymbol{C}_A^B = \boldsymbol{I} - 2q_{As}^B[\boldsymbol{q}_{Av\times}^B] + 2[\boldsymbol{q}_{Av\times}^B]^2. \tag{3}$$

Denote $\boldsymbol{\omega}_{B/A}^C \in \mathbb{R}^3$ as the angular velocity of frame B w.r.t. frame A expressed in frame C. Then, the rotational kinematics for $\boldsymbol{q}_A^B$ is given by:

$$\dot{\boldsymbol{q}}_A^B = \frac{1}{2}\boldsymbol{\omega}_{B/A}^B \otimes \boldsymbol{q}_A^B. \tag{4}$$

For an angular velocity $\boldsymbol{\omega}_{B/A}^B$, we denote its magnitude $\Omega_{B/A}$ and its direction $\overrightarrow{\boldsymbol{\omega}}_{B/A}^B$, such that:

$$\Omega_{B/A} \triangleq \left\| \boldsymbol{\omega}_{B/A}^B \right\|, \qquad \overrightarrow{\boldsymbol{\omega}}_{B/A}^B \triangleq \frac{\boldsymbol{\omega}_{B/A}^B}{\Omega_{B/A}}.$$

Assuming a constant angular velocity $\boldsymbol{\omega}_{B/A}^B$ throughout a period $\Delta t = t_f - t_0$, then the solution to the kinematic differential equation in Eq. 4 is given by $\boldsymbol{q}_A^B(t_f) = \boldsymbol{F}(\boldsymbol{\omega}_{B/A}^B) \cdot \boldsymbol{q}_A^B(t_0)$, where:

$$\boldsymbol{F}\left(\boldsymbol{\omega}_{B/A}^B\right) = \exp\left[\tfrac{\Delta t}{2}\boldsymbol{\omega}_{B/A}^B \otimes\right] = \cos\frac{\Omega_{B/A}\Delta t}{2} \cdot \boldsymbol{I} + \sin\frac{\Omega_{B/A}\Delta t}{2} \cdot \overrightarrow{\boldsymbol{\omega}}_{B/A}^B \otimes . \tag{5}$$

Using the subscript I to denote inertial frame and O for the frame of the object of interest, the remainder of this paper will denote $\boldsymbol{q}_i \triangleq \boldsymbol{q}_I^O(t_i)$, $\boldsymbol{\omega} \triangleq \boldsymbol{\omega}_{O/I}^O$, $\overrightarrow{\boldsymbol{\omega}} \triangleq \overrightarrow{\boldsymbol{\omega}}_{O/I}^O$, and $\Omega \triangleq \Omega_{O/I}$.

B. Measurement Model

In this section, we present the assumed measurement model for the problem. The assumptions and derivations herein presented are crucial for posing and solving the AOR optimal estimation within QuateRA.

We employ the quaternion measurement model given by:

$$\bar{\boldsymbol{q}}_i = \boldsymbol{q}_i \otimes \boldsymbol{q}_{Ni}, \tag{6}$$

where $\boldsymbol{q}_i = \begin{bmatrix} q_{si} & \boldsymbol{q}_{vi}^T \end{bmatrix}^T$ is the true quaternion and $\boldsymbol{q}_{Ni}$ is the noise quaternion:

$$\boldsymbol{q}_{Ni} \triangleq \begin{bmatrix} \cos\frac{\theta_i}{2} \\ \boldsymbol{e}_{Ni}\sin\frac{\theta_i}{2} \end{bmatrix}, \tag{7}$$

in which θ_i and $\boldsymbol{e}_{Ni}$ are independent random variables. We assume that θ_i is Gaussian (Although it might be unrealistic to assume that angles are distributed as Gaussian, Ref. [20] has shown that this is a reasonable approximation for double-precision machines as long as $\sigma_\theta \leq 22\,\text{deg}$) such that $\theta_i \sim \mathcal{N}(0, \sigma_\theta^2)$, and $\boldsymbol{e}_{Ni} \in \mathbb{S}^2$ is a unit-norm random vector uniformly distributed* in $\mathbb{S}^2 = \{\boldsymbol{x} \in \mathbb{R}^3 : ||\boldsymbol{x}|| = 1\}$ and has the characteristics $\mathbb{E}[\boldsymbol{e}_{Ni}] = \boldsymbol{0}$ and $\mathbb{E}[\boldsymbol{e}_{Ni}\boldsymbol{e}_{Ni}^T] = \frac{1}{3}\boldsymbol{I}$ (see Appendix A).

Assuming that all $\boldsymbol{q}_{Ni}$, $i \in \{1, \cdots, n\}$ are idependent and identically distributed, we define the quantities $\boldsymbol{\mu}_N$ and $\boldsymbol{P}_N$ as the mean and covariance for the noise quaternion, respectively:

$$\begin{aligned}
\boldsymbol{\mu}_N &\triangleq \mathbb{E}\left[\boldsymbol{q}_{Ni}\right] = \mathbb{E}\begin{bmatrix} \cos\frac{\theta_i}{2} \\ \boldsymbol{e}_{Ni}\sin\frac{\theta_i}{2} \end{bmatrix} = \begin{bmatrix} \mathbb{E}\left[\cos\frac{\theta_i}{2}\right] \\ \mathbb{E}\left[\boldsymbol{e}_{Ni}\sin\frac{\theta_i}{2}\right] \end{bmatrix} = \begin{bmatrix} \mathbb{E}\left[\cos\frac{\theta_i}{2}\right] \\ \mathbb{E}\left[\boldsymbol{e}_{Ni}\right]\mathbb{E}\left[\sin\frac{\theta_i}{2}\right] \end{bmatrix} = \mathbb{E}\left[\cos\frac{\theta_i}{2}\right]\begin{bmatrix} 1 \\ \boldsymbol{0} \end{bmatrix} \\
\boldsymbol{P}_N &\triangleq \mathbb{E}\left[\left(\boldsymbol{q}_{Ni} - \boldsymbol{\mu}_N\right)\left(\boldsymbol{q}_{Ni} - \boldsymbol{\mu}_N\right)^T\right] = \mathbb{E}\left[\boldsymbol{q}_{Ni}\boldsymbol{q}_{Ni}^T\right] - \boldsymbol{\mu}_N\boldsymbol{\mu}_N^T \\
&= \begin{bmatrix} \mathbb{E}\left[\cos^2\frac{\theta_i}{2}\right] - \mathbb{E}^2\left[\cos\frac{\theta_i}{2}\right] & \mathbb{E}\left[\boldsymbol{e}_{Ni}^T\cos\frac{\theta_i}{2}\sin\frac{\theta_i}{2}\right] \\ \mathbb{E}\left[\boldsymbol{e}_{Ni}\cos\frac{\theta_i}{2}\sin\frac{\theta_i}{2}\right] & \mathbb{E}\left[\boldsymbol{e}_{Ni}\boldsymbol{e}_{Ni}^T\right]\mathbb{E}\left[\sin^2\frac{\theta_k}{2}\right] \end{bmatrix} = \begin{bmatrix} \mathbb{E}\left[\cos^2\frac{\theta_i}{2}\right] - \mathbb{E}^2\left[\cos\frac{\theta_i}{2}\right] & \boldsymbol{0} \\ \boldsymbol{0} & \frac{1}{3}\mathbb{E}\left[\sin^2\frac{\theta_i}{2}\right]\boldsymbol{I}_3 \end{bmatrix}
\end{aligned}$$

The expected values above can be calculated according with Ref. [20]: $\mathbb{E}\left[\cos\frac{\theta_i}{2}\right] = e^{-\sigma_\theta^2/8}$, $\mathbb{E}\left[\cos^2\frac{\theta_i}{2}\right] = \frac{1}{2}\left(1 + e^{-\sigma_\theta^2/2}\right)$, and $\mathbb{E}\left[\sin^2\frac{\theta_i}{2}\right] = \frac{1}{2}\left(1 - e^{-\sigma_\theta^2/2}\right)$. Defining $\sigma_s^2 \triangleq \mathbb{E}\left[\cos^2\frac{\theta_i}{2}\right] - \mathbb{E}^2\left[\cos\frac{\theta_i}{2}\right] = \frac{1}{2}\left(1 + e^{-\sigma_\theta^2/2}\right) - e^{-\sigma_\theta^2/4}$ and $\sigma_v^2 \triangleq \frac{1}{3}\mathbb{E}\left[\sin^2\frac{\theta_i}{2}\right] = \frac{1}{2}\left(1 - e^{-\sigma_\theta^2/2}\right)$, then the noise covariance matrix takes the form:

$$\boldsymbol{P}_N = \begin{bmatrix} \sigma_s^2 & \boldsymbol{0} \\ \boldsymbol{0} & \sigma_v^2\boldsymbol{I}_3 \end{bmatrix}. \tag{8}$$

*The reader should note that this is a simplification, given that it is not always true that the angle randomness is as likely in any direction. For instance, star trackers tend to have different noise characteristics in the boresight direction w.r.t. the ones perpendicular to it.

We define the covariance for the measured quaternion as:

$$\begin{aligned}\boldsymbol{P}_q &\triangleq \mathbb{E}\left[\left(\bar{\boldsymbol{q}}_i - \mathbb{E}\left[\bar{\boldsymbol{q}}_i\right]\right)\left(\bar{\boldsymbol{q}}_i - \mathbb{E}\left[\bar{\boldsymbol{q}}_i\right]\right)^T\right] = \mathbb{E}\left[\bar{\boldsymbol{q}}_i\bar{\boldsymbol{q}}_i^T\right] - \mathbb{E}\left[\bar{\boldsymbol{q}}_i\right]\mathbb{E}\left[\bar{\boldsymbol{q}}_i\right]^T \\ &= \left(\boldsymbol{q}_i\otimes\right)\mathbb{E}\left[\boldsymbol{q}_{Ni}\boldsymbol{q}_{Ni}^T\right]\left(\boldsymbol{q}_i\otimes\right)^T - \left(\boldsymbol{q}_i\otimes\right)\boldsymbol{\mu}_N\boldsymbol{\mu}_N^T\left(\boldsymbol{q}_i\otimes\right)^T = \left(\boldsymbol{q}_i\otimes\right)\left[\mathbb{E}\left[\boldsymbol{q}_{Ni}\boldsymbol{q}_{Ni}^T\right] - \boldsymbol{\mu}_N\boldsymbol{\mu}_N^T\right]\left(\boldsymbol{q}_i\otimes\right)^T \\ &= \left(\boldsymbol{q}_i\otimes\right)\boldsymbol{P}_N\left(\boldsymbol{q}_i\otimes\right)^T.\end{aligned}$$

If we make the notation relaxation $\boldsymbol{q}_i = \begin{bmatrix} q_s & \boldsymbol{q}_v^T \end{bmatrix}^T$, and use Eqs. 1 and 8, we can further expand $\boldsymbol{P}_q$ as:

$$\boldsymbol{P}_q = \begin{bmatrix} \sigma_s^2 q_s^2 + \sigma_v^2\boldsymbol{q}_v^T\boldsymbol{q}_v & \sigma_s^2 q_s\boldsymbol{q}_v^T - \sigma_v^2 q_s\boldsymbol{q}_v^T \\ \sigma_s^2 q_s\boldsymbol{q}_v - \sigma_v^2 q_s\boldsymbol{q}_v & \sigma_s^2\boldsymbol{q}_v\boldsymbol{q}_v^T - \sigma_v^2\left(q_s^2\boldsymbol{I}_3 - \left[\boldsymbol{q}_{v\times}\right]^2\right) \end{bmatrix}.$$

Using the properties $\left[\boldsymbol{q}_{v\times}\right]^2 = \boldsymbol{q}_v\boldsymbol{q}_v^T - \boldsymbol{q}_v^T\boldsymbol{q}_v\boldsymbol{I}_3$, and $q_s^2 + \boldsymbol{q}_v^T\boldsymbol{q}_v = 1$, we have that:

$$\boldsymbol{P}_q = \begin{bmatrix} \sigma_v^2 + \left(\sigma_s^2 - \sigma_v^2\right)q_s^2 & \left(\sigma_s^2 - \sigma_v^2\right)q_s\boldsymbol{q}_v^T \\ \left(\sigma_s^2 - \sigma_v^2\right)q_s\boldsymbol{q}_v & \sigma_v^2\boldsymbol{I}_3 + \left(\sigma_s^2 - \sigma_v^2\right)\boldsymbol{q}_v\boldsymbol{q}_v^T \end{bmatrix} = \sigma_v^2\boldsymbol{I}_4 + \left(\sigma_s^2 - \sigma_v^2\right)\boldsymbol{q}_i\boldsymbol{q}_i^T. \tag{9}$$

Using the statistics above, if one desires to perform a quaternion measurement normalization, it is necessary to decompose the covariance matrix in the form $\boldsymbol{P}_q = \boldsymbol{L}_q\boldsymbol{L}_q^T$. There are multiple ways of proceeding with the decomposition, but here we derive the *square root* decomposition, i.e., $\boldsymbol{P}_q = \boldsymbol{L}_q\boldsymbol{L}_q$, where $\boldsymbol{L}_q = \boldsymbol{L}_q^T$. Starting from Eq. 9, we add and subtract $2\sigma_v^2\boldsymbol{q}_i\boldsymbol{q}_i^T$ and $2\sigma_v\sigma_s\boldsymbol{q}_i\boldsymbol{q}_i^T$ on the right-hand side of the equation:

$$\begin{aligned}\boldsymbol{P}_q &= \sigma_v^2\boldsymbol{I}_4 - 2\sigma_v^2\boldsymbol{q}_i\boldsymbol{q}_i^T + 2\sigma_v\sigma_s\boldsymbol{q}_i\boldsymbol{q}_i^T + \left(\sigma_s^2 - 2\sigma_v\sigma_s + \sigma_v^2\right)\boldsymbol{q}_i\boldsymbol{q}_i^T \\ &= \sigma_v^2\boldsymbol{I}_4 - 2\sigma_v\left(\sigma_v - \sigma_s\right)\boldsymbol{q}_i\boldsymbol{q}_i^T + \left(\sigma_v - \sigma_s\right)^2\boldsymbol{q}_i\boldsymbol{q}_i^T.\end{aligned}$$

Defining $\sigma_q \triangleq \sigma_v - \sigma_s$ and using the property $\boldsymbol{q}_i\boldsymbol{q}_i^T = \boldsymbol{q}_i\boldsymbol{q}_i^T\boldsymbol{q}_i\boldsymbol{q}_i^T$ then:

$$\begin{aligned}\boldsymbol{P}_q &= \sigma_v^2\boldsymbol{I}_4 - 2\sigma_v\sigma_q\boldsymbol{q}_i\boldsymbol{q}_i^T + \sigma_q^2\boldsymbol{q}_i\boldsymbol{q}_i^T\boldsymbol{q}_i\boldsymbol{q}_i^T = \sigma_v^2\boldsymbol{I}_4 - 2\sigma_v\sigma_q\boldsymbol{q}_i\boldsymbol{q}_i^T + \sigma_q^2\left(\boldsymbol{q}_i\boldsymbol{q}_i^T\right)^2 \\ &= \left(\sigma_v\boldsymbol{I}_4 - \sigma_q\boldsymbol{q}_i\boldsymbol{q}_i^T\right)^2\end{aligned} \tag{10}$$

Therefore, the matrix square-root of $\boldsymbol{P}_q$ is given by $\boldsymbol{L}_q = \sigma_v\boldsymbol{I}_4 - \sigma_q\boldsymbol{q}_i\boldsymbol{q}_i^T$, where $\sigma_q = \sigma_v - \sigma_s$. The inverse of the

square-root matrix is given by:

$$\boldsymbol{L}_q^{-1} = \frac{1}{\sigma_s \sigma_v}\left(\sigma_s \boldsymbol{I}_4 + \sigma_q \boldsymbol{q}_i \boldsymbol{q}_i^T\right).$$

Post-multiplying $\boldsymbol{L}_q^{-1}$ by $\boldsymbol{q}_i$, we get that:

$$\boldsymbol{L}_q^{-1}\boldsymbol{q}_i = \frac{1}{\sigma_s \sigma_v}\left(\sigma_s \boldsymbol{q}_i + \sigma_q \boldsymbol{q}_i\right) = \frac{\sigma_v}{\sigma_s \sigma_v}\boldsymbol{q}_i = \frac{1}{\sigma_s}\boldsymbol{q}_i.$$

Therefore, $\boldsymbol{q}_i$ is an eigenvector of $\boldsymbol{L}_q^{-1}$, and the corresponding eigenvalue is given by $\lambda_q = 1/\sigma_s$. Having that in mind, if we perform a Taylor Expansion on Eq. 6 around $\theta_i = 0$, and pre-multiply by $\boldsymbol{L}_q^{-1}$, we get that:

$$\begin{aligned}
\boldsymbol{L}_q^{-1}\bar{\boldsymbol{q}}_i &= \boldsymbol{L}_q^{-1}\boldsymbol{q}_i \otimes \boldsymbol{q}_{Ni} = \boldsymbol{L}_q^{-1}\left(\boldsymbol{q}_i\otimes\right)\left(\boldsymbol{q}_I + \left.\frac{\partial \boldsymbol{q}_{Ni}}{\partial \theta_i}\right|_0 \theta_i + \left.\frac{\partial^2 \boldsymbol{q}_{Ni}}{\partial \theta_i^2}\right|_0 \theta_i^2 + \cdots\right) \\
&= \boldsymbol{L}_q^{-1}\boldsymbol{q}_i + \boldsymbol{L}_q^{-1}\left(\boldsymbol{q}_i\otimes\right)\left(\left.\frac{\partial \boldsymbol{q}_{Ni}}{\partial \theta_i}\right|_0 \theta_i + \left.\frac{\partial^2 \boldsymbol{q}_{Ni}}{\partial \theta_i^2}\right|_0 \theta_i^2 + \cdots\right) \\
&= \frac{1}{\sigma_s}\boldsymbol{q}_i + \boldsymbol{L}_q^{-1}\left(\boldsymbol{q}_i\otimes\right)\left(\left.\frac{\partial \boldsymbol{q}_{Ni}}{\partial \theta_i}\right|_0 \theta_i + \left.\frac{\partial^2 \boldsymbol{q}_{Ni}}{\partial \theta_i^2}\right|_0 \theta_i^2 + \cdots\right),
\end{aligned}$$

where $\boldsymbol{q}_I$ is the identity quaternion defined in Eq. 2.

Therefore, if we consider only the *0-th* order approximation for the measurement normalization performed by the operation $\boldsymbol{L}_q^{-1}\bar{\boldsymbol{q}}_k$, then this operation is just a scaling operation on the true quaternion. In practice, it is impossible to perform the measurement normalization $\boldsymbol{L}_q^{-1}\bar{\boldsymbol{q}}_i$ because $\boldsymbol{L}_q$ is a function of the true quaternion $\boldsymbol{q}_i$ (not the measured one), which is unknown. Alternatively, if we make the practical approximation [21]:

$$\boldsymbol{P}_q \approx \left(\sigma_v \boldsymbol{I}_4 - \sigma_q \bar{\boldsymbol{q}}_i \bar{\boldsymbol{q}}_i^T\right)^2 \qquad \Longrightarrow \qquad \boldsymbol{L}_q^{-1} \approx \frac{1}{\sigma_s \sigma_v}\left(\sigma_s \boldsymbol{I}_4 + \sigma_q \bar{\boldsymbol{q}}_i \bar{\boldsymbol{q}}_i^T\right), \tag{11}$$

then the measurement normalization leads to $\boldsymbol{L}_q^{-1}\bar{\boldsymbol{q}}_i = \lambda_q \bar{\boldsymbol{q}}_i$.

III. Problem Formulation

This section poses the problem that we solve with QuateRA. Let $\bar{\boldsymbol{q}}_i$ be a quaternion measurement at time t_i, and assume that we have n measurements. We assume that the measured quaternions represent rotations relative to an inertial frame, hence we pursue an angular velocity estimate that is relative to the same inertial frame (and expressed in

the inertial frame). Defining $\hat{\boldsymbol{q}}_i \in \mathbb{S}^3$ as the quaternion estimate at time t_i, we want to minimize the cost function:

$$J_{LS} = \frac{1}{2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{q}}_i - \bar{\boldsymbol{q}}_i\right)^T \boldsymbol{P}_q^{-1}\left(\hat{\boldsymbol{q}}_i - \bar{\boldsymbol{q}}_i\right) = \frac{1}{2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{q}}_i - \bar{\boldsymbol{q}}_i\right)^T \boldsymbol{L}_q^{-T}\boldsymbol{L}_q^{-1}\left(\hat{\boldsymbol{q}}_i - \bar{\boldsymbol{q}}_i\right)$$
$$= \frac{1}{2}\sum_{i=1}^{n}\left(\boldsymbol{L}_q^{-1}\hat{\boldsymbol{q}}_i - \boldsymbol{L}_q^{-1}\bar{\boldsymbol{q}}_i\right)^T \left(\boldsymbol{L}_q^{-1}\hat{\boldsymbol{q}}_i - \boldsymbol{L}_q^{-1}\bar{\boldsymbol{q}}_i\right).$$

Making the practical approximations $\boldsymbol{L}_q^{-1}\bar{\boldsymbol{q}}_i \approx \lambda_q \bar{\boldsymbol{q}}_i$ and $\boldsymbol{L}_q^{-1}\hat{\boldsymbol{q}}_i \approx \lambda_q \hat{\boldsymbol{q}}_i$, then:

$$J_{LS} \approx \frac{1}{2}\lambda_q^2 \sum_{i=1}^{n}\left\|\hat{\boldsymbol{q}}_i - \bar{\boldsymbol{q}}_i\right\|_2^2. \tag{12}$$

Dropping the constant gain λ_q from the cost function (as it shouldn't impact the optimal solution), and using the property $\bar{\boldsymbol{q}}_i^T\bar{\boldsymbol{q}}_i = \hat{\boldsymbol{q}}_i^T\hat{\boldsymbol{q}}_i = 1$, then J_{LS} can be further simplified as:

$$J_{LS} = \frac{1}{2}\sum_{i=1}^{n}\left\|\hat{\boldsymbol{q}}_i - \bar{\boldsymbol{q}}_i\right\|_2^2 = \frac{1}{2}\sum_{i=1}^{n}\left(\hat{\boldsymbol{q}}_i - \bar{\boldsymbol{q}}_i\right)^T\left(\hat{\boldsymbol{q}}_i - \bar{\boldsymbol{q}}_i\right) = \frac{1}{2}\sum_{i=1}^{n}\left(2 - 2\hat{\boldsymbol{q}}_i^T\bar{\boldsymbol{q}}_i\right) = n - \sum_{i=1}^{n}\hat{\boldsymbol{q}}_i^T\bar{\boldsymbol{q}}_i. \tag{13}$$

additionally, the optimal estimation problem has to be subject to the quaternion kinematic equation of Eq. 4:

$$\dot{\boldsymbol{q}} = \frac{1}{2}\boldsymbol{\omega}\otimes\boldsymbol{q}.$$

Assuming that $\boldsymbol{\omega} = \Omega\overrightarrow{\omega}$ is constant, the rotational kinematics is solved by using the state transition matrix of Eq. 5:

$$\boldsymbol{q}(t) = \left[\cos\tfrac{\Omega\Delta t}{2}\cdot\boldsymbol{I} + \sin\tfrac{\Omega\Delta t}{2}\cdot\overrightarrow{\omega}\otimes\right]\boldsymbol{q}_0 = \cos\frac{\Omega\Delta t}{2}\cdot\boldsymbol{q}_0 + \sin\frac{\Omega\Delta t}{2}\cdot\overrightarrow{\omega}\otimes\boldsymbol{q}_0, \tag{14}$$

where $\Delta t \triangleq t - t_0$. In summary, we are searching for estimates of $\hat{\omega} = \hat{\Omega}\hat{\overrightarrow{\omega}}$, and $\hat{\boldsymbol{q}}_i$ satisfying:

$$\begin{cases} \min_{\hat{\omega},\hat{\boldsymbol{q}}_i} & J_{LS} = n - \sum_{i=1}^{n}\hat{\boldsymbol{q}}_i^T\bar{\boldsymbol{q}}_i \\ s.t. & \hat{\boldsymbol{q}}_{i+1} = \cos\frac{\hat{\Omega}\delta_i}{2}\cdot\hat{\boldsymbol{q}}_i + \sin\frac{\hat{\Omega}\delta_i}{2}\cdot\hat{\overrightarrow{\omega}}\otimes\hat{\boldsymbol{q}}_i, \ \forall\ i\in\{1,\cdots,n-1\} \\ & ||\hat{\boldsymbol{q}}_i|| = 1, \qquad\qquad \forall\ i\in\{1,\cdots,n\} \end{cases}, \tag{15}$$

where $\delta_i \triangleq t_{i+1} - t_i$.

As mentioned in the Introduction, QuateRA is a two step algorithm: it first estimates the AOR $\hat{\overrightarrow{\omega}}$, and then uses its knowledge to estimate for the AVM $\hat{\Omega}$. In order to estimate the AOR, QuateRA uses a geometric interpretation based on the solution to the quaternion kinematic equation of Eq. 14.

Defining the vectors $\boldsymbol{u}_1 \in \mathbb{S}^3 = \boldsymbol{q}_0$ and $\boldsymbol{u}_2 \in \mathbb{S}^3 = \overrightarrow{\boldsymbol{\omega}} \otimes \boldsymbol{q}_0$, we have that $\boldsymbol{u}_1^T \boldsymbol{u}_2 = \boldsymbol{q}_0 \cdot \overrightarrow{\boldsymbol{\omega}} \otimes \cdot \boldsymbol{q}_0$. Since $\overrightarrow{\boldsymbol{\omega}}\otimes$ is a skew-symmetric matrix (see Eq. II.A) then $\boldsymbol{u}_1^T \boldsymbol{u}_2 = 0$, i.e., $\boldsymbol{u}_1 \perp \boldsymbol{u}_2$. Defining $\alpha(t) \triangleq \frac{\Omega \Delta t}{2}$, we can write Eq. 14 as:

$$\boldsymbol{q}(t) = \cos\alpha(t) \cdot \boldsymbol{u}_1 + \sin\alpha(t) \cdot \boldsymbol{u}_2. \tag{16}$$

Clearly, any $\boldsymbol{q}(t)$ described by Eq. 16 is a linear combination of $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$, for all $t \in \mathbb{R}$. Hence, if we define the 4D hyperplane $\mathbb{P}(\boldsymbol{u}_1, \boldsymbol{u}_2) = \text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$, then $\boldsymbol{q}(t) \in \mathbb{P}(\boldsymbol{u}_1, \boldsymbol{u}_2)$, $\forall\, t \in \mathbb{R}$. Thus, the optimally estimated quaternions should belong to a single plane of rotation: $\hat{\boldsymbol{q}}_i \in \mathbb{P}(\boldsymbol{u}_1, \boldsymbol{u}_2),\ \forall i \in \{1, \cdots, n\}$. In addition, we have that $\boldsymbol{\omega} = \boldsymbol{u}_2 \otimes \boldsymbol{u}_1^{-1}$.

Therefore, if we have a sequence of measurements $\bar{\boldsymbol{q}}_i$, $i \in \{1, \cdots, n\}$, with $n \in \mathbb{N}_{\geq 2}$, then we can estimate the axis of rotation by finding the optimal hyperplane that fits the measured quaternions. Classically speaking, plane-fitting is a Total Least Squares (TLS) problem [22]. We define $\hat{\boldsymbol{q}}_i^{TLS}$ as the TLS best in-plane estimate for the i-th measurement. Defining the matrices $\bar{\boldsymbol{Q}}$ and $\hat{\boldsymbol{Q}}^{TLS}$ as:

$$\bar{\boldsymbol{Q}} \triangleq \begin{bmatrix} \bar{\boldsymbol{q}}_1 & \bar{\boldsymbol{q}}_2 & \cdots & \bar{\boldsymbol{q}}_n \end{bmatrix}, \qquad \hat{\boldsymbol{Q}}^{TLS} \triangleq \begin{bmatrix} \hat{\boldsymbol{q}}_1^{TLS} & \hat{\boldsymbol{q}}_2^{TLS} & \cdots & \hat{\boldsymbol{q}}_n^{TLS} \end{bmatrix}, \tag{17}$$

then our plane-fitting problem can be cast in the following TLS form:

$$\begin{cases} \min_{\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2, \hat{\boldsymbol{q}}_i} & J_{TLS} = \left\| \hat{\boldsymbol{Q}}^{TLS} - \bar{\boldsymbol{Q}} \right\|_F^2 \\ s.t. & \hat{\boldsymbol{q}}_i^{TLS} \in \mathbb{P}(\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2), \quad \forall\, i \in \{1, \cdots, n\} \\ & ||\hat{\boldsymbol{q}}_i^{TLS}|| = 1, \qquad \forall\, i \in \{1, \cdots, n\} \end{cases}, \tag{18}$$

where the $||\cdot||_F$ denotes the *Frobenius norm*. Notice that the optimization problem of Eq. 18 is related to the classical TLS problem, except for the unit norm constraint $||\hat{\boldsymbol{q}}_i^{TLS}|| = 1$, and the fact that the solution is a 2D space (instead of a vector). Hence, although we start from a TLS cost function for estimating the quaternion plane of rotation, the solution is not related to the textbook solutions on TLS.

Once we solve the optimization problem of Eq. 18, we are able to obtain estimates for the axis of rotation $\hat{\overrightarrow{\boldsymbol{\omega}}}$, the plane of rotation $\mathbb{P}(\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2)$, and the quaternion estimates $\hat{\boldsymbol{q}}_i^{TLS} \in \mathbb{P}(\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2)$. Given those estimates, we recast the

optimization problem of Eq. 15 as:

$$\begin{cases} \min_{\hat{\Omega},\hat{\boldsymbol{q}}_i} & J_{LS} = n - \sum_{i=1}^{n} \hat{\boldsymbol{q}}_i^T \hat{\boldsymbol{q}}_i^{TLS} \\ s.t. & \hat{\boldsymbol{q}}_{i+1} = \cos\frac{\hat{\Omega}\delta_i}{2} \cdot \hat{\boldsymbol{q}}_i + \sin\frac{\hat{\Omega}\delta_i}{2} \cdot \widehat{\vec{\omega}} \otimes \hat{\boldsymbol{q}}_i, \ \forall\ i \in \{1, \cdots, n-1\} \\ & ||\hat{\boldsymbol{q}}_i|| = 1, \qquad \forall\ i \in \{1, \cdots, n\} \\ & \hat{\boldsymbol{q}}_i \in \mathbb{P}(\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2), \qquad \forall\ i \in \{1, \cdots, n\} \end{cases}, \tag{19}$$

In a nutshell, the problem of Eq. 18 is solved by taking the Singular Value Decomposition on $\boldsymbol{Z} \triangleq \bar{\boldsymbol{Q}}\bar{\boldsymbol{Q}}^T$, whose two first left singular vectors determine the quaternion plane of rotation. The AOR direction is uniquely identified from the plane of rotation. In order to solve the problem of Eq. 19, we observe that a unit quaternion $\boldsymbol{q}_i$ on a plane can be uniquely identified by a single angle Φ_i. Hence, if we assume that this angle is evolving linearly as in $\Phi_i = \Phi_1 + \Omega\Delta t$, we can perform least squares to solve for optimal $\hat{\Phi}_1$ and $\hat{\Omega}$ that determine the quaternion evolution on that plane. QuateRA's algorithm is summarized in Section IV.D.

IV. The Quaternion Regression Algorithm

In this section, we develop the QuateRA algorithm. The remainder of this section is structured as follows: Section IV.A derives the AOR estimation algorithm, while Section IV.B derives the AVM estimator. A method for estimating the covariance matrix is given in Section IV.C. Section IV.D summarizes QuateRA into a few steps, and Section IV.E presents some insights and analysis to the overall algorithm.

A. Estimation of the Axis of Rotation

In order to estimate the AOR, the goal is to find a plane $\hat{\mathbb{P}}(\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2) = \text{span}\{\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2\}$ and a set of estimated quaternions $\hat{\boldsymbol{q}}_i \in \hat{\mathbb{P}}(\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2)$, $i \in \{1, \cdots, n\}$ that minimizes the TLS cost function:

$$J_1 = \frac{1}{2}\left\|\bar{\boldsymbol{Q}} - \hat{\boldsymbol{Q}}^{TLS}\right\|_F^2. \tag{20}$$

In order to reduce heavy notation, the remainder of this subsection will denote $\hat{\boldsymbol{Q}} \equiv \hat{\boldsymbol{Q}}^{TLS}$ and $\hat{\boldsymbol{q}}_i \equiv \hat{\boldsymbol{q}}_i^{TLS}$.

Starting from the definition of $\bar{\boldsymbol{Q}}$ in Eq. 17, we can derive the following property:

$$\text{tr}\left(\bar{\boldsymbol{Q}}\bar{\boldsymbol{Q}}^T\right) = \text{tr}\left(\sum_{i=1}^{n} \bar{\boldsymbol{q}}_i\bar{\boldsymbol{q}}_i^T\right) = \sum_{i=1}^{n} \text{tr}\left(\bar{\boldsymbol{q}}_i\bar{\boldsymbol{q}}_i^T\right) = \sum_{i=1}^{n} \left\|\bar{\boldsymbol{q}}_i\right\|^2 = n \tag{21}$$

From the Frobenius norm definition, we have that:

$$J_1 = \frac{1}{2}\,\mathrm{tr}\left[\left(\bar{\boldsymbol{Q}} - \hat{\boldsymbol{Q}}\right)\left(\bar{\boldsymbol{Q}} - \hat{\boldsymbol{Q}}\right)^T\right] = \frac{1}{2}\,\mathrm{tr}\left[\bar{\boldsymbol{Q}}\bar{\boldsymbol{Q}}^T - \bar{\boldsymbol{Q}}\hat{\boldsymbol{Q}}^T - \hat{\boldsymbol{Q}}\bar{\boldsymbol{Q}}^T + \hat{\boldsymbol{Q}}\hat{\boldsymbol{Q}}^T\right]$$
$$= \frac{1}{2}\,\mathrm{tr}\left(\bar{\boldsymbol{Q}}\bar{\boldsymbol{Q}}^T\right) - \frac{1}{2}\,\mathrm{tr}\left(\bar{\boldsymbol{Q}}\hat{\boldsymbol{Q}}^T\right) - \frac{1}{2}\,\mathrm{tr}\left(\hat{\boldsymbol{Q}}\bar{\boldsymbol{Q}}^T\right) + \frac{1}{2}\,\mathrm{tr}\left(\hat{\boldsymbol{Q}}\hat{\boldsymbol{Q}}^T\right). \tag{22}$$

Using the trace property $\mathrm{tr}(\boldsymbol{AB}) = \mathrm{tr}(\boldsymbol{BA})$, and the property of Eq. 21, we have that:

$$J_1 = n - \mathrm{tr}\left(\bar{\boldsymbol{Q}}\hat{\boldsymbol{Q}}^T\right) = n - \mathrm{tr}\left(\textstyle\sum_{i=1}^{n} \bar{\boldsymbol{q}}_i\hat{\boldsymbol{q}}_i^T\right) = n - \displaystyle\sum_{i=1}^{n} \mathrm{tr}\left(\bar{\boldsymbol{q}}_i\hat{\boldsymbol{q}}_i^T\right) = n - \sum_{i=1}^{n} \bar{\boldsymbol{q}}_i^T\,\hat{\boldsymbol{q}}_i. \tag{23}$$

Minimizing the cost function of Eq. 23 is equivalent to *maximizing* the following cost function:

$$J_2 = \sum_{i=1}^{k} \bar{\boldsymbol{q}}_i^T\,\hat{\boldsymbol{q}}_i. \tag{24}$$

Theorem 1 *Given a quaternion $\boldsymbol{q} \in \mathbb{S}^3$ and a plane spanned by the unit vectors $\boldsymbol{u}_1 \in \mathbb{S}^3$ and $\boldsymbol{u}_2 \in \mathbb{S}^3$ such that $\boldsymbol{u}_1^T\boldsymbol{u}_2 = 0$. Denoting this plane as $\mathbb{P}(\boldsymbol{u}_1, \boldsymbol{u}_2)$, the quaternion $\boldsymbol{q}_p \in \mathbb{S}^3$ that belongs to the plane $\mathbb{P}(\boldsymbol{u}_1, \boldsymbol{u}_2)$ and minimizes the cost function:*

$$J_0 = \frac{1}{2}\left\|\boldsymbol{q} - \boldsymbol{q}_p\right\|_2^2 = \frac{1}{2}\left\|\boldsymbol{q} - \boldsymbol{q}_p\right\|_F^2 \tag{25}$$

is given by:

$$\boldsymbol{q}_p = \frac{1}{\sqrt{\left(\boldsymbol{q}^T\boldsymbol{u}_1\right)^2 + \left(\boldsymbol{q}^T\boldsymbol{u}_2\right)^2}}\left[\left(\boldsymbol{q}^T\boldsymbol{u}_1\right)\boldsymbol{u}_1 + \left(\boldsymbol{q}^T\boldsymbol{u}_2\right)\boldsymbol{u}_2\right] \tag{26}$$

Proof. The cost function of Eq. 25 can be written as:

$$J_0 = \frac{1}{2}\left\|\boldsymbol{q} - \boldsymbol{q}_p\right\|_2^2 = \frac{1}{2}\left(\boldsymbol{q}^T\boldsymbol{q} - 2\boldsymbol{q}^T\boldsymbol{q}_p + \boldsymbol{q}_p^T\boldsymbol{q}_p\right) = 1 - \boldsymbol{q}^T\boldsymbol{q}_p. \tag{27}$$

Minimizing the cost function of Eq. 27 is the same as maximizing the following cost function:

$$J_1 = \boldsymbol{q}^T\boldsymbol{q}_p. \tag{28}$$

Every quaternion that belongs to the plane $\mathbb{P}(\boldsymbol{u}_1, \boldsymbol{u}_2)$ can be written as a linear combination of $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$:

$$\boldsymbol{q}_p = a\boldsymbol{u}_1 + b\boldsymbol{u}_2. \tag{29}$$

In order to satisfy the norm condition for $||\boldsymbol{q}_p|| = 1$, the following holds:

$$||\boldsymbol{q}_p|| = \boldsymbol{q}_p^T\boldsymbol{q}_p = a^2\boldsymbol{u}_1^T\boldsymbol{u}_1 + 2ab\boldsymbol{u}_1^T\boldsymbol{u}_2 + b^2\boldsymbol{u}_2^T\boldsymbol{u}_2 = a^2 + b^2 = 1$$

Hence, the coefficients a and b from Eq. 29 are constrained such that $a^2 + b^2 = 1$. We rewrite the optimization problem as:

$$\begin{cases} \max_{a,b} J_1 = \boldsymbol{q}^T\boldsymbol{q}_p = a\boldsymbol{q}^T\boldsymbol{u}_1 + b\boldsymbol{q}^T\boldsymbol{u}_2 \\ s.t. \quad a^2 + b^2 = 1. \end{cases} \tag{30}$$

Introducing the Lagrange multiplier λ, the Lagragian related to the problem above is written as:

$$\mathcal{L} = a\boldsymbol{q}^T\boldsymbol{u}_1 + b\boldsymbol{q}^T\boldsymbol{u}_2 + \frac{1}{2}\lambda(a^2 + b^2 - 1) \quad \implies \begin{cases} \frac{\partial\mathcal{L}}{\partial a} = \boldsymbol{q}^T\boldsymbol{u}_1 + \lambda a \\ \frac{\partial\mathcal{L}}{\partial b} = \boldsymbol{q}^T\boldsymbol{u}_2 + \lambda b \end{cases}.$$

From the first-order necessary optimality conditions, we get that:

$$\boldsymbol{q}^T\boldsymbol{u}_1 + \lambda a = 0 \implies a = -\frac{\boldsymbol{q}^T\boldsymbol{u}_1}{\lambda}, \qquad \boldsymbol{q}^T\boldsymbol{u}_2 + \lambda b = 0 \implies b = -\frac{\boldsymbol{q}^T\boldsymbol{u}_2}{\lambda}. \tag{31}$$

Substituting a and b from Eq. 31 into $a^2 + b^2 = 1$, we get that:

$$\frac{(\boldsymbol{q}^T\boldsymbol{u}_1)^2}{\lambda^2} + \frac{(\boldsymbol{q}^T\boldsymbol{u}_2)^2}{\lambda^2} = 1 \implies \lambda = \pm\sqrt{(\boldsymbol{q}^T\boldsymbol{u}_1)^2 + (\boldsymbol{q}^T\boldsymbol{u}_2)^2}. \tag{32}$$

Therefore, we have that:

$$a = -\frac{\boldsymbol{q}^T\boldsymbol{u}_1}{\lambda} = \pm\frac{1}{\sqrt{(\boldsymbol{q}^T\boldsymbol{u}_1)^2 + (\boldsymbol{q}^T\boldsymbol{u}_2)^2}}\boldsymbol{q}^T\boldsymbol{u}_1, \qquad b = -\frac{\boldsymbol{q}^T\boldsymbol{u}_2}{\lambda} = \pm\frac{1}{\sqrt{(\boldsymbol{q}^T\boldsymbol{u}_1)^2 + (\boldsymbol{q}^T\boldsymbol{u}_2)^2}}\boldsymbol{q}^T\boldsymbol{u}_2. \tag{33}$$

We can notice that this problem has two extremum points: a maximizing solution and a minimizing one. By

inspecting the cost function in Eq. 30, the maximizing solution has to be the one given by:

$$\begin{cases} a = \frac{1}{\sqrt{(\boldsymbol{q}^T\boldsymbol{u}_1)^2+(\boldsymbol{q}^T\boldsymbol{u}_2)^2}}\boldsymbol{q}^T\boldsymbol{u}_1 \\ b = \frac{1}{\sqrt{(\boldsymbol{q}^T\boldsymbol{u}_1)^2+(\boldsymbol{q}^T\boldsymbol{u}_2)^2}}\boldsymbol{q}^T\boldsymbol{u}_2 \end{cases} \quad \Longrightarrow \quad \boldsymbol{q}_p = \frac{1}{\sqrt{\left(\boldsymbol{q}^T\boldsymbol{u}_1\right)^2 + \left(\boldsymbol{q}^T\boldsymbol{u}_2\right)^2}} \left[\left(\boldsymbol{q}^T\boldsymbol{u}_1\right)\boldsymbol{u}_1 + \left(\boldsymbol{q}^T\boldsymbol{u}_2\right)\boldsymbol{u}_2\right].$$

□

Using Theorem 1, then $\hat{\boldsymbol{q}}$ can be written as a linear combination of the optimal plane vectors $\hat{\boldsymbol{u}}_1$ and $\hat{\boldsymbol{u}}_2$. Hence, the cost function J_2 from Eq. 24 can be written as:

$$\begin{aligned} J_2 &= \sum_{i=1}^{n} \frac{1}{\sqrt{\left(\bar{\boldsymbol{q}}_i^T\hat{\boldsymbol{u}}_1\right)^2 + \left(\bar{\boldsymbol{q}}_i^T\hat{\boldsymbol{u}}_2\right)^2}} \bar{\boldsymbol{q}}_i^T \left[\left(\bar{\boldsymbol{q}}_i^T\hat{\boldsymbol{u}}_1\right)\hat{\boldsymbol{u}}_1 + \left(\bar{\boldsymbol{q}}_i^T\hat{\boldsymbol{u}}_2\right)\hat{\boldsymbol{u}}_2\right] = \sum_{i=1}^{n} \frac{1}{\sqrt{\left(\bar{\boldsymbol{q}}_i^T\hat{\boldsymbol{u}}_1\right)^2 + \left(\bar{\boldsymbol{q}}_i^T\hat{\boldsymbol{u}}_2\right)^2}} \left[\left(\bar{\boldsymbol{q}}_i^T\hat{\boldsymbol{u}}_1\right)^2 + \left(\bar{\boldsymbol{q}}_i^T\hat{\boldsymbol{u}}_2\right)^2\right] \\ &= \sum_{i=1}^{n} \sqrt{\left(\bar{\boldsymbol{q}}_i^T\hat{\boldsymbol{u}}_1\right)^2 + \left(\bar{\boldsymbol{q}}_i^T\hat{\boldsymbol{u}}_2\right)^2} \end{aligned} \tag{34}$$

Note that in the total absence of measurement noise, and assuming $\hat{\boldsymbol{u}}_1 \in \text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$, $\hat{\boldsymbol{u}}_2 \in \text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ with $\hat{\boldsymbol{u}}_1^T\hat{\boldsymbol{u}}_2 = 0$, the following holds:

$$\sqrt{\left(\bar{\boldsymbol{q}}_i^T\hat{\boldsymbol{u}}_1\right)^2 + \left(\bar{\boldsymbol{q}}_i^T\hat{\boldsymbol{u}}_2\right)^2} = 1, \quad \forall\, i \in \{1, \cdots, n\}.$$

Defining the variable $x \triangleq \left(\bar{\boldsymbol{q}}_i^T\hat{\boldsymbol{u}}_1\right)^2 + \left(\bar{\boldsymbol{q}}_i^T\hat{\boldsymbol{u}}_2\right)^2$, the First order Taylor Expansion of $\sqrt{x}$ around $x = 1$ is given by:

$$\sqrt{x} \approx \frac{1}{2} + \frac{x}{2} \quad \Longrightarrow \quad \sqrt{\left(\bar{\boldsymbol{q}}_i^T\hat{\boldsymbol{u}}_1\right)^2 + \left(\bar{\boldsymbol{q}}_i^T\hat{\boldsymbol{u}}_2\right)^2} \approx \frac{1}{2} + \frac{1}{2}\left(\bar{\boldsymbol{q}}_i^T\hat{\boldsymbol{u}}_1\right)^2 + \frac{1}{2}\left(\bar{\boldsymbol{q}}_i^T\hat{\boldsymbol{u}}_2\right)^2$$

Therefore, under the small angle approximation for the measurement noise, we have that the cost function J_2 can be approximated to:

$$J_2 \approx \frac{n}{2} + \frac{1}{2}\sum_{i=1}^{n}\left[\left(\bar{\boldsymbol{q}}_i^T\hat{\boldsymbol{u}}_1\right)^2 + \left(\bar{\boldsymbol{q}}_i^T\hat{\boldsymbol{u}}_2\right)^2\right] \tag{35}$$

For simplicity of notation, we define a new cost function whose maximization is equivalent to the maximization of

Eq. 35:

$$J = \sum_{i=1}^{n} \left[\left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1 \right)^2 + \left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2 \right)^2 \right] = \sum_{i=1}^{n} \left[\hat{\boldsymbol{u}}_1^T \bar{\boldsymbol{q}}_i \bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1 + \hat{\boldsymbol{u}}_2^T \bar{\boldsymbol{q}}_i \bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2 \right] = \hat{\boldsymbol{u}}_1^T \sum_{i=1}^{n} \bar{\boldsymbol{q}}_i \bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1 + \hat{\boldsymbol{u}}_2^T \sum_{i=1}^{n} \bar{\boldsymbol{q}}_i \bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2 \tag{36}$$

$$= \hat{\boldsymbol{u}}_1^T \bar{\boldsymbol{Q}} \bar{\boldsymbol{Q}}^T \hat{\boldsymbol{u}}_1 + \hat{\boldsymbol{u}}_2^T \bar{\boldsymbol{Q}} \bar{\boldsymbol{Q}}^T \hat{\boldsymbol{u}}_2 \tag{37}$$

Defining $\bar{\boldsymbol{Z}} \triangleq \bar{\boldsymbol{Q}} \bar{\boldsymbol{Q}}^T$, the optimization problem can be stated in the following form:

$$\begin{cases} \max\limits_{\hat{\boldsymbol{u}}_1 \in \mathbb{S}^3, \hat{\boldsymbol{u}}_2 \in \mathbb{S}^3} \hat{\boldsymbol{u}}_1^T \bar{\boldsymbol{Z}} \hat{\boldsymbol{u}}_1 + \hat{\boldsymbol{u}}_2^T \bar{\boldsymbol{Z}} \hat{\boldsymbol{u}}_2 \\ \text{s.t. } \hat{\boldsymbol{u}}_1^T \hat{\boldsymbol{u}}_2 = 0 \end{cases} . \tag{38}$$

The optimization problem of Eq. 38 does not admit a unique solution. This should be an obvious statement, since there are infinitely many pairs of orthogonal vectors that define a plane. Still, this is not an issue for QuateRA, since the axis of rotation direction can be uniquely determined from the hyperplane, regardless of which particular optimal solution has been obtained for $\hat{\boldsymbol{u}}_1$ and $\hat{\boldsymbol{u}}_2$. Lemma 2 introduces a particular optimal solution to the problem above.

Lemma 2 *A solution to the optimization problem in Eq. 38 can be obtained from the Singular Value Decomposition (SVD) of* $\bar{\boldsymbol{Z}} = \hat{\boldsymbol{U}} \hat{\boldsymbol{\Sigma}} \hat{\boldsymbol{U}}^T$, *where* $\hat{\boldsymbol{U}} \in \mathbb{R}^{4\times4} = \begin{bmatrix} \hat{\boldsymbol{u}}_1 & \hat{\boldsymbol{u}}_2 & \hat{\boldsymbol{u}}_3 & \hat{\boldsymbol{u}}_4 \end{bmatrix}$ *contains the singular vectors of* $\bar{\boldsymbol{Z}}$, *and* $\hat{\boldsymbol{\Sigma}} = \text{diag}\left(\hat{\sigma}_1, \hat{\sigma}_2, \hat{\sigma}_3, \hat{\sigma}_4 \right)$ *contains the singular values of* $\bar{\boldsymbol{Z}}$, *wherein* $\hat{\sigma}_1 \geq \hat{\sigma}_2 \geq \hat{\sigma}_3 \geq \hat{\sigma}_4 \geq 0$. *If* $\hat{\sigma}_2 > \hat{\sigma}_3$, *then* $\hat{\boldsymbol{u}}_1$ *and* $\hat{\boldsymbol{u}}_2$ *compose a solution to the optimization problem in Eq. 38 and the optimal cost is given by* $J^*(\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2) = \hat{\sigma}_1 + \hat{\sigma}_2$, *with* $\hat{\sigma}_1 = \hat{\boldsymbol{u}}_1^T \bar{\boldsymbol{Z}} \hat{\boldsymbol{u}}_1$ *and* $\hat{\sigma}_2 = \hat{\boldsymbol{u}}_2^T \bar{\boldsymbol{Z}} \hat{\boldsymbol{u}}_2$.

Proof. This follows from common knowledge in SVD, as the best-fit k-dimensional subspace for a matrix is the subspace spanned by the first k singular vectors [23]. As we are looking for a 2-dimensional subspace that best approximates $\bar{\boldsymbol{Z}}$, then the solution to the optimization problem of Eq. 38 is given by the first two left singular vectors of $\bar{\boldsymbol{Z}}$. □

Having the optimal hyperplane estimate $\hat{\mathbb{P}}(\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2)$, we still need to calculate the AOR $\hat{\overrightarrow{\boldsymbol{\omega}}}$ that leads to rotation on that plane. As previously observed in Eq. 14, the optimal hyperplane can be written as $\hat{\mathbb{P}}(\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2) = \hat{\mathbb{P}}(\hat{\boldsymbol{u}}_1, \overrightarrow{\boldsymbol{\omega}} \otimes \hat{\boldsymbol{u}}_1)$. This implies that $\hat{\boldsymbol{u}}_2 = \hat{\overrightarrow{\boldsymbol{\omega}}} \otimes \hat{\boldsymbol{u}}_1$. Therefore, the optimal estimate for the AOR is given by:

$$\hat{\overrightarrow{\boldsymbol{\omega}}} = \hat{\boldsymbol{u}}_2 \otimes \hat{\boldsymbol{u}}_1^{-1}. \tag{39}$$

An important observation is that $\hat{\overrightarrow{\boldsymbol{\omega}}}$ is an ambiguous estimate of $\overrightarrow{\boldsymbol{\omega}}$ up to a sign error, i.e, it estimates the direction of $\overrightarrow{\boldsymbol{\omega}}$, but the sense might be wrong. This ambiguity is eliminated when estimating the AVM Ω, whose estimate $\hat{\Omega}$ will be negative when $\hat{\overrightarrow{\boldsymbol{\omega}}}$ is an estimate of $-\overrightarrow{\boldsymbol{\omega}}$. In any case, the product $\hat{\boldsymbol{\omega}} = \hat{\Omega} \hat{\overrightarrow{\boldsymbol{\omega}}}$ is consistent with $\boldsymbol{\omega} = \Omega \overrightarrow{\boldsymbol{\omega}}$.

Using the result from Theorem 1, the optimally estimated quaternions on the plane $\hat{\mathbb{P}}(\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2)$ are given by:

$$\hat{\boldsymbol{q}}_i^{TLS} = \frac{1}{\sqrt{\left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1\right)^2 + \left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2\right)^2}} \left[\left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1\right) \hat{\boldsymbol{u}}_1 + \left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2\right) \hat{\boldsymbol{u}}_2\right]. \tag{40}$$

B. Estimation of the Angular Velocity Magnitude

In this section, we use the estimated in-plane quaternions $\hat{\boldsymbol{q}}_i^{TLS} \in \mathbb{P}(\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2)$ to solve the optimization problem of Eq. 19, where $\hat{q}_i^{TLS}$ is given by Eq. 40.

We make the observation that a unit quaternion belonging to a plane $\mathbb{P}(\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2)$ can be fully specified simply by an angle on that plane. We define the quaternion angle on $\mathbb{P}(\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2)$ as having a zero-angle when aligned with $\boldsymbol{u}_1$, and it grows positive as the quaternion rotates from $\boldsymbol{u}_1$ towards $\boldsymbol{u}_2$. We make the definitions:

$$\hat{\Phi}_i = 2 \cdot \text{atan2}\left(\hat{\boldsymbol{u}}_2^T \hat{\boldsymbol{q}}_i,\ \ \hat{\boldsymbol{u}}_1^T \hat{\boldsymbol{q}}_i\right), \qquad \bar{\Phi}_i = 2 \cdot \text{atan2}\left(\hat{\boldsymbol{u}}_2^T \hat{\boldsymbol{q}}_i^{TLS},\ \ \hat{\boldsymbol{u}}_1^T \hat{\boldsymbol{q}}_i^{TLS}\right), \tag{41}$$

where $\bar{\Phi}_i$ is the respective angle of the quaternion $\hat{\boldsymbol{q}}_i^{TLS}$, and $\hat{\Phi}_i$ is the angle of the quaternion that we are estimating $\hat{\boldsymbol{q}}_i$.

If we define $\psi_i \triangleq \hat{\Phi}_i - \bar{\Phi}_i$, then we have that $\hat{\boldsymbol{q}}_i^T \hat{\boldsymbol{q}}_i^{TLS} = \cos\frac{|\psi|}{2}$. Using Taylor series around the origin, we can approximate $\hat{\boldsymbol{q}}_i^T \hat{\boldsymbol{q}}_i^{TLS} \approx 1 - \frac{\psi^2}{8}$. Hence, for sufficiently small ψ (i.e. low noise characteristics), the cost function of Eq. 19 can be approximated as:

$$J_{LS} = n - \sum_{i=1}^{n} \hat{\boldsymbol{q}}_i^T \hat{\boldsymbol{q}}_i^{TLS} \approx n - n + \frac{1}{8}\sum_{i=1}^{n} \psi^2 = \frac{1}{8}\sum_{i=1}^{n} (\hat{\Phi}_i - \bar{\Phi}_i)^2. \tag{42}$$

If we assume the system model:

$$\Phi_i = \Phi_1 + \Omega \Delta t_i = \begin{bmatrix} 1 & \Delta t_i \end{bmatrix} \begin{bmatrix} \Phi_0 \\ \Omega \end{bmatrix},$$

and the measurement model:

$$\bar{\Phi}_i = \Phi_i + \nu_i,$$

where ν_i is the measurement noise such that $\mathbb{E}[\nu_i] = 0$, $\mathbb{E}[\nu_i \nu_j] = 0$, $i \neq j$, and $\mathbb{E}[\nu_i^2] = P_\nu$, then we can use least

squares to estimate for $\hat{\Phi}_0$ and $\hat{\Omega}$:

$$\hat{\boldsymbol{X}} \triangleq \begin{bmatrix} \hat{\Phi}_1 \\ \hat{\Omega} \end{bmatrix} = \left(\boldsymbol{H}^T \boldsymbol{H} \right)^{-1} \boldsymbol{H}^T \hat{\boldsymbol{\Phi}}, \tag{43}$$

where:

$$\boldsymbol{H} \triangleq \begin{bmatrix} 1 & \cdots & 1 \\ \Delta t_1 & \cdots & \Delta t_n \end{bmatrix}^T, \qquad \hat{\boldsymbol{\Phi}} \triangleq \begin{bmatrix} \hat{\Phi}_1 & \cdots & \hat{\Phi}_n \end{bmatrix}^T. \tag{44}$$

Given that the measurement noise assumed to be uncorrelated between two measurements ($\mathbb{E}[\nu_i \nu_j] = 0,\ i \neq j$), then the covariance matrix of the estimate $\hat{\boldsymbol{X}}$ is given by $\text{cov}[\hat{\boldsymbol{X}}] = P_\nu \left(\boldsymbol{H}^T \boldsymbol{H} \right)^{-1}$.

The optimally estimated quaternions $\hat{\boldsymbol{q}}_i \in \mathbb{P}(\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2)$ can be retrieved as:

$$\hat{\boldsymbol{q}}_{i+1} = \cos \frac{\hat{\Omega}\delta_i}{2} \cdot \hat{\boldsymbol{q}}_i + \sin \frac{\hat{\Omega}\delta_i}{2} \cdot \hat{\vec{\omega}} \otimes \hat{\boldsymbol{q}}_i,\ \forall\ i \in \{1, \cdots, n-1\}. \tag{45}$$

Note that the quaternion estimates in Eq. 45 satisfies all the constraints of the optimization problem in Eq. 19.

Theorem 3 below proves that the noise ν_i is actually zero mean and that $P_\nu = \frac{1}{3}\sigma_\theta^2$, where σ_θ is the noise standard deviation for the measurement noise as defined in Eqs. 6 and 7.

Theorem 3 *Assume that* $\boldsymbol{q}_N = \begin{bmatrix} \cos \frac{\theta}{2} & \boldsymbol{e}_N^T \sin \frac{\theta}{2} \end{bmatrix}^T$ *is a noise quaternion, where* θ *is a zero-mean gaussian random variable with* $\mathbb{E}[\theta^2] = \sigma_\theta^2$*, and* $\boldsymbol{e}_N \in \mathbb{R}^2$ *is a unit vector uniformly distributed in the 3D sphere. Also, define a plane* $\mathbb{P}(\boldsymbol{q}_I, \boldsymbol{q}_v)$ *as the hyperplane spanned by the unit vectors* $\boldsymbol{q}_I$ *(identity quaternion) and* $\boldsymbol{q}_v \triangleq \begin{bmatrix} 0 & \boldsymbol{v}^T \end{bmatrix}^T$ *with* $\boldsymbol{v} \in \mathbb{S}^2$ *such that* $\boldsymbol{q}_v^T \boldsymbol{q}_I = 0$*. Now, assume that* $\boldsymbol{q}_{Np} \in \mathbb{P}(\boldsymbol{q}_I, \boldsymbol{q}_v)$ *is the quaternion that belongs to* $\mathbb{P}(\boldsymbol{q}_I, \boldsymbol{q}_v)$ *and is closest to* $\boldsymbol{q}_N$ *such as in Theorem 1. Then, if we assume the small angle approximation on* $\theta = 0$*, the quaternion* $\boldsymbol{q}_{Np}$ *has the form:*

$$\boldsymbol{q}_{Np} = \begin{bmatrix} \cos \frac{\Phi}{2} \\ \boldsymbol{v} \sin \frac{\Phi}{2} \end{bmatrix}, \tag{46}$$

where Φ *has the approximate statistics* $\mathbb{E}[\Phi] = 0$*, and* $\sigma_\Phi^2 \triangleq \mathbb{E}[\Phi^2] = \frac{1}{3}\sigma_\theta^2$.

Proof. According with Theorem 1, $\boldsymbol{q}_{NP}$ is given by:

$$\hat{\boldsymbol{q}}_{NP} = \frac{1}{\sqrt{\left(\boldsymbol{q}_N^T \boldsymbol{q}_I\right)^2 + \left(\boldsymbol{q}_N^T \boldsymbol{q}_v\right)^2}} \left[\left(\boldsymbol{q}_N^T \boldsymbol{q}_I\right) \boldsymbol{q}_I + \left(\boldsymbol{q}_N^T \boldsymbol{q}_v\right) \boldsymbol{q}_v\right] = \frac{1}{\sqrt{\left(\boldsymbol{q}_N^T \boldsymbol{q}_I\right)^2 + \left(\boldsymbol{q}_N^T \boldsymbol{q}_v\right)^2}} \begin{bmatrix} \boldsymbol{q}_N^T \boldsymbol{q}_I \\ \boldsymbol{v} \cdot \boldsymbol{q}_N^T \boldsymbol{q}_v \end{bmatrix} \tag{47}$$

Comparing Eq. 47 with Eq. 46, we get that:

$$\cos\frac{\Phi}{2} = \frac{\boldsymbol{q}_N^T \boldsymbol{q}_I}{\sqrt{\left(\boldsymbol{q}_N^T \boldsymbol{q}_I\right)^2 + \left(\boldsymbol{q}_N^T \boldsymbol{q}_v\right)^2}} \tag{48}$$

From the definition of the identity quaternion (Eq. 2), we get that $\boldsymbol{q}_N^T \boldsymbol{q}_I = \cos\frac{\theta}{2}$. In addition, we have that $\boldsymbol{q}_N^T \boldsymbol{q}_v = \boldsymbol{e}_N^T \boldsymbol{v} \sin\frac{\theta}{2}$. Defining γ as the angle between the vectors $\boldsymbol{e}_N^T$ and $\boldsymbol{v}$, then we can define $\cos\gamma \triangleq \boldsymbol{e}_N^T \boldsymbol{v}$. Given that $\boldsymbol{e}_N$ is uniformly distributed in a 3D sphere, then Appendix A shows that $\cos\gamma \sim \mathcal{U}[-1, 1]$. Therefore, we have that $\boldsymbol{q}_N^T \boldsymbol{q}_v = \cos\gamma \sin\frac{\theta}{2}$. Plugging these values into Eq. 48, and performing Taylor series expansion on both sides around $\Phi = 0$ and $\theta = 0$, we get to:

$$\cos\frac{\Phi}{2} = \frac{\cos\frac{\theta}{2}}{\sqrt{\cos^2\frac{\theta}{2} + \cos^2\gamma \sin^2\frac{\theta}{2}}} \qquad \text{(Taylor Series on both sides)}$$

$$1 - \frac{\Phi^2}{8} \approx 1 - \cos^2\gamma \frac{\theta^2}{8} \tag{49}$$

Inspecting Eq. 49, we can approximate $\Phi \approx \theta \cdot \cos\gamma$. Therefore, we have that $\mathbb{E}[\Phi] = \mathbb{E}[\theta]\mathbb{E}[\cos\gamma] = 0$ and $\mathbb{E}[\Phi^2] = \mathbb{E}[\theta^2]\mathbb{E}[\cos^2\gamma] = \frac{1}{3}\sigma_\theta^2$. □

C. Covariance Estimate

This section presents a covariance estimate for the estimated angular velocity through a Fisher Information approach. We compute how much information is added to the estimates when a new orientation measurement is processed. We base our Information propagation on the MEKF equations derived in Appendix B.

Assuming that the attitude error is in the Gibbs vector format (see Eq. 64), we define the estimation error vector as $\boldsymbol{X} = \begin{bmatrix} \delta\boldsymbol{g}^T & \delta\boldsymbol{\omega}^T \end{bmatrix}^T$, where $\delta\boldsymbol{\omega} \triangleq \hat{\boldsymbol{\omega}} - \boldsymbol{\omega}$. Each orientation measurement has the error covariance $R = \frac{1}{3}\sigma_\theta^2 \boldsymbol{I}_3$. Defining the covariance matrix $\boldsymbol{P}_X = \mathbb{E}[\boldsymbol{X}\boldsymbol{X}^T]$, the related Fisher information matrix is given by $\mathbb{I} = \boldsymbol{P}_X^{-1}$.

Assuming that one orientation measurement has been already processed, the information matrix can be initialized as:

$$\mathbb{I}_1 = \begin{pmatrix} \boldsymbol{R}^{-1} & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 \end{pmatrix}. \tag{50}$$

The information for all subsequent measurement updates can be processed iteratively as:

$$\mathbb{I}_{k+1} = \boldsymbol{\Gamma}_k^T \mathbb{I}_k \boldsymbol{\Gamma}_k + \boldsymbol{H}_k \boldsymbol{R} \boldsymbol{H}_k^T, \qquad \forall\ k \in \{1, \cdots, n-1\}, \tag{51}$$

where $\boldsymbol{H}_k = \begin{bmatrix} \boldsymbol{I}_3 & \mathbf{0}_3 \end{bmatrix}$, $\boldsymbol{A}_d[k] \triangleq e^{-A\delta_k}$, $\delta_k \triangleq t_{k+1} - t_k$, and:

$$\boldsymbol{A} = \begin{bmatrix} -[\boldsymbol{\omega}_\times] & \boldsymbol{I}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 \end{bmatrix}.$$

Finally, the estimated final covariance matrix is given by $\hat{\boldsymbol{P}}_X = \mathbb{I}_n^{-1}$.

D. Algorithm Summary

In this section, we summarize the algorithm steps for QuateRA.

1) Construct the measurement matrix $\bar{\boldsymbol{Q}}$ as in Eq. 17 and calculate $\bar{\boldsymbol{Z}} = \bar{\boldsymbol{Q}}\bar{\boldsymbol{Q}}^T$.
2) Compute the SVD $\bar{\boldsymbol{Z}} = \hat{\boldsymbol{U}}\hat{\boldsymbol{\Sigma}}\hat{\boldsymbol{U}}^T$. The plane of rotation is defined by the first two columns of $\hat{\boldsymbol{U}} = \begin{bmatrix} \hat{\boldsymbol{u}}_1 & \hat{\boldsymbol{u}}_2 & \hat{\boldsymbol{u}}_3 & \hat{\boldsymbol{u}}_4 \end{bmatrix}$.
3) The optimal axis of rotation is defined as in Eq. 39: $\hat{\boldsymbol{\omega}} = \hat{\boldsymbol{u}}_2 \otimes \hat{\boldsymbol{u}}_1^{-1}$.
4) Compute the optimally estimated quaternions $\hat{\boldsymbol{q}}_i, i \in \{1, \cdots, n\}$ on the plane $\hat{\mathbb{P}}(\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2)$ using Eq. 40.
5) For each quaternion $\hat{\boldsymbol{q}}_i$ on the plane $\hat{\mathbb{P}}(\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2)$, compute the quaternion angle within the plane $\hat{\boldsymbol{\Phi}}_i$ using Eq. 41.
6) Estimate the angular velocity $\hat{\boldsymbol{\Omega}}$ and its associated covariance using Eqs. 43 and 44. Note that the angles $\bar{\boldsymbol{\Phi}}$ need to be unwrapped before performing the least squares estimation.
7) Initialize the Fisher information matrix as in Eq. 50, and update through Eq. 51. Compute the final covariance as $\hat{\boldsymbol{P}}_X = \mathbb{I}_n^{-1}$.

E. QuateRA Analysis

In this section, we provide some critical analysis and insights about the derivation of QuateRA.

We have converted the initial optimization problem of Eq. 15 into two subproblems: one that estimates the AOR by estimating the quaternion plane of rotation (Eq. 18), and then we use the plane of rotation knowledge to estimate for the AVM (Eq. 19). The items below provide a critical view on our derivations and solutions:

- When estimating the plane of rotation, we employ a TLS cost function. Classical formulations of TLS leads to strongly consistent estimates, i.e., the TLS estimate converges to the true value (with probability 1) as the number of measurements n tend to infinity [21, 24], meaning that it is asymptotically unbiased. Monte Carlo analysis suggest that the bias of classical TLS is statistically negligible when signal-to-noise ratio is high, and n is large [22].
- Although we have employed the TLS cost function, the optimization problem is not a classical TLS problem, as we constrain the optimized variables to be unit norm. Hence, we cannot affirm that all TLS statistical properties are transferred to QuateRA.
- The Least Squares estimate of the AVM $||\boldsymbol{\Omega}||$ assumes that the velocity direction $\overrightarrow{\boldsymbol{\omega}}$ is precisely known. However, as already mentioned, TLS can provide a biased estimate $\overrightarrow{\bar{\boldsymbol{v}}} = \boldsymbol{u}_1$, which can also implicate on a biased estimation of $||\boldsymbol{\Omega}||$.
- Many portions of our derivations assume sufficiently small measurement noise. This implicates that QuateRA might not be a reasonable estimator for problems with too large orientation measurement noise.

Given the concerns above, Section V presents a Monte Carlo analysis of QuateRA, comparing its results with a Multiplicative Extended Kalman Filter. The Monte Carlo results indicate that QuateRA carry the *strong consistency* property of classical TLS, and it even outperforms MEKF in some situations, specially for situations with large angular velocities and low sampling frequency. On the other hand, MEKF seems to be a slightly better estimator for high sampling frequencies and small angular velocities. Note, however, that the average discrepancy between QuateRA and MEKF disappear as the number of measurements increase.

A few important remarks that should be noted on QuateRA are highlighted below:

- When $n = 2$, QuateRA computes the solution that leads to $J_{LS} = 0$, i.e., $n = 2$ leads to a perfect fit of the data.
- It doesn't matter if $\bar{\boldsymbol{Q}}$ is constructed with $\boldsymbol{q}_i$ or $-\boldsymbol{q}_i$. Plane-fitting is agnostic to the quaternion direction, and the AVM estimation is not affected as long as the angles are unwrapped prior to solving the LS problem.
- The quaternion averaging problem described in Ref. [17] is a special solution for the problem herein presented. Note the similarity between the cost function in Eq. 36 with respect to Eq. 12 within Ref. [17] when all the weights are unity. This implies that $\hat{\boldsymbol{u}}_1$ has the geometric meaning of an *average quaternion* among all the measurements.

V. QuateRA Monte Carlo Analysis

This section provides a Monte Carlo analysis of QuateRA, endorsing the statistical properties derived in the previous sections, as well as providing a comparison with an MEKF (see Appendix B for referencing the used formulation). We perform extensive simulations for multiple values of n (number of measurements) and σ_θ (standard deviation for the angle in the noise quaternion).

In all simulations, we used an angular velocity with direction $\overrightarrow{\omega} = \frac{1}{\sqrt{14}}\begin{bmatrix}1 & 2 & 3\end{bmatrix}^T$. The standard deviation for the measurement noise are chosen as $\sigma_\theta = 1°$, $\sigma_\theta = 2°$, $\sigma_\theta = 3°$, $\sigma_\theta = 4°$, and $\sigma_\theta = 5°$ (large values, when compared to star-tracker technology. The analysis of this section would be quite uninteresting for σ_θ values expected for Star Trackers, since QuateRA's performance would not change much as a function of the number of measurements n). The number of measurements range from $n = 5$ to $n = 50$ in increments of 5. Each Monte Carlo result is obtained after $n_{MC} = 10000$ executions. We denote $\overrightarrow{\omega}_\perp \in \mathbb{S}^2$ as an arbitrary unit vector perpendicular to $\overrightarrow{\omega}$, i.e., $\overrightarrow{\omega}^T\overrightarrow{\omega}_\perp = 0$.

In order to evaluate the AOR estimation, we calculate the mean and standard deviation of the estimated AOR $\hat{\overrightarrow{\omega}}$ along $\overrightarrow{\omega}_\perp$. Defining $\hat{\overrightarrow{\omega}}_i^T$ as the estimation of $\overrightarrow{\omega}$ at the i^{th} Monte Carlo trial, and $e_{i\perp} \triangleq \hat{\overrightarrow{\omega}}_i^T\overrightarrow{\omega}_\perp$ as the respective projected error, then the mean $\mu_\perp$ and variance $\sigma_\perp^2$ for $e_{i\perp}$ is calculated as:

$$\mu_\perp \triangleq \frac{1}{n_{MC}}\sum_{i=1}^{n_{MC}} e_{i\perp}, \qquad \sigma_\perp^2 \triangleq \frac{1}{n_{MC}-1}\sum_{i=1}^{n_{MC}}\left(e_{i\perp}-\mu_\perp\right)^2. \tag{52}$$

A sample mean around $\mu_\perp = 0$ indicates that the AOR is an unbiased estimator. The standard deviation has to belong to the range $0 < \sigma_\perp \leq 1/\sqrt{3} \approx 0.5774$, where $\sigma_\perp \to 1/\sqrt{3}$ indicates that the estimator is obtaining solutions uniformly distributed in the unit sphere (see Appendix A). In our experience, the AOR estimator provides acceptable estimates when $\sigma_\perp \leq 0.1$.

In order to evaluate the AVM estimation, we define the AVM error as $e_{i\Omega} \triangleq \hat{\Omega}_i - \Omega$, where $\hat{\Omega}_i$ is the estimated AVM for the i^{th} Monte Carlo execution. We calculate the mean μ_Ω and variance σ_Ω^2 of $e_{i\Omega}$ as:

$$\mu_\Omega \triangleq \frac{1}{n_{MC}}\sum_{i=1}^{n_{MC}} e_{i\Omega}, \qquad \sigma_\Omega^2 \triangleq \frac{1}{n_{MC}-1}\sum_{i=1}^{n_{MC}}\left(e_{i\Omega}-\mu_\Omega\right)^2. \tag{53}$$

First, we evaluate QuateRA's performance in a degenerate scenario. We start with measurements taken at 10hz, with an AVM of $\Omega = 0.1$rad/s. Notice that when the measurement is as high as $5° = 0.0873$rad and the number of measurements are as low as $n = 5$, the change in orientation throughout that period is of 0.05rad, and hence the signal to noise ratio is extremely low for accurately estimating the angular velocity. Figure 1 presents the Monte Carlo results for the AOR estimation, indicating that the estimator is asymptotically unbiased and that the standard deviations decrease as the number of measurements increase. Figure 2 shows that the mean error μ_Ω converges to zero as the number of measurements n increase. The standard deviation also decreases as n increases. One should be aware that these solutions only make sense if the AOR make sense, i.e., if $\sigma_\perp$ is small enough.

Given the estimate error for the i-th Monte Carlo execution as $\boldsymbol{\omega}_{ei} = \hat{\boldsymbol{\omega}} - \boldsymbol{\omega}$, we compute the sample standard deviation on $\boldsymbol{\omega}_{ei}$, defined as $\boldsymbol{\sigma}_\omega = \begin{bmatrix}\boldsymbol{\sigma}_{\omega x} & \boldsymbol{\sigma}_{\omega y} & \boldsymbol{\sigma}_{\omega z}\end{bmatrix}^T$. We compare $\boldsymbol{\sigma}_\omega$ with the standard deviations estimated in

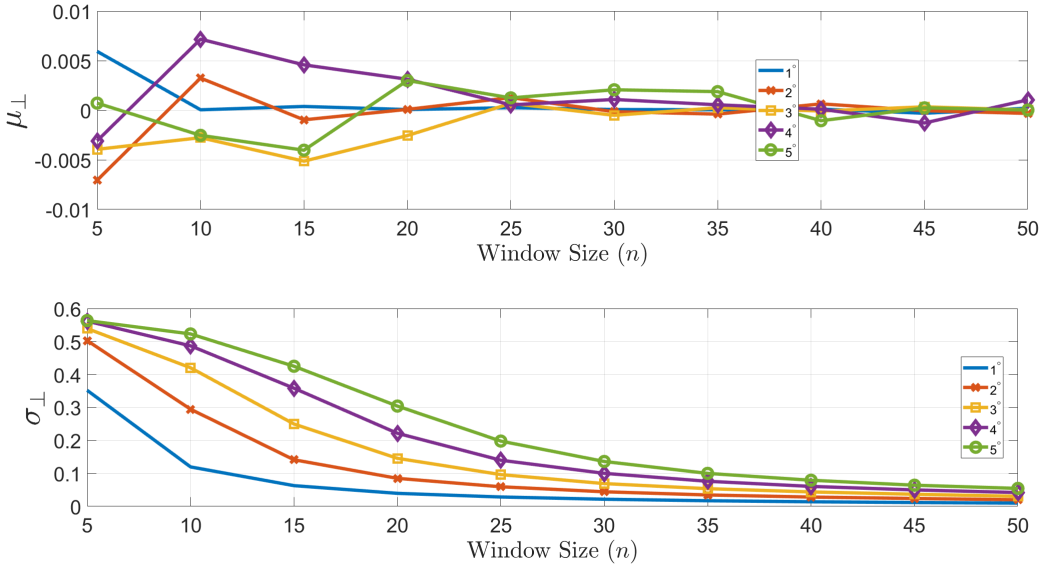

Fig. 1 Sample Mean and Standard Deviation of the projection of the estimated AOR along a direction perpendicular to the true AOR. Measurements taken at 10Hz, with an AVM of $\Omega = 0.1$rad/s. Results are shown as a function of the number of measurements (x axis) and the standard deviations σ_θ (different plots).

Section IV.C, denoted as $\hat{\boldsymbol{\sigma}}_\omega = \begin{bmatrix} \hat{\boldsymbol{\sigma}}_{\omega x} & \hat{\boldsymbol{\sigma}}_{\omega y} & \hat{\boldsymbol{\sigma}}_{\omega z} \end{bmatrix}^T$. We compare both in a Percent Deviation sense:

$$PD_{\sigma x} \triangleq 100 \cdot \frac{\hat{\boldsymbol{\sigma}}_{\omega x} - \boldsymbol{\sigma}_{\omega x}}{\boldsymbol{\sigma}_{\omega x}}, \qquad PD_{\sigma y} \triangleq 100 \cdot \frac{\hat{\boldsymbol{\sigma}}_{\omega y} - \boldsymbol{\sigma}_{\omega y}}{\boldsymbol{\sigma}_{\omega y}}, \tag{54}$$
$$PD_{\sigma z} \triangleq 100 \cdot \frac{\hat{\boldsymbol{\sigma}}_{\omega z} - \boldsymbol{\sigma}_{\omega z}}{\boldsymbol{\sigma}_{\omega z}}.$$

Figure 3 shows how the covariance estimates are biased for a small number of measurements, but the bias diminishes as the number of measurements increase.

QuateRA's performance is improved drastically (compared to the example from before) in a scenario for which measurements are taken at 1Hz, still with an AVM of $\Omega = 0.1$rad/s. We can see that both the bias and the standard deviations (Figures 4 and 5) are reduced compared with the previous scenario, and the estimated covariance is very close to the sample covariance (Figure 6). Our reasoning for improvement is based upon the fact that TLS can provide better planar estimates when the quaternion measurements are more sparsely distributed along the plane, whereas the previous scenario had many quaternions close to each other, making it harder to determine the plane of rotation from the given measurements.

In order to compare QuateRA with the MEKF, we will analyze varying values for sampling frequency δt and angular velocity magnitude. We compare both estimators by evaluating the Least Squares cost function of Eq. 15 $J_{LS} = n - \sum \hat{\boldsymbol{q}}_i^T \bar{\boldsymbol{q}}_i$. In order to obtain the MEKF quaternion estimates $\hat{\boldsymbol{q}}$, we first execute the MEKF algorithm

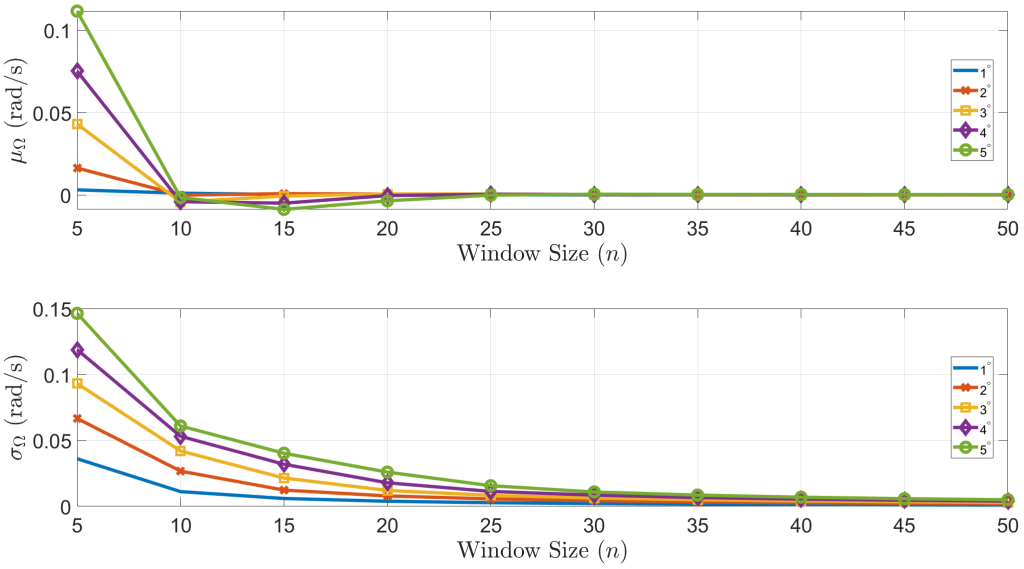

Fig. 2 Sample Mean and Standard Deviation of the estimated AVM error for measurements taken at 10Hz, and an AVM of $\Omega = 0.1$rad/s. Results are shown as a function of the number of measurements (x axis) and the standard deviations σ_θ (different plots).

- processing all orientation measurements - and then we use the estimated angular velocity to propagate the final orientation backwards in time to obtain previous orientations (smoothing procedure). We compare Quatera with MEKF as a percent deviation:

$$PD(\%) = 100 \cdot \frac{J_{LS}(MEKF) - J_{LS}(QuateRA)}{J_{LS}(MEKF)}, \tag{55}$$

where MEKF outperforms QuateRA when $PD(\%) < 0$ and QuateRA ourperforms MEKF otherwise.

Table 1 presents the average percent deviation for $dt = 0.1$s and $\Omega = 0.1$rad. We notice that MEKF outperforms QuateRA most of the time for this scenario. The performance between both is quite similar when the number of measurements is in the range $n \geq 25$. Table 2 presents the percent deviation for $dt = 1$s and $\Omega = 0.1$rad, and we notice that there is no clear winner when comparing both in this scenario. In contrast, QuateRA outperforms MEKF largely when $dt = 1$s and $\Omega = 1$rad, as shown in Table 3. We attribute the poor performance of MEKF in this last scenario due to the fact that MEKF is just a first order filter, and its performance degrades when nonlinearities become dominant when measurements are taken sparsely.

Finally, QuateRA is compared with a solution obtained from a batch nonlinear solver for the optimization problem of Eq. 15. In order to perform a fair comparison, the optimization problem is initialized in two ways:

- First, we provide an initial guess given by the exact true value of the angular velocity and initial orientation.
- Second, we provide an initialization assuming the approximations $\dot{\boldsymbol{q}} \approx \frac{\bar{\boldsymbol{q}}_2 - \bar{\boldsymbol{q}}_1}{t_2 - t_1}$ and $\dot{\boldsymbol{q}} \approx \frac{1}{2}\boldsymbol{\omega} \otimes \bar{\boldsymbol{q}}_1$. We then initialize

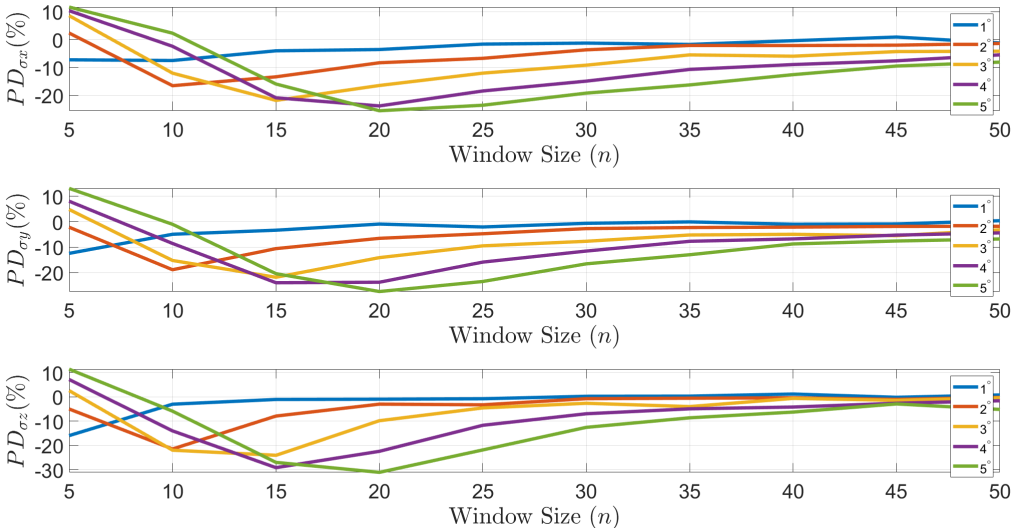

Fig. 3 Percentual deviation of the average estimated standard deviation for $\hat{\sigma}_\omega$ w.r.t. the sample standard deviation σ_ω, assuming sampling frequency of 10Hz, and an AVM of $\Omega = 0.1$rad/s. Results are shown as a function of the number of measurements (x axis) and the measurement noise standard deviations σ_θ (different plots).

the nonlinear solver with the estimates $\hat{\boldsymbol{\omega}} = 2\frac{\bar{\boldsymbol{q}}_2 - \bar{\boldsymbol{q}}_1}{t_2 - t_1} \otimes \bar{\boldsymbol{q}}_1^{-1}$ and $\hat{\boldsymbol{q}}_1 = \bar{\boldsymbol{q}}_1$.

We have used Matlab's [25] function *fmincon* [26] using the *interior-point* algorithm with constraint tolerance of 10^{-6}, maximum of 1000 iterations and optimality tolerance of 10^{-6}. Again, we perform 10000 Monte Carlo executions for the same scenarios as in the previous comparison: first with $dt = 0.1$s and $\Omega = 0.1$rad, then $dt = 1$s and $\Omega = 0.1$rad, and $dt = 1$s and $\Omega = 1$rad.

Again, we use a Percent Deviation as a metric of comparison between QuateRA and *Fmincon*. The following expression is used:

$$PD(\%) = 100 \cdot \frac{J_{LS}(Fmincon) - J_{LS}(QuateRA)}{J_{LS}(Fmincon)}, \tag{56}$$

where *Fmincon* outperforms QuateRA when $PD(\%) < 0$ and QuateRA ourperforms *Fmincon* otherwise. In principle, *Fmincon* should always outperform QuateRA, and this analysis helps us in understanding how far is QuateRA from the optimal solution. Note that the analysis hereinafter is done only for *Fmincon's* converged solutions, and the non-converged ones are discarded.

Tables 4-6 show the number of times that *Fmincon* converged for each scenario over 10000 executions with perfect initial guesses. We see that *Fmincon* always had trouble to converge when the number of input measurements were $n = 5$. When $n \neq 5$, we see some variability for the number of convergences depending on the scenario for the sampling frequency and angular velocity. This hints at the idea that a nonlinear optimizer won't always converge, and we see a particular case in which it converged only 67% of the time.

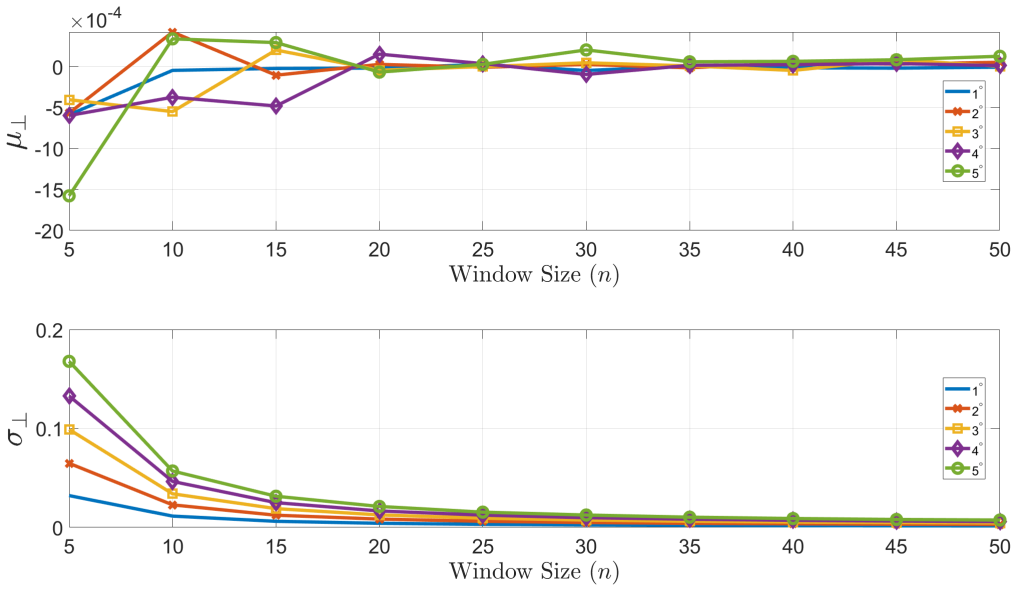

Fig. 4 Sample Mean and Standard Deviation of the projection of the estimated AOR along a direction perpendicular to the true AOR. Measurements taken at 1Hz, with an AVM of $\Omega = 0.1$rad/s. Results are shown as a function of the number of measurements (x axis) and the standard deviations σ_θ (different plots).

Tables 7-9 show the percent deviation between both methods. Just as in the comparison with the MEKF, QuateRA is slightly outperformed when n is low, $dt = 0.1$s and $\Omega = 0.1$rad. As for the remaining scenarios (Tables 14 and 15), QuateRA seems to perform just as good as *Fmincon*, indicating that QuateRA is close to the optimal solution for these explored cases.

Tables 10-12 show the number of times that *Fmincon* converged for each scenario when the optimizer is initialized with an initial guess that is not the true values for the angular velocity and initial orientation. Again, *Fmincon* always had trouble to converge when the number of input measurements were $n = 5$. When $n \neq 5$, we see that the optimizer converged more often with the given initial values, where the worst case converged only 83.8% of the time. Although the convergence rate for non-perfect initialization is higher than when perfectly initialized, it seems that many of the converged solutions actually converge to a local minimum that is not close to the global solution.

Tables 13-15 show the percent deviation between both methods. Just as in the previous comparisons, QuateRA is slightly outperformed when n is low, $dt = 0.1$s and $\Omega = 0.1$rad. On the other hand, the results in Tables 14 and 15 show that QuateRA largely outperforms *Fmincon* for large number of input measurements (and performs similarly for a low number of measurements). The reason for that is because *Fmincon* converges to a local minimum, showing how a nonlinear optimizer is susceptible to perform poorly if not well initialized for the problem at hand.

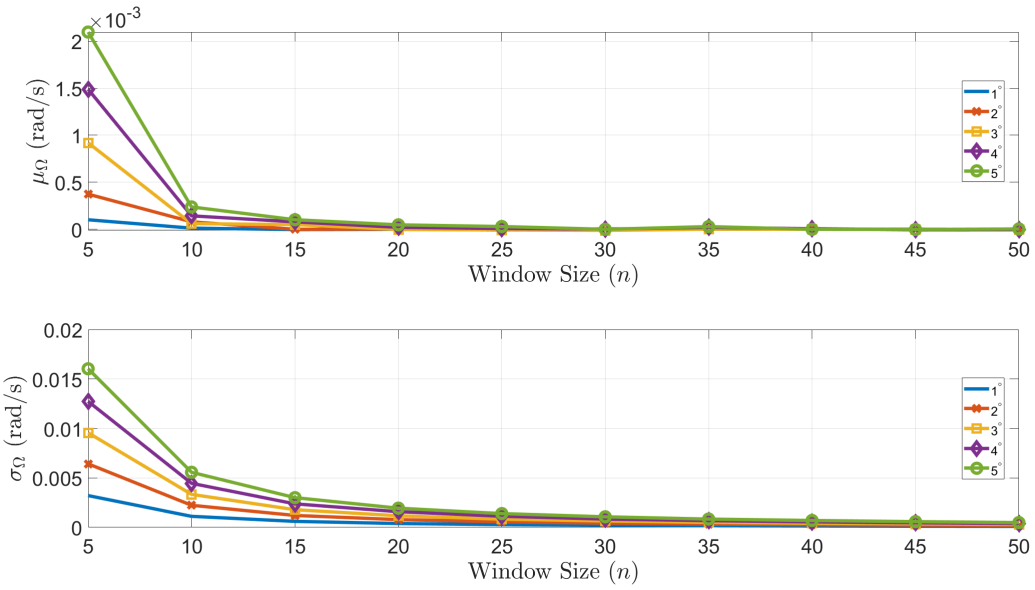

Fig. 5 Sample Mean and Standard Deviation of the estimated AVM error for measurements taken at 10Hz, and an AVM of $\Omega = 0.1$rad/s. Results are shown as a function of the number of measurements (x axis) and the standard deviations σ_θ (different plots).

VI. Conclusions

This work presented a batch estimation procedure for the determination of a constant angular velocity from quaternion measurements. In the constant angular velocity scenario, we show that the orientation quaternion evolves without departing from a fixed plane of rotation. With this insight, we are able to estimate the axis of rotation. Given the plane of rotation, the quaternions can be reprojected onto this plane, being parametrized as a single evolving angle on the plane. The angular velocity magnitude is then estimated from the evolution of the quaternion angle on the plane.

As we show in our Monte Carlo section, the performance of the Quaternion Regression Algorithm (QuateRA) is a function of n and the expected amplitude of the measurement noise. Our results indicate asymptotic unbiasedness of QuateRA, and we are able to accurately determine the standard deviation of the angular velocity estimation for sufficiently large sample sets. We show that QuateRA performs very close to a Multiplicative Extended Kalman Filter (MEKF), even outperforming the latter when nonlinearities are dominant (as is tipically the case with large angular rates), as MEKF is a first order estimator.

When compared with a nonlinear optimization solver, QuateRA performs very close to *Fmincon* when the latter is initialized with the exact truth values. When *Fmincon* is initialized with an educated guess (but not necessarily truth values), QuateRA outperforms *Fmincon* specially for large measurement sample sets.

Our earlier contributions have also demonstrated the application of preliminary versions of QuateRA for estimating a non-constant angular velocity. These works introduced tuning parameters for adapting the size of the sliding window and for tuning the Angular Velocity Magnityde (AVM) estimator. In contrast, the current work presents a method for

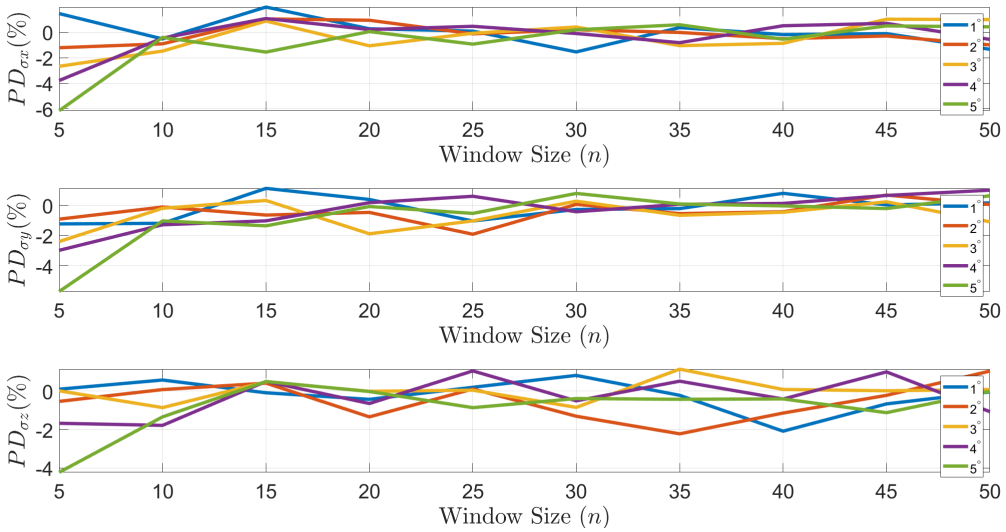

Fig. 6 Percentual deviation of the average estimated standard deviation for $\hat{\sigma}_\omega$ w.r.t. the sample standard deviation σ_ω, assuming sampling frequency of 10Hz, and an AVM of $\Omega = 0.1$rad/s. Results are shown as a function of the number of measurements (x axis) and the measurement noise standard deviations σ_θ (different plots).

estimating the AVM that is free of tuning parameters, and it does produce a covariance estimate for the estimate (provided a sufficiently large sample set). These contributions are relevant for the overall problem of estimating a time-varying Axis of Rotation (AOR) without the need for heuristic tuning. In the case of non-constant angular velocity with unknown torques and inertia matrix, filtering techniques as an MEKF are not appropriate solutions because the dynamics are not fully modeled. On the other hand, a self-tuning algorithm such as a QuateRA-based sliding window with statistically adaptive window size can figure out how many measurements can be taken without breaking the assumption that the angular velocity is approximately constant. Hence, QuateRA is applicable not only for constant angular velocity (or even the pure-spin case), but also in the presence of unmodeled attitude dynamics (large uncertainties in the inertial properties and possible presence of unknown external disturbance torques).

An interesting path of future work would be to determine a covariance estimate associated with the estimated AOR. Classically, it is possible to estimate asymptotic covariances for TLS solutions provided that the solution is unique. As shown in Section IV, the TLS solution for this problem is not unique (still, the estimate of the angular velocity direction is unique) and we cannot determine the covariance of $\hat{\boldsymbol{u}}_1$ and $\hat{\boldsymbol{u}}_2$ using classical methods in TLS. Since the AOR estimate is determined from $\hat{\boldsymbol{u}}_1$ and $\hat{\boldsymbol{u}}_2$, computing the covariance of the estimated AOR is not trivial. Therefore, establishing the AOR covariance would be a meaningful contribution for future work.

Another interesting path for future research would be to expand QuateRA for non-constant measurement covariance over multiple measurements. Additionally, we have assumed that the axis of the noise quaternion is distributed in a uniform spherical distribution, whereas this is not always true in practice. For instance, star trackers typically have different covariances associated with the roll, pitch and yaw directions. Hence, it would also be meaningful to adapt

QuateRA to accommodate for a more accurate measurement model.

VII. Appendices

A. Statistics of the Spherical Uniform Distribution

In this section we prove that if $\boldsymbol{e} \in \mathbb{S}^2$ is a unit vector uniformly distributed in the 3-D unit sphere, then: $\mathbb{E}[\boldsymbol{e}] = \mathbf{0}$ and $\mathbb{E}[\boldsymbol{e}\boldsymbol{e}^T] = \frac{1}{3}\boldsymbol{I}$.

Assume a unit radius sphere and a cylinder of radius $r = 1$ and height $h = 2$. According with Archimedes' Hat-Box Theorem [27], if we slice both the cylinder and the sphere at the same height as shown on Fig. 7, then the lateral surface area of the spherical segment (S_1) is equal to the lateral surface area of the cylindrical segment (S_2).

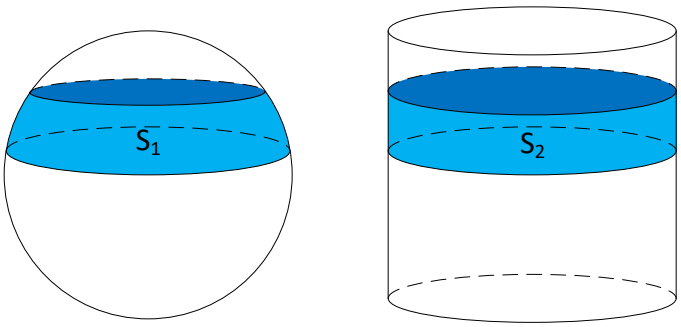

Fig. 7 Illustration of Archimedes' Hat-Box Theorem.

More specifically, the surface area S of the cylinder parametrized with radius $r = 1$ and height $h = 2$ is the same as the unit-radius sphere, i.e, $S = 4\pi$. A commonly used method [28] to generate uniformly distributed samples on a sphere $\boldsymbol{e} \in \mathbb{S}^2$ is to uniformly sample a point in the cylinder through a height value $z \sim \mathcal{U}[-1,\ 1]$, and an angle value $\phi \sim \mathcal{U}[-\pi,\ \pi]$, and then map it to the sphere through the transformation:

$$\boldsymbol{e} = \begin{bmatrix} \sqrt{1 - z^2}\cos(\phi) \\ \sqrt{1 - z^2}\sin(\phi) \\ z \end{bmatrix}. \tag{57}$$

The transformation of Eq. 57 guarantees that areas in the cylinder are preserved in the sphere after the projection. Therefore, if a random variable is uniformly distributed in the prior space (cylindrical space), then it should still be uniformly distributed in the posterior space (spherical space).

Denoting $P_z(x)$ and $P_\phi(x)$ as the probability distributions of the scalar variables z and ϕ respectively, then:

$$\begin{aligned}
\mathbb{E}[z] &= \int_{-1}^{1} xP_z(x)\,dx = \frac{1}{2}\int_{-1}^{1} x\,dx = \frac{1}{4}x^2\Big|_{-1}^{1} = 0 \\
\mathbb{E}[z^2] &= \int_{-1}^{1} x^2P_z(x)\,dx = \frac{1}{2}\int_{-1}^{1} x^2\,dx = \frac{1}{6}x^3\Big|_{-1}^{1} = \frac{1}{3} \\
\mathbb{E}[1-z^2] &= 1 - \frac{1}{3} = \frac{2}{3} \\
\mathbb{E}[\cos\phi] &= \int_{-\pi}^{\pi} \cos x P_\phi(x)\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos x\,dx = \frac{1}{2\pi}\sin x\Big|_{-\pi}^{\pi} = 0 \\
\mathbb{E}[\sin\phi] &= \int_{-\pi}^{\pi} \sin x P_\phi(x)\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} \sin x\,dx = -\frac{1}{2\pi}\cos x\Big|_{-\pi}^{\pi} = 0 \\
\mathbb{E}[\cos\phi\sin\phi] &= \int_{-\pi}^{\pi} \cos x \sin x P_\phi(x)\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos x\sin x\,dx = -\frac{1}{4\pi}cos^2x\Big|_{-\pi}^{\pi} = 0 \\
\mathbb{E}[\cos^2\phi] &= \int_{-\pi}^{\pi} \cos^2 x P_\phi(x)\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos^2 x\,dx = \frac{1}{8\pi}(2x+\sin 2x)\Big|_{-\pi}^{\pi} = \frac{1}{2} \\
\mathbb{E}[\sin^2\phi] &= \int_{-\pi}^{\pi} \sin^2 x P_\phi(x)\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} \sin^2 x\,dx = \frac{1}{8\pi}(2x-\sin 2x)\Big|_{-\pi}^{\pi} = \frac{1}{2}
\end{aligned}$$

Therefore, given that z and ϕ are independently distributed, we have that:

$$\mathbb{E}[\boldsymbol{e}] = \begin{bmatrix} \mathbb{E}[\sqrt{1-z^2}\cos(\phi)] \\ \mathbb{E}[\sqrt{1-z^2}\sin(\phi)] \\ \mathbb{E}[z] \end{bmatrix} = \begin{bmatrix} \mathbb{E}[\sqrt{1-z^2}]\mathbb{E}[\cos(\phi)] \\ \mathbb{E}[\sqrt{1-z^2}]\mathbb{E}[\sin(\phi)] \\ \mathbb{E}[z] \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{58}$$

Also, we have that:

$$\mathbb{E}[\boldsymbol{e}\boldsymbol{e}^T] = \mathbb{E}\left[\begin{bmatrix} (1-z^2)\cos^2\phi & (1-z^2)\cos\phi\sin\phi & (1-z^2)z\cos\phi \\ (1-z^2)\cos\phi\sin\phi & (1-z^2)\sin^2\phi & (1-z^2)z\sin\phi \\ (1-z^2)z\cos\phi & (1-z^2)z\sin\phi & z^2 \end{bmatrix}\right] = \frac{1}{3}\boldsymbol{I}. \tag{59}$$

B. Multiplicative Extended Kalman Filter Formulation

We present the Multiplicative Extended Kalman Filter (MEKF) formulation for the problem in hand. The filter herein presented is based on the formulations in [11] and [13]. However, our equations differ from the works cited since we do not have gyroscope measurements, and we do not estimate the gyroscope measurement bias. In addition, the assumption that the angular velocity is constant implies no process noise in the dynamics propagation.

We define the reference trajectory kinematics:

$$\dot{\boldsymbol{q}}_R = \frac{1}{2}\boldsymbol{\omega}_R \otimes \boldsymbol{q}_R, \tag{60}$$

where $\boldsymbol{q}_R \triangleq \begin{bmatrix} q_{Rs} & \boldsymbol{q}_{Rv}^T \end{bmatrix}^T$ is the reference quaternion and $\boldsymbol{\omega}_R$ is the reference angular velocity of the reference attitude. The true attitude $\boldsymbol{q}$ can be represented as:

$$\boldsymbol{q} = \delta\boldsymbol{q} \otimes \boldsymbol{q}_R, \tag{61}$$

where $\delta\boldsymbol{q} \triangleq \begin{bmatrix} \delta q_s & \delta\boldsymbol{q}_v^T \end{bmatrix}^T$ represents the rotation from $\boldsymbol{q}_R$ to the true rotation.

Differentiating Eq. 61, we get:

$$\dot{\boldsymbol{q}} = \delta\dot{\boldsymbol{q}} \otimes \boldsymbol{q}_R + \delta\boldsymbol{q} \otimes \dot{\boldsymbol{q}}_R \quad \Longrightarrow \quad \frac{1}{2}\boldsymbol{\omega} \otimes \boldsymbol{q} = \delta\dot{\boldsymbol{q}} \otimes \boldsymbol{q}_R + \frac{1}{2}\delta\boldsymbol{q} \otimes \boldsymbol{\omega}_R \otimes \boldsymbol{q}_R. \tag{62}$$

Post-multiplying Eq. 62 by $\boldsymbol{q}_R^{-1}$ and isolating $\delta\dot{\boldsymbol{q}}$, we get:

$$\begin{aligned} \delta\dot{\boldsymbol{q}} &= \frac{1}{2}\left(\boldsymbol{\omega} \otimes \boldsymbol{q} \otimes \boldsymbol{q}_R^{-1} - \delta\boldsymbol{q} \otimes \boldsymbol{\omega}_R\right) = \frac{1}{2}\left(\boldsymbol{\omega} \otimes \delta\boldsymbol{q} - \delta\boldsymbol{q} \otimes \boldsymbol{\omega}_R\right) \\ &= \frac{1}{2}\left(\begin{bmatrix} 0 & -\boldsymbol{\omega} \\ \boldsymbol{\omega} & -[\boldsymbol{\omega}_\times] \end{bmatrix}\begin{bmatrix} \delta q_s \\ \delta\boldsymbol{q}_v \end{bmatrix} - \begin{bmatrix} \delta q_s & -\delta\boldsymbol{q}_v^T \\ \delta\boldsymbol{q}_v & \delta q_s\boldsymbol{I} - [\delta\boldsymbol{q}_{v\times}] \end{bmatrix}\begin{bmatrix} 0 \\ \boldsymbol{\omega}_R \end{bmatrix}\right) \end{aligned} \tag{63}$$

After some algebraic manipulations, we get that:

$$\delta\dot{\boldsymbol{q}}_s = \left(\boldsymbol{\omega}_R - \boldsymbol{\omega}\right)^T \delta\boldsymbol{q}_v, \qquad \delta\dot{\boldsymbol{q}}_v = \left(\boldsymbol{\omega} - \boldsymbol{\omega}_R\right)\delta\boldsymbol{q}_s - \left(\boldsymbol{\omega} + \boldsymbol{\omega}_R\right) \times \delta\boldsymbol{q}_v.$$

We define the scaled attitude error Gibbs vector:

$$\delta\boldsymbol{g} \triangleq 2\frac{\delta\boldsymbol{q}_v}{\delta\boldsymbol{q}_s}, \tag{64}$$

The Gibbs vector associated with the noise quaternion of Eq. 7 is given by:

$$\boldsymbol{g}_{Nk} = 2\boldsymbol{e}_{Nk}\tan\frac{\theta_k}{2}. \tag{65}$$

The transformation from *Gibbs vector* to quaternion is done as follows:

$$\delta q_s = \sqrt{\frac{2}{2 + ||\delta \boldsymbol{g}||^2}}, \qquad \delta \boldsymbol{q}_v = \frac{1}{2}\delta q_s \delta \boldsymbol{g}. \tag{66}$$

One should notice that $\mathbb{E}[\boldsymbol{g}_N] = 0$. In addition, assuming a small angle approximation, then $\tan^2 \frac{\theta_k}{2} \approx \frac{\theta_k^2}{4}$, leading to $\mathbb{E}[\boldsymbol{g}_N \boldsymbol{g}_N^T] = \frac{1}{3}\sigma_\theta^2 \boldsymbol{I}$.

The Gibbs error kinematics is described as:

$$\delta \dot{\boldsymbol{g}} = 2\frac{\delta \dot{\boldsymbol{q}}_v}{\delta \boldsymbol{q}_s} - 2\frac{\delta \boldsymbol{q}_v}{\delta \boldsymbol{q}_s}\frac{\delta \dot{\boldsymbol{q}}_s}{\delta \boldsymbol{q}_s} = \left[\boldsymbol{I} + \tfrac{1}{4}\delta \boldsymbol{g} \delta \boldsymbol{g}^T\right]\left(\boldsymbol{\omega} - \boldsymbol{\omega}_R\right) - \frac{1}{2}\left(\boldsymbol{\omega} + \boldsymbol{\omega}_R\right) \times \delta \boldsymbol{g}.$$

Assuming the first order approximations $\delta \boldsymbol{g} \delta \boldsymbol{g}^T \approx 0$, and $\delta \boldsymbol{\omega} \times \delta \boldsymbol{g} \approx 0$, we get to:

$$\delta \dot{\boldsymbol{g}} \approx \delta \boldsymbol{\omega} - \boldsymbol{\omega}_R \times \delta \boldsymbol{g}. \tag{67}$$

Defining the state vector $\boldsymbol{X} \triangleq \begin{bmatrix} \delta \boldsymbol{g}^T & \boldsymbol{\omega}^T \end{bmatrix}^T$ and the dynamics of Eq. 67, then we have the linearized state dynamics:

$$\dot{\boldsymbol{X}} = \underbrace{\begin{bmatrix} -[\boldsymbol{\omega}_\times] & \boldsymbol{I}_3 \\ \boldsymbol{0}_3 & \boldsymbol{0}_3 \end{bmatrix}}_{\triangleq \boldsymbol{A}} \boldsymbol{X}. \tag{68}$$

We define the state transition matrix:

$$\boldsymbol{A}_d[k] \triangleq e^{A\delta_k}, \quad \delta_k \triangleq t_{k+1} - t_k. \tag{69}$$

In the propagation step, the following equations are used:

$$\begin{aligned} \boldsymbol{q}_{k+1|k} &= \boldsymbol{F}(\boldsymbol{\omega}_{k|k})\boldsymbol{q}_{k|k}, \\ \boldsymbol{\omega}_{k+1|k} &= \boldsymbol{\omega}_{k|k}, \\ \boldsymbol{P}_{k+1|k} &= \boldsymbol{A}_d[k]\boldsymbol{P}_{k|k}\boldsymbol{A}_d^T[k], \end{aligned}$$

where $\boldsymbol{P}_{k|k} \triangleq \mathbb{E}[\boldsymbol{X}_{k|k}\boldsymbol{X}_{k|k}^T]$, and $\boldsymbol{P}_{k+1|k} \triangleq \mathbb{E}[\boldsymbol{X}_{k+1|k}\boldsymbol{X}_{k+1|k}^T]$, and $\boldsymbol{F}(\boldsymbol{\omega}_{k|k})$ is defined in Eq. 5.

As for the measurement model, only quaternion measurements are available. The innovation term is given by:

$$\boldsymbol{\nu}_k = 2\frac{\tilde{\boldsymbol{q}}_v[k]}{\tilde{q}_s[k]}, \tag{70}$$

where $\tilde{q}_s[k]$ and $\tilde{\boldsymbol{q}}_v[k]$ are, respectively, the scalar and vector parts of $\tilde{\boldsymbol{q}}_k$, defined as:

$$\tilde{\boldsymbol{q}}_k \triangleq \begin{bmatrix} \tilde{q}_s[k] \\ \tilde{\boldsymbol{q}}_v[k] \end{bmatrix} \triangleq \hat{\boldsymbol{q}}_k \otimes \boldsymbol{q}_{k|k-1}^{-1}.$$

Assuming the measurement noise defined in Eq. 65, the measurement covariance is given by $\boldsymbol{R}_k \triangleq \mathbb{E}[\boldsymbol{g}_N \boldsymbol{g}_N^T] = \frac{1}{3}\sigma_\theta^2 \boldsymbol{I}$.

The measurement update step uses the following expressions:

$$\begin{aligned}
\boldsymbol{H}_k &= \begin{bmatrix} \boldsymbol{I}_3 & \boldsymbol{0}_3 \end{bmatrix}, \\
\boldsymbol{S}_k &= \boldsymbol{H}_k \boldsymbol{P}_{k|k-1} \boldsymbol{H}_k^T + \boldsymbol{R}_k, \\
\boldsymbol{K}_k &= \boldsymbol{P}_{k|k-1} \boldsymbol{H}_k^T \boldsymbol{S}_k^{-1}, \\
\Delta\boldsymbol{x}_{k|k} &= \boldsymbol{K}_k \boldsymbol{\nu}_k, \\
\boldsymbol{P}_{k|k} &= (\boldsymbol{I} - \boldsymbol{K}_k \boldsymbol{H}_k)\boldsymbol{P}_{k|k-1}(\boldsymbol{I} - \boldsymbol{K}_k \boldsymbol{H}_k)^T + \boldsymbol{K}_k \boldsymbol{R}_k \boldsymbol{K}_k^T,
\end{aligned}$$

where $\Delta\boldsymbol{x}_{k|k} \triangleq \begin{bmatrix} \Delta x_1 & \Delta x_2 & \Delta x_3 & \Delta x_4 & \Delta x_5 & \Delta x_6 \end{bmatrix}^T$ is the incremental state update typical for standard EKF formulations.

The updated state $\delta\boldsymbol{g}_{k|k}$ can be obtained from $\Delta\boldsymbol{x}_{k|k}$ as $\delta\boldsymbol{g}_{k|k} = \begin{bmatrix} \Delta x_1 & \Delta x_2 & \Delta x_3 \end{bmatrix}^T$. Bearing in mind that $\delta\boldsymbol{g}_{k|k}$ represents the attitude error respective to $\delta\boldsymbol{q}_k$ (see Eqs. 61 and 64), then $\delta\boldsymbol{q}_{k|k}$ can be obtained from $\delta\boldsymbol{g}_{k|k}$ using the transformation in Eq. 66. Defining $\Delta\boldsymbol{\omega} = \begin{bmatrix} \Delta x_4 & \Delta x_5 & \Delta x_6 \end{bmatrix}^T$, the updated states are given by:

$$\begin{aligned}
\boldsymbol{q}_{k|k} &= \delta\boldsymbol{q}_{k|k} \otimes \boldsymbol{q}_{k|k-1}, \\
\boldsymbol{\omega}_{k|k} &= \boldsymbol{\omega}_{k|k-1} + \Delta\boldsymbol{\omega}.
\end{aligned}$$

References

[1] Bar-Shalom, Y., Li, X. R., and Kirubarajan, T., *Estimation with applications to tracking and navigation: theory algorithms and software*, John Wiley & Sons, 2004. doi:10.1002/0471221279.

[2] Salcudean, S., "A globally convergent angular velocity observer for rigid body motion," *IEEE transactions on Automatic*

Control, Vol. 36, No. 12, 1991, pp. 1493–1497. doi:10.1109/9.106169.

[3] Oshman, Y., Franc-para, and Dellus, o., "Spacecraft angular velocity estimation using sequential observations of a single directional vector," *Journal of Spacecraft and Rockets*, Vol. 40, No. 2, 2003, pp. 237–247. doi:10.2514/2.3938.

[4] Bar-Itzhack, I. Y., "Classification of algorithms for angular velocity estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 24, No. 2, 2001, pp. 214–218. doi:10.2514/2.4731.

[5] Bar-Itzhack, I. Y., Harman, R. R., and Thienel, J. K., "Rigid body rate inference from attitude variation," *Journal of guidance, control, and dynamics*, Vol. 30, No. 1, 2007, pp. 275–281. doi:10.2514/1.23955.

[6] Psiaki, M. L., "Generalized Wahba problems for spinning spacecraft attitude and rate determination," *The Journal of the Astronautical Sciences*, Vol. 57, No. 1-2, 2009, pp. 73–92. doi:10.1007/BF03321495.

[7] Markley, F. L., and Mortari, D., "Quaternion attitude estimation using vector observations." *Journal of the Astronautical Sciences*, Vol. 48, No. 2, 2000, pp. 359–380.

[8] Saunderson, J., Parrilo, P. A., and Willsky, A. S., "Convex Solution to a Joint Attitude and Spin-Rate Estimation Problem," *Journal of Guidance, Control, and Dynamics*, Vol. 39, No. 1, 2015, pp. 118–127. doi:10.2514/1.G001107.

[9] Hinks, J. C., and Psiaki, M. L., "Solution Strategies for an Extension of Wahba's Problem to a Spinning Spacecraft," *Journal of Guidance, Control, and Dynamics*, Vol. 34, No. 6, 2011, pp. 1734–1745. doi:10.2514/1.53530.

[10] Psiaki, M. L., and Hinks, J. C., "Numerical solution of a generalized Wahba problem for a spinning spacecraft," *Journal of Guidance, Control, and Dynamics*, Vol. 35, No. 3, 2012, pp. 764–773. doi:10.2514/1.56151.

[11] Lefferts, E. J., Markley, F. L., and Shuster, M. D., "Kalman filtering for spacecraft attitude estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 5, No. 5, 1982, pp. 417–429. doi:10.2514/3.56190.

[12] Gai, E., Daly, K., Harrison, J., and Lemos, L., "Star-sensor-based satellite attitude/attitude rate estimator," *Journal of Guidance, Control, and Dynamics*, Vol. 8, No. 5, 1985, pp. 560–565. doi:10.2514/3.56393.

[13] Markley, F. L., "Attitude error representations for Kalman filtering," *Journal of guidance, control, and dynamics*, Vol. 26, No. 2, 2003, pp. 311–317. doi:10.2514/2.5048.

[14] Psiaki, M. L., "Backward-smoothing extended Kalman filter," *Journal of guidance, control, and dynamics*, Vol. 28, No. 5, 2005, pp. 885–894. doi:10.2514/1.12108.

[15] Mortari, D., and Akella, M., "Discrete and Continuous Time Adaptive Angular Velocity Estimators," *Proceedings of the AAS/AIAA Space Flight Mechanics Conference, Williamsburg, VA*, 2015.

[16] Lerner, G. M., "Three-axis attitude determination," *Spacecraft Attitude Determination and Control*, Vol. 73, 1978, pp. 420–428.

[17] Markley, F. L., Cheng, Y., Crassidis, J. L., and Oshman, Y., "Averaging quaternions," *Journal of Guidance, Control, and Dynamics*, Vol. 30, No. 4, 2007, pp. 1193–1197. doi:10.2514/1.28949.

[18] Almeida, M., and Akella, M., "Discrete Adaptive Angular Velocity Estimation - An Experimental Analysis," *2016 Space Flight Mechanics Meeting*, 2016.

[19] Almeida, M., Zanetti, R., Mortari, D., and Akella, M., "Real-time Angular Velocity Estimation of Non-Cooperative Space Objects Using Camera Measurements," *2018 AAS/AIAA Astrodynamics Specialist Conference*, 2018.

[20] Mortari, D., and Majji, M., "Multiplicative measurement model," *The Journal of the Astronautical Sciences*, Vol. 57, No. 1-2, 2009, pp. 47–60. doi:10.1007/BF03321493.

[21] Crassidis, J. L., and Cheng, Y., "Error-covariance analysis of the total least-squares problem," *Journal of Guidance, Control, and Dynamics*, Vol. 37, No. 4, 2014, pp. 1053–1063. doi:10.2514/1.62959.

[22] Van Huffel, S., and Vandewalle, J., *The total least squares problem: computational aspects and analysis*, Vol. 9, Siam, 1991. doi:10.2307/2670017.

[23] Späth, H., "Orthogonal least squares fitting with linear manifolds," *Numerische Mathematik*, Vol. 48, No. 4, 1986, pp. 441–445. doi:10.1007/BF01389650.

[24] Huffel, S. V., and Vandewalle, J., "Analysis and properties of the generalized total least squares problem AX=B when some or all columns in A are subject to error," *SIAM Journal on Matrix Analysis and Applications*, Vol. 10, No. 3, 1989, pp. 294–315. doi:10.1137/0610023.

[25] MATLAB, *version 2019a*, The MathWorks Inc., Natick, Massachusetts, 2019.

[26] MATLAB, "fmincon Find minimum of constrained nonlinear multivariable function," `https://www.mathworks.com/help/optim/ug/fmincon.html`, 2019. Accessed: 2019-12-04.

[27] Cundy, H., and Rollett, A., "Sphere and cylinder—Archimedes' theorem," *Mathematical Models*, 1989, pp. 172–173.

[28] Weisstein, E. W., "Sphere Point Picking. From MathWorld–A Wolfram Web Resource." `http://mathworld.wolfram.com/SpherePointPicking.html`, 2018. Accessed: 2018-08-08.

Table 1 Percent deviation between MEKF and QuateRA for $dt = 0.1$s and $\Omega = 0.1$rad.

	$n = 5$	$n = 10$	$n = 15$	$n = 20$	$n = 25$	$n = 30$	$n = 35$	$n = 40$	$n = 45$	$n = 50$
$\sigma = 1°$	-9.91	-1.18	-0.41	-0.15	-0.06	0.04	0.07	-0.35	0.11	-0.31
$\sigma = 2°$	-9.56	-6.65	-1.71	-0.35	-0.80	-0.85	-0.08	0.08	-0.31	0.14
$\sigma = 3°$	-10.50	-8.27	-4.73	-2.83	-1.49	-0.95	-0.26	0.08	-0.40	-0.12
$\sigma = 4°$	-9.62	-9.04	-7.19	-3.53	-1.35	-1.04	-0.79	-0.16	-0.19	-0.36
$\sigma = 5°$	-9.04	-7.49	-7.08	-5.32	-2.93	-1.67	-0.86	-0.29	-0.20	-0.52

Table 2 Percent deviation between MEKF and QuateRA for $dt = 1$s and $\Omega = 0.1$rad.

	$n = 5$	$n = 10$	$n = 15$	$n = 20$	$n = 25$	$n = 30$	$n = 35$	$n = 40$	$n = 45$	$n = 50$
$\sigma = 1°$	1.18	-0.65	0.10	-0.59	0.09	-0.67	0.22	-0.80	-0.07	-0.22
$\sigma = 2°$	-0.61	-0.66	0.10	0.25	0.22	0.47	-0.04	0.01	0.05	-0.39
$\sigma = 3°$	0.81	-0.10	0.96	-0.35	-0.66	-0.33	-0.14	-0.58	0.13	-0.31
$\sigma = 4°$	-0.02	-0.32	0.35	0.55	-0.09	0.13	-0.39	-0.28	0.21	0.08
$\sigma = 5°$	-2.09	-0.66	-0.80	-0.15	0.03	0.56	0.38	-0.12	0.15	-0.25

Table 3 Percent deviation between MEKF and QuateRA for $dt = 1$s and $\Omega = 1$rad.

	$n = 5$	$n = 10$	$n = 15$	$n = 20$	$n = 25$	$n = 30$	$n = 35$	$n = 40$	$n = 45$	$n = 50$
$\sigma = 1°$	92.60	77.54	61.35	47.93	37.35	29.31	23.73	19.12	16.20	13.15
$\sigma = 2°$	76.43	47.19	29.11	19.51	13.79	10.08	7.66	5.74	5.26	4.27
$\sigma = 3°$	60.81	29.69	16.15	10.18	7.26	5.23	4.43	3.11	2.41	2.11
$\sigma = 4°$	47.12	20.07	10.47	6.75	4.95	3.35	2.47	1.93	1.67	1.45
$\sigma = 5°$	37.31	14.65	7.94	5.20	4.09	1.62	1.79	1.22	1.27	0.79

Table 4 Number of times that *Fmincon* converged for $dt = 0.1$s and $\Omega = 0.1$rad.

	$n = 5$	$n = 10$	$n = 15$	$n = 20$	$n = 25$	$n = 30$	$n = 35$	$n = 40$	$n = 45$	$n = 50$
$\sigma = 1°$	0	10000	10000	10000	10000	10000	10000	10000	10000	10000
$\sigma = 2°$	0	10000	10000	10000	10000	10000	10000	10000	10000	10000
$\sigma = 3°$	0	10000	10000	10000	10000	10000	10000	10000	10000	10000
$\sigma = 4°$	0	10000	10000	10000	10000	10000	10000	10000	10000	10000
$\sigma = 5°$	0	10000	10000	10000	10000	10000	10000	10000	10000	10000

Table 5 Number of times that *Fmincon* converged for $dt = 1$s and $\Omega = 0.1$rad.

	$n = 5$	$n = 10$	$n = 15$	$n = 20$	$n = 25$	$n = 30$	$n = 35$	$n = 40$	$n = 45$	$n = 50$
$\sigma = 1°$	0	10000	9994	9113	7116	6701	7472	8100	7204	6703
$\sigma = 2°$	0	10000	9995	9038	7642	7499	7548	7838	7550	7163
$\sigma = 3°$	0	10000	9988	9073	7928	7816	7665	7926	7842	7778
$\sigma = 4°$	0	10000	9989	9136	8215	7990	7841	8047	8144	8148
$\sigma = 5°$	0	10000	9986	9291	8541	8125	8071	8211	8338	8394

Table 6 Number of times that *Fmincon* converged for $dt = 1$s and $\Omega = 1$rad.

	$n = 5$	$n = 10$	$n = 15$	$n = 20$	$n = 25$	$n = 30$	$n = 35$	$n = 40$	$n = 45$	$n = 50$
$\sigma = 1°$	0	10000	10000	10000	9999	10000	9989	9970	9907	9831
$\sigma = 2°$	0	10000	10000	10000	9999	10000	9998	9988	9935	9824
$\sigma = 3°$	0	10000	10000	9999	10000	10000	9997	9974	9914	9789
$\sigma = 4°$	0	10000	10000	10000	9999	10000	9999	9969	9881	9753
$\sigma = 5°$	0	10000	10000	10000	10000	10000	9996	9970	9885	9743

Table 7 Percent deviation between *Fmincon* and QuateRA for $dt = 0.1$s and $\Omega = 0.1$rad (there is no data for $n = 5$).

	$n = 5$	$n = 10$	$n = 15$	$n = 20$	$n = 25$	$n = 30$	$n = 35$	$n = 40$	$n = 45$	$n = 50$
$\sigma = 1°$	x	-2.24	-0.70	-0.86	0.23	-0.15	-0.47	-0.09	0.24	0.06
$\sigma = 2°$	x	-6.49	-1.22	-1.55	-0.35	0.15	-0.34	-0.35	-0.07	-0.13
$\sigma = 3°$	x	-8.15	-5.81	-2.76	-0.68	-0.40	-0.72	0.18	-0.21	0.22
$\sigma = 4°$	x	-8.01	-6.88	-4.35	-2.00	-1.07	-0.32	-0.45	-0.15	0.14
$\sigma = 5°$	x	-6.62	-6.77	-5.03	-2.42	-0.96	-1.06	-0.66	-0.42	-0.61

Table 8 Percent deviation between *Fmincon* and QuateRA for $dt = 1$s and $\Omega = 0.1$rad (there is no data for $n = 5$).

	$n = 5$	$n = 10$	$n = 15$	$n = 20$	$n = 25$	$n = 30$	$n = 35$	$n = 40$	$n = 45$	$n = 50$
$\sigma = 1°$	x	-0.45	-0.50	0.28	0.55	0.32	0.42	0.07	0.04	-0.57
$\sigma = 2°$	x	-1.45	0.40	-0.20	-0.23	0.63	0.06	0.51	0.71	0.17
$\sigma = 3°$	x	0.45	0.00	0.21	0.12	0.47	0.20	-0.71	0.32	0.52
$\sigma = 4°$	x	-0.76	0.63	0.08	-0.01	0.05	0.46	-0.42	0.42	0.16
$\sigma = 5°$	x	-0.19	0.75	-0.12	-0.24	-0.15	0.93	-0.23	-0.23	0.15

Table 9 Percent deviation between *Fmincon* and QuateRA for $dt = 1$s and $\Omega = 1$rad (there is no data for $n = 5$).

	$n = 5$	$n = 10$	$n = 15$	$n = 20$	$n = 25$	$n = 30$	$n = 35$	$n = 40$	$n = 45$	$n = 50$
$\sigma = 1°$	x	1.09	-0.12	0.53	-0.18	0.09	-0.05	0.36	0.26	-0.41
$\sigma = 2°$	x	0.35	-0.13	0.23	0.34	0.32	27.99	0.38	0.28	0.09
$\sigma = 3°$	x	-0.56	-0.64	-0.14	-0.19	-0.06	0.21	-0.08	-0.33	-0.23
$\sigma = 4°$	x	-0.21	0.21	-0.28	-0.16	-0.02	-0.09	0.28	-0.12	-0.16
$\sigma = 5°$	x	-0.38	-0.21	-0.26	0.96	-0.12	-0.61	-0.22	-0.04	-0.40

Table 10 Number of times that *Fmincon* converged for $dt = 0.1$s and $\Omega = 0.1$rad.

	$n = 5$	$n = 10$	$n = 15$	$n = 20$	$n = 25$	$n = 30$	$n = 35$	$n = 40$	$n = 45$	$n = 50$
$\sigma = 1°$	0	10000	10000	10000	10000	10000	10000	10000	10000	10000
$\sigma = 2°$	0	10000	10000	10000	10000	10000	10000	10000	10000	10000
$\sigma = 3°$	0	10000	10000	10000	10000	10000	10000	10000	10000	10000
$\sigma = 4°$	0	10000	10000	10000	10000	10000	10000	10000	10000	10000
$\sigma = 5°$	0	10000	10000	10000	10000	10000	10000	10000	10000	10000

Table 11 Number of times that *Fmincon* converged for $dt = 1$s and $\Omega = 0.1$rad.

	$n = 5$	$n = 10$	$n = 15$	$n = 20$	$n = 25$	$n = 30$	$n = 35$	$n = 40$	$n = 45$	$n = 50$
$\sigma = 1°$	0	10000	9984	9200	8544	8545	8837	8664	8455	8384
$\sigma = 2°$	0	10000	9982	9216	8619	8663	8795	8827	8946	9005
$\sigma = 3°$	0	10000	9964	9327	8847	8835	8973	9088	9231	9326
$\sigma = 4°$	0	10000	9974	9448	8961	8984	9057	9198	9354	9518
$\sigma = 5°$	0	10000	9972	9519	9138	9116	9185	9296	9454	9568

Table 12 Number of times that *Fmincon* converged for $dt = 1$s and $\Omega = 1$rad.

	$n = 5$	$n = 10$	$n = 15$	$n = 20$	$n = 25$	$n = 30$	$n = 35$	$n = 40$	$n = 45$	$n = 50$
$\sigma = 1°$	0	10000	10000	10000	9998	9998	9999	9997	9996	9991
$\sigma = 2°$	0	10000	10000	10000	10000	9997	9998	9997	9995	9990
$\sigma = 3°$	0	10000	10000	9998	10000	9999	9997	9993	9991	9987
$\sigma = 4°$	0	10000	10000	9999	9998	9995	9994	9990	9992	9985
$\sigma = 5°$	0	10000	9999	10000	10000	9999	9997	9996	9987	9991

Table 13 Percent deviation between *Fmincon* and QuateRA for $dt = 0.1$s and $\Omega = 0.1$rad (there is no data for $n = 5$).

	$n = 5$	$n = 10$	$n = 15$	$n = 20$	$n = 25$	$n = 30$	$n = 35$	$n = 40$	$n = 45$	$n = 50$
$\sigma = 1°$	x	-0.54	-0.76	-0.72	-0.97	0.15	-0.51	-0.10	0.04	0.29
$\sigma = 2°$	x	-7.35	-1.17	-1.34	-0.18	0.20	0.25	-0.12	-0.32	-0.24
$\sigma = 3°$	x	-9.42	-4.97	-2.05	-0.58	-0.19	-0.25	0.36	-0.24	0.22
$\sigma = 4°$	x	-7.34	-7.54	-4.40	-1.87	-0.51	-0.32	-0.36	-0.24	-0.25
$\sigma = 5°$	x	-6.61	-6.49	-4.87	-2.42	-0.65	-1.53	-0.57	-0.76	-0.63

Table 14 Percent deviation between *Fmincon* and QuateRA for $dt = 1$s and $\Omega = 0.1$rad (there is no data for $n = 5$).

	$n = 5$	$n = 10$	$n = 15$	$n = 20$	$n = 25$	$n = 30$	$n = 35$	$n = 40$	$n = 45$	$n = 50$
$\sigma = 1°$	x	-0.64	-0.43	0.23	0.20	-0.06	-0.43	0.10	0.03	0.10
$\sigma = 2°$	x	-0.46	1.21	0.74	0.28	0.82	0.13	21.77	21.42	62.82
$\sigma = 3°$	x	0.49	0.11	0.04	-0.72	19.72	19.32	77.92	81.65	92.21
$\sigma = 4°$	x	-0.43	0.49	-0.51	29.00	66.02	74.48	89.94	93.99	96.35
$\sigma = 5°$	x	0.63	0.49	7.79	64.48	85.36	88.25	95.41	96.64	97.49

Table 15 Percent deviation between *Fmincon* and QuateRA for $dt = 1$s and $\Omega = 1$rad (there is no data for $n = 5$).

	$n = 5$	$n = 10$	$n = 15$	$n = 20$	$n = 25$	$n = 30$	$n = 35$	$n = 40$	$n = 45$	$n = 50$
$\sigma = 1°$	x	0.97	0.06	0.47	-0.15	0.09	0.03	0.41	0.23	47.83
$\sigma = 2°$	x	0.53	-0.11	0.20	0.36	0.37	-0.39	0.39	31.96	18.91
$\sigma = 3°$	x	-0.65	-0.50	-0.23	-0.19	-0.06	0.18	12.43	16.98	65.44
$\sigma = 4°$	x	-0.21	0.16	-0.34	-0.20	-0.04	12.63	35.66	52.67	71.67
$\sigma = 5°$	x	-0.20	-0.31	-0.17	9.34	14.84	32.94	58.70	65.84	81.74