**AAS 19-654**

# QUATERA: THE QUATERNION REGRESSION ALGORITHM

**Marcelino M. de Almeida,* Daniele Mortari,† Renato Zanetti,‡ Maruthi Akella§**

This work proposes a batch solution to the problem of estimating fixed angular velocity using orientation measurements. Provided that the angular velocity remains constant, 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 derive certain important statistical properties of the problem, and draw parallels to the relatively straightforward problem of estimating constant translational velocity from position measurements. We also perform a Monte Carlo analysis of the proposed algorithm, validating our method.

## INTRODUCTION

This paper presents a batch solution to the problem of angular velocity estimation using a time-sequence of orientation measurements. Our approach is motivated by the constant translational velocity estimation problem, whose solution is well known and has well-understood statistical properties [6]. Surprisingly, the rotational counterpart is significantly more challenging and has not yet been solved in a batch sense (to the best of our knowledge). Based on reasonable assumptions for the quaternion noise measurement model, we employ a Total Least Squares (TLS) cost function to derive a closed-form solution of the constant angular velocity estimation problem without the need to use iterative algorithms with no closed form solution.

The problem of estimating the angular velocity under pure spin is a very 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 can be 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 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 [26, 22] 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.

*PhD Candidate, 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.

In Ref. [24], the author proposes generalizations to Wahba's problem by accepting sequential vector measurements instead of the traditional simultaneous ones (see Ref. [18] 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. [24] 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. [27] 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. [24]. A solution to the three-vector measurement problem is provided in Ref. [11] and a numerical solution is provided in Ref. [25] 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)[14, 10, 16], since these rely on kinematics only. These methods should usually converge if properly initialized and iterated through a backward smoothing process [23]. Iterated nonlinear programming methods present the drawback that these might converge to local minima, unless proven otherwise. Our 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). Our QuateRA builds upon the work of Ref. [20], and it is a batch solver that 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 an 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 [15, 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, validating those with Monte Carlo simulations.

QuateRA's AOR estimation was first introduced by Ref. [20], and experimental validation was presented in Ref. [2]. Ref. [3] 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 the quaternion unit norms), and reach the same solution suggested by Ref. [20] under the assumption of small angle approximation for the quaternion measurement noise.
- None of the previous works 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. [20] suggested the use of performing "*dirty*" derivatives on the most recently measured quaternions. In contrast, the work of [2] showed that one can often obtain better results by pre-filtering the measured quaternions before employing the derivative. The AVM estimation in Ref. [3] is performed by using a modified MEKF. The AVM estimation suggested by Ref. [20] is actually biased under mild measurement noise, while the solutions presented in Refs. [2] and [3] 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: the following sections presents a motivation to the angular velocity estimation problem by introducing solutions to the simple problem of estimating constant linear velocity from position measurements. Some of the insights therein are crucial for understanding QuateRA's statistical properties. The subsequent section introduces the rotational attitude kinematics, describing some notations and parametrizations, as well as the assumed measurement model. Next, we present the QuateRA algorithm, followed by a Monte Carlo analysis of the proposed algorithm. Finally, the final section presents conclusions for this work.

## MOTIVATION: BATCH ESTIMATION OF LINEAR VELOCITY FROM POSITION MEASUREMENTS

Assume a point mass moving on the $xy$ plane with unknown constant velocity $\boldsymbol{v} = \begin{bmatrix} v_x & v_y \end{bmatrix}^T$. The position of the body is denoted as $\boldsymbol{p} = \begin{bmatrix} x & y \end{bmatrix}^T$. The kinematics of the problem is described as:

$$\boldsymbol{p}(t) = \boldsymbol{p}_0 + \boldsymbol{v}t, \tag{1}$$

where $t$ denotes time and $\boldsymbol{p}_0 \triangleq \begin{bmatrix} x_0 & y_0 \end{bmatrix}^T$ is the position of the system at time $t = 0$. The goal of this section is to estimate the vector $\boldsymbol{X} = \begin{bmatrix} \boldsymbol{p}_0^T & \boldsymbol{v}^T \end{bmatrix}^T$ through LS and TLS, drawing parallels between the two approaches.

We denote an estimated variable as $\hat{(\cdot)}$ ($\hat{x}(t)$ is an estimate of $x(t)$ and $\hat{y}(t)$ is an estimate of $y(t)$), and a measured variable as $\bar{(\cdot)}$ ($\bar{\boldsymbol{p}}$ is a measurement of $\boldsymbol{p}$ and $\bar{\boldsymbol{v}}$ is a measurement of $\boldsymbol{v}$). We use *star* notation $(\cdot)^*$ with variables with general covariance to distinguish them from their counterpart with normalized covariance ($\text{cov}[\boldsymbol{p}^*]$ is a positive-definite matrix, while $\text{cov}[\boldsymbol{p}] = \boldsymbol{I}$, where $\boldsymbol{I}$ is an identity matrix). The notation $\overrightarrow{(\cdot)}$ is used to denote unit-norm vectors ($\overrightarrow{\boldsymbol{x}}$ satisfies $||\overrightarrow{\boldsymbol{x}}||_2 = 1$). In addition, for simplicity of notation, we denote $\boldsymbol{p}_i = \boldsymbol{p}(t_i)$.

Assume that we measure the position of this system at $n$ different instants of time $t_i, i \in \{1, \cdots, n\}$. The measurement model is given by:

$$\begin{cases} \bar{x}(t_i) = x(t_i) + \epsilon_x(t_i) \\ \bar{y}(t_i) = y(t_i) + \epsilon_y(t_i) \end{cases}, \tag{2}$$

where $\boldsymbol{\epsilon}_i^* \triangleq \begin{bmatrix} \epsilon_x(t_i) & \epsilon_y(t_i) \end{bmatrix}^T$ is assumed to be a normally distributed random variable with mean $\mathbb{E}[\epsilon_i^*] = \boldsymbol{0}$ and covariance $\boldsymbol{P}_\epsilon \triangleq \text{cov}[\boldsymbol{\epsilon}_i^*] = \mathbb{E}[\boldsymbol{\epsilon}_i^* \boldsymbol{\epsilon}_i^{*T}]$, with $\boldsymbol{P}_\epsilon > 0$. We denote the measured position $\bar{\boldsymbol{p}}_i^* = \boldsymbol{p}_i + \boldsymbol{\epsilon}_i^*$, which is a random variable with mean $\mathbb{E}[\bar{\boldsymbol{p}}_i^*] = \boldsymbol{p}_i$ and covariance $\text{cov}[\bar{\boldsymbol{p}}_i^*] = \boldsymbol{P}_\epsilon$. Decomposing the covariance matrix as $\boldsymbol{P}_\epsilon = \boldsymbol{L}\boldsymbol{L}^T$, we define the normalized measurements $\bar{\boldsymbol{p}}_i = \boldsymbol{L}^{-1}\bar{\boldsymbol{p}}_i^*$ such that $\bar{\boldsymbol{p}}_i = \boldsymbol{L}^{-1}\boldsymbol{p}_i + \boldsymbol{L}^{-1}\boldsymbol{\epsilon}_i^*$. Defining $\boldsymbol{\epsilon}_i = \boldsymbol{L}^{-1}\boldsymbol{\epsilon}_i^*$, we have that $\mathbb{E}[\epsilon_i] = \boldsymbol{0}$ and $\text{cov}[\boldsymbol{\epsilon}_i] = \mathbb{E}[\boldsymbol{\epsilon}_i \boldsymbol{\epsilon}_i^T] = \boldsymbol{I}_2$, where $\boldsymbol{I}_2$ is the two-dimension identity matrix.

The vector of normalized measured positions is written as $\bar{\boldsymbol{P}} \triangleq \begin{bmatrix} \bar{\boldsymbol{p}}_1^T & \bar{\boldsymbol{p}}_2^T & \cdots & \bar{\boldsymbol{p}}_n^T \end{bmatrix}^T$, and the equivalent vector of normalized true positions is given by $\boldsymbol{P} \triangleq \begin{bmatrix} (\boldsymbol{L}^{-1}\boldsymbol{p}_1)^T & (\boldsymbol{L}^{-1}\boldsymbol{p}_2)^T & \cdots & (\boldsymbol{L}^{-1}\boldsymbol{p}_n)^T \end{bmatrix}^T$. The measurement error vector is written as $\boldsymbol{\varepsilon} \triangleq \begin{bmatrix} \boldsymbol{\epsilon}_1^T & \boldsymbol{\epsilon}_2^T & \cdots & \boldsymbol{\epsilon}_n^T \end{bmatrix}^T$, implying $\bar{\boldsymbol{P}} = \boldsymbol{P} + \boldsymbol{\varepsilon}$. Since $\mathbb{E}[\varepsilon] = \boldsymbol{0}$, then we have that $\mathbb{E}[\bar{\boldsymbol{P}}] = \boldsymbol{P}$ and $\text{cov}[\bar{\boldsymbol{P}}] = \text{cov}[\boldsymbol{\varepsilon}] = \boldsymbol{I}_{2n}$.

Given the measurement vector $\bar{\boldsymbol{P}}$, we want to optimally estimate the system's initial position $\boldsymbol{p}_0$ and velocity $\boldsymbol{v}$. A common method to solve this problem is to use the least squares solution, which pursues to find optimal $\boldsymbol{p}_0$ and $\boldsymbol{v}$ that minimizes the cost function:

$$J = \frac{1}{2}\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon} = \frac{1}{2}\left(\bar{\boldsymbol{P}} - \boldsymbol{P}\right)^T\left(\bar{\boldsymbol{P}} - \boldsymbol{P}\right). \tag{3}$$

The solution to this problem is very well known in the literature. Constructing the matrix $\boldsymbol{H} \in \mathbb{R}^{2n \times 4}$:

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{L}^{-1} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{L}^{-1} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{L}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_2 & t_1\boldsymbol{I}_2 \\ \boldsymbol{I}_2 & t_2\boldsymbol{I}_2 \\ \vdots & \vdots \\ \boldsymbol{I}_2 & t_n\boldsymbol{I}_2 \end{bmatrix}, \tag{4}$$

we have that $\boldsymbol{P} = \boldsymbol{H}\boldsymbol{X}$. The optimal solution* $\hat{\boldsymbol{X}}_{LS} = \begin{bmatrix}\hat{\boldsymbol{p}}_0^T & \hat{\boldsymbol{v}}^T\end{bmatrix}$ for the cost function in Eq. 3 is given by:

$$\hat{\boldsymbol{X}}_{LS} = \left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\bar{\boldsymbol{P}}. \tag{5}$$

Although Eq. 5 is a very common method to estimate the unknowns from Eq. 1, it is also possible to obtain solutions that minimize cost functions different from Eq. 3.

In particular, one can pursue a solution through Total Least Squares (TLS - also referred to as Orthogonal Least Squares), as opposed to Least Squares (LS). Starting from Eq. 1, we have that:

$$x(t) = x_0 + v_x \cdot t \tag{6}$$
$$y(t) = y_0 + v_x \cdot t \tag{7}$$

Isolating the $t$ term in Eq. 6 and substituting it into Eq. 7 leads to:

$$y = y_0 + \frac{v_y}{v_x}\left(x - x_0\right) = \left(y_0 - \tfrac{v_y}{v_x}x_0\right) + \frac{v_y}{v_x}x \tag{8}$$

Defining $\alpha \triangleq y_0 - \frac{v_y}{v_x}x_0$ and $\beta = \frac{v_y}{v_x}$, then Eq. 8 can be written in the compact form:

$$y = \alpha + \beta x, \tag{9}$$

and the unknowns to be found are now $\alpha$ and $\beta$. The problem can be recast as finding the Cartesian line $L(\boldsymbol{l}_0, \overrightarrow{\boldsymbol{l}})$ ($\boldsymbol{l}_0 \in \mathbb{R}^2$ is a point belonging to the line, and $\overrightarrow{\boldsymbol{l}} \in \mathbb{S}^1$ is the line direction) such that the distance squared between the regularized measured points $\bar{\boldsymbol{p}}_i, i \in \{1, \cdots, n\}$ and the line $L(\boldsymbol{l}_0, \overrightarrow{\boldsymbol{l}})$ are minimized. The distance function used in TLS is not necessarily the *Euclidian* distance between a point and a line, unless the error covariance is of the form $\boldsymbol{P}_\epsilon = \sigma^2\boldsymbol{I}_2$, where $\sigma \in \mathbb{R}_{>0}$.

For general values of the covariance matrix, we pursue as in LS by covariance-normalizing the measurements $\bar{\boldsymbol{p}}_i = \boldsymbol{L}^{-1}\bar{\boldsymbol{p}}_i^*$, where $\boldsymbol{L}$ comes from the decomposition of $\boldsymbol{P}_\epsilon = \boldsymbol{L}\boldsymbol{L}^T$. Defining $d\left(\bar{\boldsymbol{p}}_i, L\right)$ as the *Euclidian* distance between $\bar{\boldsymbol{p}}_i$ and $L(\boldsymbol{l}_0, \overrightarrow{\boldsymbol{l}})$, the TLS cost function is given by:

$$J_{TLS} = \sum_{i=1}^{n} d\left(\bar{\boldsymbol{p}}_i, L\right)^2. \tag{10}$$

The regression problem for the cost of Eq. 10 was first proposed and solved in [1] for the special case in which $\boldsymbol{P}_\epsilon = \sigma^2\boldsymbol{I}_2$. Many solution formulations have been presented to this problem for the general case (see [29, 19, 7] for literature review), but here we present the solution presented in [28] due to its connections to the QuateRA problem. First, we calculate the centroid of all the data-points:

$$\boldsymbol{\mu}_p \triangleq \frac{1}{n}\sum_{i=1}^{n}\bar{\boldsymbol{p}}_i. \tag{11}$$

It turns out that the optimal line $\hat{L}(\boldsymbol{l}_0, \overrightarrow{\boldsymbol{l}})$ passes through the centroid $\boldsymbol{\mu}_p$. Since a line is defined as a point and a direction, the solution is complete once the line direction is found. To this purpose, we define the translated vectors $\underline{\boldsymbol{p}}_i$:

$$\underline{\boldsymbol{p}}_i \triangleq \bar{\boldsymbol{p}}_i - \boldsymbol{\mu}_p, \quad \forall i \in \{1, \cdots, n\} \tag{12}$$

Clearly, the centroid of the set of vectors $\underline{\boldsymbol{p}}_i, i \in \{1, \cdots, n\}$ is at the origin. Then, we define the matrix $\boldsymbol{B} \in \mathbb{R}^{2\times n}$ as a concatenation of all translated vectors $\underline{\boldsymbol{p}}_i$:

$$\boldsymbol{B} \triangleq \begin{bmatrix}\underline{\boldsymbol{p}}_1 & \underline{\boldsymbol{p}}_2 & \cdots & \underline{\boldsymbol{p}}_n\end{bmatrix} \tag{13}$$

*We use the subscript LS to indicate that this is the Linear Squares solution to the problem.

Taking the Singular Value Decomposition (SVD) on the matrix $\boldsymbol{B}$, we get $\boldsymbol{B} = \boldsymbol{U}\tilde{\boldsymbol{\Sigma}}\boldsymbol{V}^T$, where $\boldsymbol{U} = \begin{bmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2\end{bmatrix}$ contains the *left singular vectors* of $\boldsymbol{B}$, $\boldsymbol{V} = \begin{bmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots \boldsymbol{v}_n\end{bmatrix}$ contains the *right singular vectors* of $\boldsymbol{B}$, and $\tilde{\boldsymbol{\Sigma}} = \begin{bmatrix}\boldsymbol{\Sigma} & \boldsymbol{0}_{2\times n-2}\end{bmatrix}$ contains the singular values of $\boldsymbol{B}$ within $\boldsymbol{\Sigma} = \mathrm{diag}\left(\sigma_1, \sigma_2\right)$.

As shown in Ref. [28], the line that minimizes the cost function of Eq. 10 is parameterized as $\hat{L}(\boldsymbol{\mu}_p, \boldsymbol{u}_1)$, where $\boldsymbol{u}_1 \in \mathbb{S}^1$ is the first left singular vector of $\boldsymbol{B}$, and the optimal cost is given by $\hat{J}_{TLS} = \sigma_2$. The problem can then be mapped back into the original coordinates:

$$\boldsymbol{\mu}_p^* = \begin{bmatrix}\mu_{px}^* \\ \mu_{py}^*\end{bmatrix} = \boldsymbol{L}\boldsymbol{\mu}_p, \qquad \boldsymbol{u}_1^* = \begin{bmatrix}u_{1x}^* \\ u_{1y}^*\end{bmatrix} = \frac{\boldsymbol{L}\boldsymbol{u}_1}{\left\|\boldsymbol{L}\boldsymbol{u}_1\right\|}. \tag{14}$$

Thus, the TLS minimizer that fits the model of Eq. 9 given the measurements of Eq. 2 and the measurement noise covariance $\boldsymbol{P}_\epsilon$ is given by the line $\hat{L}^*(\boldsymbol{\mu}_p^*, \boldsymbol{u}_1^*)$. The constants $\alpha$ and $\beta$ from Eq. 9 can be calculated as:

$$\beta = \frac{u_{1y}^*}{u_{1x}^*}, \qquad \alpha = \mu_{py}^* - \beta\mu_{px}^*. \tag{15}$$

Going back to the original problem of estimating the velocity in Eq. 1, the vector $\boldsymbol{u}_1^*$ is an estimate of the velocity direction $\overrightarrow{v} \triangleq \boldsymbol{v}/||\boldsymbol{v}||$. In order to obtain the estimation of the velocity $\boldsymbol{v}$, one still needs to estimate the velocity magnitude $||\boldsymbol{v}||$.

We can project the measured points onto the optimal line $\boldsymbol{\mu}_p^*, \boldsymbol{u}_1^*$, obtaining the TLS estimates for these points along the line. Then, the velocity magnitude can be estimated as an average displacement along the line. Given that the measurements are distributed as $\bar{\boldsymbol{p}}_i^* \sim \mathcal{N}(\boldsymbol{p}_i, \boldsymbol{P}_\epsilon)$, it is possible to show that the marginal distribution of $\bar{\boldsymbol{p}}_i^*$ along any line $L(\boldsymbol{l}_0, \overrightarrow{\boldsymbol{l}})$ is a one-dimensional normally-distributed random variable with mean at $\boldsymbol{p}_{pi}^*$ and standard deviation $\sigma$ along the $\overrightarrow{\boldsymbol{l}}$ direction, where:

$$\boldsymbol{p}_{pi}^* = \boldsymbol{l}_0 + \frac{1}{\sigma^2}\left[\left(\bar{\boldsymbol{p}}_i^* - \boldsymbol{l}_0\right)^T \boldsymbol{P}^{-1}\overrightarrow{\boldsymbol{l}}\right]\overrightarrow{\boldsymbol{l}}, \qquad \sigma = \frac{1}{\left\|\boldsymbol{L}^{-1}\overrightarrow{\boldsymbol{l}}\right\|} = \frac{1}{\sqrt{\overrightarrow{\boldsymbol{l}}\boldsymbol{P}^{-1}\overrightarrow{\boldsymbol{l}}}}. \tag{16}$$

Hence, defining $\bar{S}_i = \boldsymbol{p}_{pi}^*\boldsymbol{u}_1^*$ as the displacement along the optimal TLS line, and admitting the distribution $\bar{S}_i \sim \mathcal{N}(S_i, \sigma^2)$, one can use LS to solve for $S_0$ and $||\boldsymbol{v}||$ in the model:

$$S_i = S_0 + ||\boldsymbol{v}|| \cdot t_i. \tag{17}$$

**Analysis**

It turns out that the solution obtained through LS (Eq. 5) is generally different from the one obtained through TLS (solution of Eq. 17 and the first left singular vector of the matrix in Eq. 13). The different solutions are expected, given that both estimators employ different cost functions.

For the particular scenario of estimating the planar system's velocity of Eq. 1, the LS solution is more advantageous than TLS in many aspects, some of which are described below. Assuming a linear model (as in Eq. 1) with additive gaussian measurement noise (as in Eq. 2), LS is a maximum likelyhood estimator, implying:

- LS is known to be the globally optimal estimator that obtains the Minimum Mean Square Error (MMSE) of the estimate, i.e., it minimizes $MSE = \mathbb{E}\left[\left(\bar{\boldsymbol{X}} - \boldsymbol{X}\right)^T\left(\bar{\boldsymbol{X}} - \boldsymbol{X}\right)\right]$. This implies that, in average, no other estimator will perform as good as LS for minimization of MSE. In other words, the LS solution will produce smaller squared error more than 50% of the time (in average) when compared with any other estimator.
- LS is well known for being an unbiased estimator given zero-mean additive noise to the measurements. On the other hand, TLS is only guaranteed to be strongly consistent, i.e., the TLS estimate converges to the true value (with probability 1) as the number of measurements $n$ tend to infinity [12, 8], meaning

that it is asymptotically unbiased. Monte Carlo analysis suggest that the bias of TLS is statistically appreciable when signal-to-noise ratio is low, and $n$ is small [29]. On the other hand, the Monte Carlo analysis in [13] indicate that $n > 20$ is large enough to neglect the TLS bias.

- The error-covariance for TLS estimates are known for $n \to \infty$, while the error-covariance of LS is known for any $n$. However, the Monte Carlo analysis in [13] suggest that the TLS error-covariance estimation for $n \to \infty$ is a good approximation for $n < \infty$ provided that $n > 20$. Ref [8] derives a TLS covariance matrix for large samples.

- The LS estimate of the velocity magnitude $||\boldsymbol{v}||$ using Eq. 17 assumes that the velocity direction $\overrightarrow{\boldsymbol{v}}$ is precisely known. However, as already mentioned, TLS provides a biased estimate $\overrightarrow{\boldsymbol{v}} = \boldsymbol{u}_1$, which can also implicate on a biased estimation of $||\boldsymbol{v}||$.

Based on the comparisons above, there is no compelling reason to convert the model of Eq. 1 into the form of Eq. 9, and then perform TLS. On the other hand, provided that measurement noise is sufficiently small, and the number of measurements are large enough (e.g., say $n > 20$), then then TLS is a competitive algorithm that matches closely the LS solution in the MSE sense (i.e., it outperforms LS in the MSE sense almost 50% of the time).

As a motivational example, assume a system moving on a line with initial position $\boldsymbol{p}_0 = \begin{bmatrix}1 & 0\end{bmatrix}^T$m and velocity $\boldsymbol{v} = \begin{bmatrix}2 & 1\end{bmatrix}^T$m. The measurement error standard deviation is given by $\sigma_\epsilon = 0.1$m. The measurements are taken once every $dt \triangleq t_{i+1} - t_i = 0.1$s and the regression is made with $n = 20$ measurements. Running a Monte Carlo simulation of $100.000$ solutions, it turns out that LS outperforms TLS $51.05\%$ of the time in the estimated squared error sense. If the measurement error standard deviation degrades to $\sigma_\epsilon = 0.5$m, then LS outperforms TLS $55.44\%$ of the time. By taking $n = 50$ measurements with $\sigma_\epsilon = 0.1$m, LS outperforms TLS $50.34\%$ of the time.

Despite of the possible limitations of TLS, we employ the TLS cost function in the development of QuateRA. This choice is made because it is then possible to decouple the estimation of the angular velocity axis of rotation from its magnitude, whereas the estimation of the coupled problem (which would be the LS counterpart) is substantially more complex. When estimating a fixed axis of rotation among sequential quaternion measurements, the estimation problem can be posed as a plane fitting problem (special case of TLS), as will be shown in the following sections.

## ATTITUDE KINEMATICS AND MEASUREMENT MODEL

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:

$$\boldsymbol{q}_A^B = \begin{bmatrix} q_{As}^B & (\boldsymbol{q}_{Av}^B)^T \end{bmatrix}^T, \tag{18}$$

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:

$$\boldsymbol{q}_B^A = \begin{bmatrix} q_{As}^B & -(\boldsymbol{q}_{Av}^B)^T \end{bmatrix}^T. \tag{19}$$

The quaternion composition rule is denoted as:

$$\boldsymbol{q}_A^C = \boldsymbol{q}_B^C \otimes \boldsymbol{q}_A^B, \qquad \boldsymbol{q}_B^C \otimes = \begin{bmatrix} q_{Bs}^C & -(\boldsymbol{q}_{Bv}^C)^T \\ \boldsymbol{q}_{Bv}^C & q_{Bs}^C \boldsymbol{I} - [\boldsymbol{q}_{Bv\times}^C] \end{bmatrix}, \tag{20}$$

where $\boldsymbol{I}$ is the $3 \times 3$ 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*:

$$\boldsymbol{q}_I \triangleq (\boldsymbol{q}_A^B)^{-1} \otimes \boldsymbol{q}_A^B = \boldsymbol{q}_A^B \otimes (\boldsymbol{q}_A^B)^{-1} = \begin{bmatrix}1 & 0 & 0 & 0\end{bmatrix}^T \tag{21}$$

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

$$\boldsymbol{v}\otimes \triangleq \begin{bmatrix}0 & -\boldsymbol{v}^T \\ \boldsymbol{v} & -[\boldsymbol{v}_\times]\end{bmatrix}. \tag{22}$$

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:

$$\boldsymbol{q} \otimes \boldsymbol{v} \triangleq \boldsymbol{q} \otimes \begin{bmatrix}0 \\ \boldsymbol{v}\end{bmatrix}. \tag{23}$$

Given a vector $\boldsymbol{v}^A \in \mathbb{R}^3$ expressed in frame $A$, its representation in frame $B$ can be obtained as:

$$\begin{bmatrix}0 \\ \boldsymbol{v}^B\end{bmatrix} = \boldsymbol{q}_A^B \otimes \boldsymbol{v}^A \otimes (\boldsymbol{q}_A^B)^{-1}. \tag{24}$$

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:

$$\dot{\boldsymbol{q}}_A^B = \frac{1}{2}\boldsymbol{\omega}_{B/A}^B \otimes \boldsymbol{q}_A^B. \tag{25}$$

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:

$$\Omega_{B/A} \triangleq \left\|\boldsymbol{\omega}_{B/A}^B\right\|, \qquad \overrightarrow{\boldsymbol{\omega}}_{B/A}^B \triangleq \frac{\boldsymbol{\omega}_{B/A}^B}{\Omega_{B/A}}. \tag{26}$$

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. 25 is given by $\boldsymbol{q}_A^B(t_f) = \boldsymbol{F}(\boldsymbol{\omega}_{B/A}^B) \cdot \boldsymbol{q}_A^B(t_0)$, where:

$$\boldsymbol{F}\left(\boldsymbol{\omega}_{B/A}^B\right) = \exp\left[\tfrac{1}{2}\boldsymbol{\omega}_{B/A}^B\otimes\right] = \cos\frac{\Omega_{B/A}\Delta t}{2} \cdot \boldsymbol{I} + \sin\frac{\Omega_{B/A}\Delta t}{2} \cdot \overrightarrow{\boldsymbol{\omega}}_{B/A}^B \otimes. \tag{27}$$

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 $\Omega \triangleq \Omega_{O/I}$.

**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:

$$\bar{\boldsymbol{q}}_i = \boldsymbol{q}_i \otimes \boldsymbol{q}_{Ni}, \tag{28}$$

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:

$$\boldsymbol{q}_{Ni} \triangleq \left[\cos\tfrac{\theta_i}{2} \quad \boldsymbol{e}_{Ni}^T \sin\tfrac{\theta_i}{2}\right]^T, \tag{29}$$

in which $\theta_i$ and $\boldsymbol{e}_{Ni}$ are independent random variables. We assume that $\theta_i$ is Gaussian (Although it might be unrealistic to assume that angles are distributed as Gaussian, Ref. [21] 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:

$$\boldsymbol{\mu}_N \triangleq \mathbb{E}\left[\boldsymbol{q}_{Ni}\right] = \mathbb{E}\begin{bmatrix}\cos\frac{\theta_i}{2}\\ \boldsymbol{e}_{Ni}\sin\frac{\theta_i}{2}\end{bmatrix} = \begin{bmatrix}\mathbb{E}\left[\cos\frac{\theta_i}{2}\right]\\ \mathbb{E}\left[\boldsymbol{e}_{Ni}\sin\frac{\theta_i}{2}\right]\end{bmatrix} = \begin{bmatrix}\mathbb{E}\left[\cos\frac{\theta_i}{2}\right]\\ \mathbb{E}\left[\boldsymbol{e}_{Ni}\right]\mathbb{E}\left[\sin\frac{\theta_i}{2}\right]\end{bmatrix} = \mathbb{E}\left[\cos\frac{\theta_i}{2}\right]\begin{bmatrix}1\\ \boldsymbol{0}\end{bmatrix} \tag{30}$$

$$\begin{aligned}\boldsymbol{P}_N &\triangleq \mathbb{E}\left[\left(\boldsymbol{q}_{Ni} - \boldsymbol{\mu}_N\right)\left(\boldsymbol{q}_{Ni} - \boldsymbol{\mu}_N\right)^T\right] = \mathbb{E}\left[\boldsymbol{q}_{Ni}\boldsymbol{q}_{Ni}^T\right] - \boldsymbol{\mu}_N\boldsymbol{\mu}_N^T\\ &= \begin{bmatrix}\mathbb{E}\left[\cos^2\frac{\theta_i}{2}\right] - \mathbb{E}^2\left[\cos\frac{\theta_i}{2}\right] & \mathbb{E}\left[\boldsymbol{e}_{Ni}\cos\frac{\theta_i}{2}\sin\frac{\theta_i}{2}\right]\\ \mathbb{E}\left[\boldsymbol{e}_{Ni}^T\cos\frac{\theta_i}{2}\sin\frac{\theta_i}{2}\right] & \mathbb{E}\left[\boldsymbol{e}_{Ni}\boldsymbol{e}_{Ni}^T\right]\mathbb{E}\left[\sin^2\frac{\theta_k}{2}\right]\end{bmatrix}\\ &= \begin{bmatrix}\mathbb{E}\left[\cos^2\frac{\theta_i}{2}\right] - \mathbb{E}^2\left[\cos\frac{\theta_i}{2}\right] & \boldsymbol{0}\\ \boldsymbol{0} & \frac{1}{3}\mathbb{E}\left[\sin^2\frac{\theta_i}{2}\right]\boldsymbol{I}_3\end{bmatrix}\end{aligned} \tag{31}$$

The expected values above can be calculated according with Ref. [21]: $\mathbb{E}\left[\cos\frac{\theta_i}{2}\right] = e^{-\sigma_\theta^2/8}$, $\mathbb{E}\left[\cos^2\frac{\theta_i}{2}\right] = \frac{1}{2}\left(1 + e^{-\sigma_\theta^2/2}\right)$, and $\mathbb{E}\left[\sin^2\frac{\theta_i}{2}\right] = \frac{1}{2}\left(1 - e^{-\sigma_\theta^2/2}\right)$. Defining $\sigma_s^2 \triangleq \mathbb{E}\left[\cos^2\frac{\theta_i}{2}\right] - \mathbb{E}^2\left[\cos\frac{\theta_k}{2}\right]$ and $\sigma_v^2 \triangleq \frac{1}{3}\mathbb{E}\left[\sin^2\frac{\theta_i}{2}\right]$, then the noise covariance matrix takes the form:

$$\boldsymbol{P}_N = \begin{bmatrix}\sigma_s^2 & \boldsymbol{0}\\ \boldsymbol{0} & \sigma_v^2\boldsymbol{I}_3\end{bmatrix}. \tag{32}$$

We define the covariance for the measured quaternion as:

$$\boldsymbol{P}_q \triangleq \mathbb{E}\left[\left(\bar{\boldsymbol{q}}_i - \mathbb{E}\left[\bar{\boldsymbol{q}}_i\right]\right)\left(\bar{\boldsymbol{q}}_i - \mathbb{E}\left[\bar{\boldsymbol{q}}_i\right]\right)^T\right] = \mathbb{E}\left[\bar{\boldsymbol{q}}_i\bar{\boldsymbol{q}}_i^T\right] - \mathbb{E}\left[\bar{\boldsymbol{q}}_i\right]\mathbb{E}\left[\bar{\boldsymbol{q}}_i\right]^T \tag{33}$$

$$= \left(\boldsymbol{q}_i\otimes\right)\mathbb{E}\left[\boldsymbol{q}_{Ni}\boldsymbol{q}_{Ni}^T\right]\left(\boldsymbol{q}_i\otimes\right)^T - \left(\boldsymbol{q}_i\otimes\right)\boldsymbol{\mu}_N\boldsymbol{\mu}_N^T\left(\boldsymbol{q}_i\otimes\right)^T = \left(\boldsymbol{q}_i\otimes\right)\left[\mathbb{E}\left[\boldsymbol{q}_{Ni}\boldsymbol{q}_{Ni}^T\right] - \boldsymbol{\mu}_N\boldsymbol{\mu}_N^T\right]\left(\boldsymbol{q}_i\otimes\right)^T \tag{34}$$

$$= \left(\boldsymbol{q}_i\otimes\right)\boldsymbol{P}_N\left(\boldsymbol{q}_i\otimes\right)^T \tag{35}$$

If we make the notation relaxation $\boldsymbol{q}_i = \begin{bmatrix}q_s & \boldsymbol{q}_v^T\end{bmatrix}^T$, and use Eqs. **??** and 32, we can further expand $\boldsymbol{P}_q$ as:

$$\boldsymbol{P}_q = \begin{bmatrix}\sigma_s^2 q_s^2 + \sigma_v^2\boldsymbol{q}_v^T\boldsymbol{q}_v & \sigma_s^2 q_s\boldsymbol{q}_v^T - \sigma_v^2 q_s\boldsymbol{q}_v^T\\ \sigma_s^2 q_s\boldsymbol{q}_v - \sigma_v^2 q_s\boldsymbol{q}_v & \sigma_s^2\boldsymbol{q}_v\boldsymbol{q}_v^T - \sigma_v^2\left(q_s^2\boldsymbol{I}_3 - \left[\boldsymbol{q}_{v\times}\right]^2\right)\end{bmatrix}. \tag{36}$$

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:

$$\boldsymbol{P}_q = \begin{bmatrix}\sigma_v^2 + \left(\sigma_s^2 - \sigma_v^2\right)q_s^2 & \left(\sigma_s^2 - \sigma_v^2\right)q_s\boldsymbol{q}_v^T\\ \left(\sigma_s^2 - \sigma_v^2\right)q_s\boldsymbol{q}_v & \sigma_v^2\boldsymbol{I}_3 + \left(\sigma_s^2 - \sigma_v^2\right)\boldsymbol{q}_v\boldsymbol{q}_v^T\end{bmatrix} = \sigma_v^2\boldsymbol{I}_4 + \left(\sigma_s^2 - \sigma_v^2\right)\boldsymbol{q}_i\boldsymbol{q}_i^T. \tag{37}$$

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. 37, 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:

$$\boldsymbol{P}_q = \sigma_v^2\boldsymbol{I}_4 - 2\sigma_v^2\boldsymbol{q}_i\boldsymbol{q}_i^T + 2\sigma_v\sigma_s\boldsymbol{q}_i\boldsymbol{q}_i^T + \left(\sigma_s^2 - 2\sigma_v\sigma_s + \sigma_v^2\right)\boldsymbol{q}_i\boldsymbol{q}_i^T \tag{38}$$

$$= \sigma_v^2\boldsymbol{I}_4 - 2\sigma_v\left(\sigma_v - \sigma_s\right)\boldsymbol{q}_i\boldsymbol{q}_i^T + \left(\sigma_v - \sigma_s\right)^2\boldsymbol{q}_i\boldsymbol{q}_i^T. \tag{39}$$

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:

$$\boldsymbol{P}_q = \sigma_v^2\boldsymbol{I}_4 - 2\sigma_q\boldsymbol{q}_i\boldsymbol{q}_i^T + \sigma_q^2\boldsymbol{q}_i\boldsymbol{q}_i^T\boldsymbol{q}_i\boldsymbol{q}_i^T = \sigma_v^2\boldsymbol{I}_4 - 2\sigma_q\boldsymbol{q}_i\boldsymbol{q}_i^T + \sigma_q^2\left(\boldsymbol{q}_i\boldsymbol{q}_i^T\right)^2 \tag{40}$$

$$= \left(\sigma_v\boldsymbol{I}_4 - \sigma_q\boldsymbol{q}_i\boldsymbol{q}_i^T\right)^2 \tag{41}$$

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:

$$\boldsymbol{L}_q^{-1} = \frac{1}{\sigma_s\sigma_v}\left(\sigma_s\boldsymbol{I}_4 + \sigma_q\boldsymbol{q}_i\boldsymbol{q}_i^T\right). \tag{42}$$

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

$$\boldsymbol{L}_q^{-1}\boldsymbol{q}_i = \frac{1}{\sigma_s\sigma_v}\left(\sigma_s\boldsymbol{q}_i + \sigma_q\boldsymbol{q}_i\right) = \frac{\sigma_v}{\sigma_s\sigma_v}\boldsymbol{q}_i = \frac{1}{\sigma_s}\boldsymbol{q}_i. \tag{43}$$

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. 28 around $\theta_i = 0$, and pre-multiply by $\boldsymbol{L}_q^{-1}$, we get that:

$$\boldsymbol{L}_q^{-1}\bar{\boldsymbol{q}}_i = \boldsymbol{L}_q^{-1}\boldsymbol{q}_i \otimes \boldsymbol{q}_{Ni} = \boldsymbol{L}_q^{-1}\left(\boldsymbol{q}_i\otimes\right)\left(\boldsymbol{q}_I + \left.\tfrac{\partial \boldsymbol{q}_{Ni}}{\partial\theta_i}\right|_0\theta_i + \left.\tfrac{\partial^2 \boldsymbol{q}_{Ni}}{\partial\theta_i^2}\right|_0\theta_i^2 + \cdots\right) \tag{44}$$

$$= \boldsymbol{L}_q^{-1}\boldsymbol{q}_i + \boldsymbol{L}_q^{-1}\left(\boldsymbol{q}_i\otimes\right)\left(\left.\tfrac{\partial \boldsymbol{q}_{Ni}}{\partial\theta_i}\right|_0\theta_i + \left.\tfrac{\partial^2 \boldsymbol{q}_{Ni}}{\partial\theta_i^2}\right|_0\theta_i^2 + \cdots\right) \tag{45}$$

$$= \frac{1}{\sigma_s}\boldsymbol{q}_i + \boldsymbol{L}_q^{-1}\left(\boldsymbol{q}_i\otimes\right)\left(\left.\tfrac{\partial \boldsymbol{q}_{Ni}}{\partial\theta_i}\right|_0\theta_i + \left.\tfrac{\partial^2 \boldsymbol{q}_{Ni}}{\partial\theta_i^2}\right|_0\theta_i^2 + \cdots\right), \tag{46}$$

where $\boldsymbol{q}_I$ is the identity quaternion defined in Eq. 21.

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 [8]:

$$\boldsymbol{P}_q \approx \left(\sigma_v\boldsymbol{I}_4 - \sigma_q\bar{\boldsymbol{q}}_i\bar{\boldsymbol{q}}_i^T\right)^2 \qquad \Longrightarrow \qquad \boldsymbol{L}_q^{-1} \approx \frac{1}{\sigma_s\sigma_v}\left(\sigma_s\boldsymbol{I}_4 + \sigma_q\bar{\boldsymbol{q}}_i\bar{\boldsymbol{q}}_i^T\right), \tag{47}$$

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

## THE QUATERNION REGRESSION ALGORITHM

In this section, we develop the QuateRA algorithm. We employ a similar approach to the TLS development previously shown, which first calculates the linear velocity direction in the optimal line, then calculates the linear velocity magnitude along the same line. Similarly, QuateRA assumes constant $\boldsymbol{\omega}$ to first estimate the AOR $\hat{\vec{\boldsymbol{\omega}}}$, then uses its knowledge to estimate for the AVM $\hat{\Omega}$. Finally, the estimated angular velocity is given by $\hat{\boldsymbol{\omega}} = \hat{\Omega}\hat{\vec{\boldsymbol{\omega}}}$.

In order to estimate the AOR, QuateRA uses a geometric interpretation based on the solution to the quaternion kinematic equation for constant $\boldsymbol{\omega}$:

$$\boldsymbol{q}(t) = \left[\cos\tfrac{\Omega\Delta t}{2}\cdot\boldsymbol{I} + \sin\tfrac{\Omega\Delta t}{2}\cdot\vec{\boldsymbol{\omega}}\otimes\right]\boldsymbol{q}_0, \tag{48}$$

with $\Delta t \triangleq t - t_0$. Defining the vectors $\boldsymbol{u}_1 \in \mathbb{S}^3 = \boldsymbol{q}_0$ and $\boldsymbol{u}_2 \in \mathbb{S}^3 = \vec{\boldsymbol{\omega}} \otimes \boldsymbol{q}_0$, we have that $\boldsymbol{u}_1^T\boldsymbol{u}_2 = \boldsymbol{q}_0 \cdot \vec{\boldsymbol{\omega}} \otimes \cdot\boldsymbol{q}_0$. Since $\vec{\boldsymbol{\omega}}\otimes$ is a skew-symmetric matrix (see Eq. 22) then $\boldsymbol{u}_1^T\boldsymbol{u}_2 = 0$, i.e., $\boldsymbol{u}_1 \perp \boldsymbol{u}_2$. Clearly, any $\boldsymbol{q}(t)$ described by Eq. 48 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 $\boldsymbol{q}(t) \in \mathbb{P}(\boldsymbol{u}_1, \boldsymbol{u}_2)$, $\forall\, t \in \mathbb{R}$. In addition, there exists a perpendicular plane $\mathbb{P}(\boldsymbol{u}_3, \boldsymbol{u}_4) = \text{span}\{\boldsymbol{u}_3, \boldsymbol{u}_4\}$, with $\boldsymbol{u}_3, \boldsymbol{u}_4 \in \mathbb{S}^3$ such that $\boldsymbol{u}_4 = \overrightarrow{\boldsymbol{\omega}} \otimes \boldsymbol{u}_3$, where $\boldsymbol{u}_3^T \boldsymbol{q}(t) = \boldsymbol{u}_4^T \boldsymbol{q}(t) = 0$, $\forall\, t \in \mathbb{R}$.

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. Plane-fitting is a classical TLS problem given a measurement model as in Eq. **??** with a non-singular measurement error covariance matrix. We use the TLS cost function to derive a quaternion plane fitting formulation. In TLS, whenever the measurement covariance matrix is not of the form $\text{cov}\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\end{bmatrix} = \sigma^2 \boldsymbol{I}$, the measurements have to be normalized by the inverse of some "square root" of the covariance matrix. We have already shown in the previous section that the square-root normalization can be approximated to a constant scaling on all measured quaternions. If we assume that all measurements are identically distributed, then this is a weighed TLS problem where all the weights are identical, implying that the weights can be neglected on the cost function.

The remainder of this section is structured as follows: next subsection presents the AOR estimation algorithm, while the following one presents the AVM estimator. The subsequent subsection summarizes QuateRA into a few steps, while the last subsection presents some insights and analysis to the overall algorithm.

### Estimation of the Axis of Rotation

Given a set of quaternion measurements $\bar{\boldsymbol{q}}_i,\ i \in \{1, \cdots, n\}$, we construct the measurement matrix $\bar{\boldsymbol{Q}}$ as:

$$\bar{\boldsymbol{Q}} \triangleq \begin{bmatrix}\bar{\boldsymbol{q}}_1 & \bar{\boldsymbol{q}}_2 & \cdots & \bar{\boldsymbol{q}}_n\end{bmatrix}. \tag{49}$$

An important property that arises from the previous definition is that:

$$\text{tr}\left(\bar{\boldsymbol{Q}}\bar{\boldsymbol{Q}}^T\right) = \text{tr}\left(\textstyle\sum_{i=1}^n \bar{\boldsymbol{q}}_i \bar{\boldsymbol{q}}_1^T\right) = \sum_{i=1}^n \text{tr}\left(\bar{\boldsymbol{q}}_i \bar{\boldsymbol{q}}_1^T\right) = \sum_{i=1}^n \left\|\bar{\boldsymbol{q}}_i\right\|^2 = n \tag{50}$$

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:

$$J_1 = \frac{1}{2}\left\|\bar{\boldsymbol{Q}} - \hat{\boldsymbol{Q}}\right\|_F^2, \tag{51}$$

where the $||\cdot||_F$ denotes the *Frobenius norm* and $\hat{\boldsymbol{Q}}$ is defined as:

$$\hat{\boldsymbol{Q}} \triangleq \begin{bmatrix}\hat{\boldsymbol{q}}_1 & \hat{\boldsymbol{q}}_2 & \cdots & \hat{\boldsymbol{q}}_k\end{bmatrix}. \tag{52}$$

From the Frobenius norm definition, we have that:

$$\begin{aligned} J_1 &= \frac{1}{2}\,\text{tr}\left[\left(\bar{\boldsymbol{Q}} - \hat{\boldsymbol{Q}}\right)\left(\bar{\boldsymbol{Q}} - \hat{\boldsymbol{Q}}\right)^T\right] = \frac{1}{2}\,\text{tr}\left[\bar{\boldsymbol{Q}}\bar{\boldsymbol{Q}}^T - \bar{\boldsymbol{Q}}\hat{\boldsymbol{Q}}^T - \hat{\boldsymbol{Q}}\bar{\boldsymbol{Q}}^T + \hat{\boldsymbol{Q}}\hat{\boldsymbol{Q}}^T\right] \\ &= \frac{1}{2}\,\text{tr}\left(\bar{\boldsymbol{Q}}\bar{\boldsymbol{Q}}^T\right) - \frac{1}{2}\,\text{tr}\left(\bar{\boldsymbol{Q}}\hat{\boldsymbol{Q}}^T\right) - \frac{1}{2}\,\text{tr}\left(\hat{\boldsymbol{Q}}\bar{\boldsymbol{Q}}^T\right) + \frac{1}{2}\,\text{tr}\left(\hat{\boldsymbol{Q}}\hat{\boldsymbol{Q}}^T\right). \end{aligned} \tag{53}$$

Using the trace property $\text{tr}(\boldsymbol{AB}) = \text{tr}(\boldsymbol{BA})$, and the property of Eq. 50, we have that:

$$J_1 = n - \text{tr}\left(\bar{\boldsymbol{Q}}\hat{\boldsymbol{Q}}^T\right) = n - \text{tr}\left(\textstyle\sum_{i=1}^n \bar{\boldsymbol{q}}_i \hat{\boldsymbol{q}}_i^T\right) = n - \sum_{i=1}^n \text{tr}\left(\bar{\boldsymbol{q}}_i \hat{\boldsymbol{q}}_i^T\right) = n - \sum_{i=1}^n \bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{q}}_i. \tag{54}$$

*Minimizing* the cost function of Eq. 54 is equivalent to *maximizing* the following cost function:

$$J_2 = \sum_{i=1}^k \bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{q}}_i. \tag{55}$$

**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:*

$$J_0 = \frac{1}{2} \left\| \boldsymbol{q} - \boldsymbol{q}_p \right\|_2^2 = \frac{1}{2} \left\| \boldsymbol{q} - \boldsymbol{q}_p \right\|_F^2 \tag{56}$$

*is given by:*

$$\boldsymbol{q}_p = \frac{1}{\sqrt{\left(\boldsymbol{q}^T \boldsymbol{u}_1\right)^2 + \left(\boldsymbol{q}^T \boldsymbol{u}_2\right)^2}} \left[ \left(\boldsymbol{q}^T \boldsymbol{u}_1\right) \boldsymbol{u}_1 + \left(\boldsymbol{q}^T \boldsymbol{u}_2\right) \boldsymbol{u}_2 \right] \tag{57}$$

*Proof.* The cost function of Eq. 56 can be written as:

$$J_0 = \frac{1}{2} \left\| \boldsymbol{q} - \boldsymbol{q}_p \right\|_2^2 = \frac{1}{2} \left( \boldsymbol{q}^T \boldsymbol{q} - 2\boldsymbol{q}^T \boldsymbol{q}_p + \boldsymbol{q}_p^T \boldsymbol{q}_p \right) = 1 - \boldsymbol{q}^T \boldsymbol{q}_p. \tag{58}$$

Minimizing the cost function of Eq. 58 is the same as maximizing the following cost function:

$$J_1 = \boldsymbol{q}^T \boldsymbol{q}_p. \tag{59}$$

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 $\boldsymbol{u}_2$:

$$\boldsymbol{q}_p = a\boldsymbol{u}_1 + b\boldsymbol{u}_2. \tag{60}$$

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. 60 are constrained such that $a^2 + b^2 = 1$. We rewrite the optimization problem as:

$$\begin{cases} \max_{a,b} J_1 = \boldsymbol{q}^T \boldsymbol{q}_p = a\boldsymbol{q}^T \boldsymbol{u}_1 + b\boldsymbol{q}^T \boldsymbol{u}_2 \\ s.t. \quad a^2 + b^2 = 1. \end{cases} \tag{61}$$

Introducing the Lagrange multiplier $\lambda$, the Lagragian related to the problem above is written as:

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

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

$$\boldsymbol{q}^T \boldsymbol{u}_1 + \lambda a = 0 \quad \Longrightarrow a = -\frac{\boldsymbol{q}^T \boldsymbol{u}_1}{\lambda}, \qquad \boldsymbol{q}^T \boldsymbol{u}_2 + \lambda b = 0 \quad \Longrightarrow b = -\frac{\boldsymbol{q}^T \boldsymbol{u}_2}{\lambda}. \tag{62}$$

Substituting $a$ and $b$ from Eq. 62 into $a^2 + b^2 = 1$, we get that:

$$\frac{(\boldsymbol{q}^T \boldsymbol{u}_1)^2}{\lambda^2} + \frac{(\boldsymbol{q}^T \boldsymbol{u}_2)^2}{\lambda^2} = 1 \quad \Longrightarrow \quad \lambda = \pm\sqrt{(\boldsymbol{q}^T \boldsymbol{u}_1)^2 + (\boldsymbol{q}^T \boldsymbol{u}_2)^2}. \tag{63}$$

Therefore, we have that:

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

We can notice that this problem has two extremum points: a maximizing solution and a minimizing one. By inspecting the cost function in Eq. 61, the maximizing solution has to be the one given by Eq. 57. □

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. 55 can be written as:

$$J_2 = \sum_{i=1}^{n} \frac{1}{\sqrt{\left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1\right)^2 + \left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2\right)^2}} \bar{\boldsymbol{q}}_i^T \left[\left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1\right) \hat{\boldsymbol{u}}_1 + \left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2\right) \hat{\boldsymbol{u}}_2\right] \tag{64}$$

$$= \sum_{i=1}^{n} \frac{1}{\sqrt{\left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1\right)^2 + \left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2\right)^2}} \left[\left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1\right)^2 + \left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2\right)^2\right] = \sum_{i=1}^{n} \sqrt{\left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1\right)^2 + \left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2\right)^2} \tag{65}$$

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

$$\sqrt{\left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1\right)^2 + \left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2\right)^2} = 1, \quad \forall\, i \in \{1, \cdots, n\}. \tag{66}$$

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

$$\sqrt{x} \approx \frac{1}{2} + \frac{x}{2} \quad \Longrightarrow \quad \sqrt{\left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1\right)^2 + \left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2\right)^2} \approx \frac{1}{2} + \frac{1}{2}\left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1\right)^2 + \frac{1}{2}\left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2\right)^2 \tag{67}$$

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

$$J_2 \approx \frac{n}{2} + \frac{1}{2} \sum_{i=1}^{n} \left[\left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1\right)^2 + \left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2\right)^2\right] \tag{68}$$

For simplicity of notation, we define a new cost function whose maximization is equivalent to the maximization of Eq. 68:

$$J = \sum_{i=1}^{n} \left[\left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1\right)^2 + \left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2\right)^2\right] = \sum_{i=1}^{n} \left[\hat{\boldsymbol{u}}_1^T \bar{\boldsymbol{q}}_i \bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1 + \hat{\boldsymbol{u}}_2^T \bar{\boldsymbol{q}}_i \bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2\right] = \hat{\boldsymbol{u}}_1^T \sum_{i=1}^{n} \bar{\boldsymbol{q}}_i \bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1 + \hat{\boldsymbol{u}}_2^T \sum_{i=1}^{n} \bar{\boldsymbol{q}}_i \bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2 \tag{69}$$

$$= \hat{\boldsymbol{u}}_1^T \bar{\boldsymbol{Q}} \bar{\boldsymbol{Q}}^T \hat{\boldsymbol{u}}_1 + \hat{\boldsymbol{u}}_2^T \bar{\boldsymbol{Q}} \bar{\boldsymbol{Q}}^T \hat{\boldsymbol{u}}_2 \tag{70}$$

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

$$\begin{cases} \arg \max\limits_{\hat{\boldsymbol{u}}_1 \in \mathbb{S}^3, \hat{\boldsymbol{u}}_2 \in \mathbb{S}^3} \hat{\boldsymbol{u}}_1^T \bar{\boldsymbol{Z}} \hat{\boldsymbol{u}}_1 + \hat{\boldsymbol{u}}_2^T \bar{\boldsymbol{Z}} \hat{\boldsymbol{u}}_2 \\ \text{s.t. } \hat{\boldsymbol{u}}_1^T \hat{\boldsymbol{u}}_2 = 0 \end{cases}. \tag{71}$$

**Theorem 2.** *The optimization problem in Eq. 71 does not a have unique solution.*

*Proof.* Assume that $\hat{\boldsymbol{u}}_1^*$ and $\hat{\boldsymbol{u}}_2^*$, with $\hat{\boldsymbol{u}}_1^{*T} \hat{\boldsymbol{u}}_2^* = 0$, are maximizers of the cost function of Eq. 71. The optimal cost is given by:

$$J^* = \sum_{i=1}^{n} \left[\left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1^*\right)^2 + \left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2^*\right)^2\right] = \sum_{i=1}^{n} J_i^*, \tag{72}$$

where $J_i^* \triangleq \left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1^*\right)^2 + \left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2^*\right)^2$. Also, define the vectors $\boldsymbol{v}_1 \in \mathbb{S}^3$ and $\boldsymbol{v}_2 \in \mathbb{S}^3$ as linear combinations of $\hat{\boldsymbol{u}}_1^*$ and $\hat{\boldsymbol{u}}_2^*$:

$$\boldsymbol{v}_1 \triangleq a\hat{\boldsymbol{u}}_1^* + b\hat{\boldsymbol{u}}_2^*, \qquad \boldsymbol{v}_2 \triangleq -b\hat{\boldsymbol{u}}_1^* + a\hat{\boldsymbol{u}}_2^*, \tag{73}$$

for any $a \in [-1,1]$, $b \in [-1,1]$ such that $a^2+b^2=1$. Now we compute the quantity $J_i \triangleq \left(\bar{\boldsymbol{q}}_i^T \boldsymbol{v}_1\right)^2 + \left(\bar{\boldsymbol{q}}_i^T \boldsymbol{v}_2\right)^2$:

$$\begin{aligned} J_i &= \left(a\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1^* + b\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2^*\right)^2 + \left(-b\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1^* + a\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2^*\right)^2 \\ &= \left(a^2+b^2\right)\left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1^*\right)^2 + \left(a^2+b^2\right)\left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2^*\right)^2 = \left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1^*\right)^2 + \left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2^*\right)^2. \end{aligned} \tag{74}$$

Therefore, $J_i = J_i^*$, implying that $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are also maximizers of the cost function in Eq. 70. □

**Theorem 3.** *A solution to the optimization problem in Eq. 71 can be obtained from the Singular Value Decomposition (SVD) $\bar{\boldsymbol{Z}} = \hat{\boldsymbol{U}}\hat{\boldsymbol{\Sigma}}\hat{\boldsymbol{U}}^T$, where $\hat{\boldsymbol{U}} \in \mathbb{R}^{4\times 4} = \begin{bmatrix}\hat{\boldsymbol{u}}_1 & \hat{\boldsymbol{u}}_2 & \hat{\boldsymbol{u}}_3 & \hat{\boldsymbol{u}}_4\end{bmatrix}$ contains the singular vectors of $\bar{\boldsymbol{Z}}$, and $\hat{\boldsymbol{\Sigma}} = \mathrm{diag}\left(\hat{\sigma}_1, \hat{\sigma}_2, \hat{\sigma}_3, \hat{\sigma}_4\right)$ 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. 71 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 $\hat{\sigma}_2 = \hat{\boldsymbol{u}}_2^T \bar{\boldsymbol{Z}} \hat{\boldsymbol{u}}_2$. It is also true that $\hat{\sigma}_3 = \hat{\boldsymbol{u}}_3^T \bar{\boldsymbol{Z}} \hat{\boldsymbol{u}}_3$ and $\hat{\sigma}_4 = \hat{\boldsymbol{u}}_4^T \bar{\boldsymbol{Z}} \hat{\boldsymbol{u}}_4$.*

*Proof.* The proof follows from the known SVD property that the best-fit k-dimensional subspace for a matrix is the subspace spanned by thee first $k$ singular vectors. □

Having the optimal hyperplane estimate $\hat{\mathbb{P}}(\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2)$, we still need to calculate the AOR $\hat{\vec{\boldsymbol{\omega}}}$ that leads to rotation on that plane. As previously observed in Eq. 48, the optimal hyperplane can be written as $\hat{\mathbb{P}}(\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2) = \hat{\mathbb{P}}(\hat{\boldsymbol{u}}_1, \vec{\boldsymbol{\omega}} \otimes \hat{\boldsymbol{u}}_1)$. This implies that $\hat{\boldsymbol{u}}_2 = \hat{\vec{\boldsymbol{\omega}}} \otimes \hat{\boldsymbol{u}}_1$. Therefore, the optimal estimate for the AOR is given by:

$$\hat{\vec{\boldsymbol{\omega}}} = \hat{\boldsymbol{u}}_2 \otimes \hat{\boldsymbol{u}}_1^{-1}. \tag{75}$$

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

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:

$$\hat{\boldsymbol{q}}_i = \frac{1}{\sqrt{\left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1\right)^2 + \left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2\right)^2}} \left[\left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1\right)\hat{\boldsymbol{u}}_1 + \left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2\right)\hat{\boldsymbol{u}}_2\right]. \tag{76}$$

### Estimation of the Angular Velocity Magnitude

Using Eq. 48, and assuming that all $\hat{\boldsymbol{q}}_i$ belong to the plane $\hat{\mathbb{P}}(\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2)$, then we can write $\hat{\boldsymbol{q}}_i$ in the form:

$$\hat{\boldsymbol{q}}_i = \left[\cos\tfrac{\Omega\Delta t_i}{2} \cdot \boldsymbol{I} + \sin\tfrac{\Omega\delta_i t}{2} \cdot \hat{\vec{\boldsymbol{\omega}}}\otimes\right]\hat{\boldsymbol{q}}_1, \tag{77}$$

where $\Delta t_i = t_i - t_1$, $\forall i \in \{1, \cdots, n\}$. If we post-multiply $\hat{\boldsymbol{q}}_i$ by $\hat{\boldsymbol{q}}_1^{-1} \otimes \hat{\boldsymbol{u}}_1$, we get:

$$\underline{\hat{\boldsymbol{q}}}_i \triangleq \hat{\boldsymbol{q}}_i \otimes \hat{\boldsymbol{q}}_1^{-1} \otimes \hat{\boldsymbol{u}}_1 = \cos\frac{\Omega\Delta t_i}{2} \cdot \hat{\boldsymbol{u}}_1 + \sin\frac{\Omega\Delta t_i}{2} \cdot \hat{\vec{\boldsymbol{\omega}}} \otimes \hat{\boldsymbol{u}}_1 \tag{78}$$

$$= \cos\frac{\Omega\Delta t_i}{2} \cdot \hat{\boldsymbol{u}}_1 + \sin\frac{\Omega\Delta t_i}{2} \cdot \hat{\boldsymbol{u}}_2 \tag{79}$$

Based on Eq. 79, the following holds:

$$\underline{\hat{\boldsymbol{q}}}_i^T \hat{\boldsymbol{u}}_1 = \cos\frac{\Omega\Delta t_i}{2}, \qquad \underline{\hat{\boldsymbol{q}}}_i^T \hat{\boldsymbol{u}}_2 = \sin\frac{\Omega\Delta t_i}{2}. \tag{80}$$

Defining the angle $\Phi_i = \Omega \Delta t_i$, then estimates for $\Phi_i$ can be obtained from:

$$\hat{\Phi}_i = 2 \cdot \operatorname{atan2}\left(\hat{\underline{\boldsymbol{q}}}_i^T \hat{\boldsymbol{u}}_2,\ \ \hat{\underline{\boldsymbol{q}}}_i^T \hat{\boldsymbol{u}}_1\right). \tag{81}$$

**Theorem 4.** *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* $\theta$ *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:*

$$\boldsymbol{q}_{Np} = \begin{bmatrix}\cos\frac{\Phi}{2} \\ \boldsymbol{v}\sin\frac{\Phi}{2}\end{bmatrix}, \tag{82}$$

*where* $\Phi$ *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:

$$\hat{\boldsymbol{q}}_{Np} = \frac{1}{\sqrt{\left(\boldsymbol{q}_N^T \boldsymbol{q}_I\right)^2 + \left(\boldsymbol{q}_N^T \boldsymbol{q}_v\right)^2}} \left[\left(\boldsymbol{q}_N^T \boldsymbol{q}_I\right) \boldsymbol{q}_I + \left(\boldsymbol{q}_N^T \boldsymbol{q}_v\right) \boldsymbol{q}_v\right] = \frac{1}{\sqrt{\left(\boldsymbol{q}_N^T \boldsymbol{q}_I\right)^2 + \left(\boldsymbol{q}_N^T \boldsymbol{q}_v\right)^2}} \begin{bmatrix}\boldsymbol{q}_N^T \boldsymbol{q}_I \\ \boldsymbol{v} \cdot \boldsymbol{q}_N^T \boldsymbol{q}_v\end{bmatrix} \tag{83}$$

Comparing Eq. 83 with Eq. 82, we get that:

$$\cos\frac{\Phi}{2} = \frac{\boldsymbol{q}_N^T \boldsymbol{q}_I}{\sqrt{\left(\boldsymbol{q}_N^T \boldsymbol{q}_I\right)^2 + \left(\boldsymbol{q}_N^T \boldsymbol{q}_v\right)^2}} \tag{84}$$

From the definition of the identity quaternion (Eq. 21), 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 $\gamma$ 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. 84, and performing Taylor series expansion on both sides around $\Phi = 0$ and $\theta = 0$, we get to:

$$\cos\frac{\Phi}{2} = \frac{\cos\frac{\theta}{2}}{\sqrt{\cos^2\frac{\theta}{2} + \cos^2\gamma \sin^2\frac{\theta}{2}}} \qquad \text{(Taylor Series on both sides)} \tag{85}$$

$$1 - \frac{\Phi^2}{8} \approx 1 - \cos^2\gamma \frac{\theta^2}{8} \tag{86}$$

Inspecting Eq. 86, 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 $\mathbb{E}[\Phi^2] = \mathbb{E}[\theta^2]\mathbb{E}[\cos^2\gamma] = \frac{1}{3}\sigma_\theta^2$. □

If we assume the model:

$$\Phi_i = \Phi_0 + \Omega \Delta t_i = \begin{bmatrix}1 & \Delta t_i\end{bmatrix} \begin{bmatrix}\Phi_0 \\ \Omega\end{bmatrix}, \tag{87}$$

and using the result from Theorem 4 for the distribution of $\hat{\Phi}_i \sim \mathcal{N}\left(\Phi_i, \frac{1}{3}\sigma_\theta^2\right)$, then we can perform the least squares estimation:

$$\hat{\boldsymbol{X}} \triangleq \begin{bmatrix}\hat{\Phi}_0 \\ \hat{\Omega}\end{bmatrix} = \left(\boldsymbol{H}^T \boldsymbol{P}_\Phi^{-1} \boldsymbol{H}\right)^{-1} \boldsymbol{H}^T \boldsymbol{P}_\Phi^{-1} \hat{\boldsymbol{\Phi}}, \tag{88}$$

where:

$$\boldsymbol{H} \triangleq \begin{bmatrix} 1 & \cdots & 1 \\ \Delta t_1 & \cdots & \Delta t_n \end{bmatrix}^T, \qquad \hat{\boldsymbol{\Phi}} \triangleq \begin{bmatrix} \hat{\Phi}_1 & \dots & \hat{\Phi}_n \end{bmatrix}^T, \qquad \boldsymbol{P}_\Phi \triangleq \text{diag}\left(\tfrac{1}{3}\sigma_\theta^2, \cdots, \tfrac{1}{3}\sigma_\theta^2\right). \tag{89}$$

The covariance matrix of the estimate $\hat{\boldsymbol{X}}$ is given by:

$$\text{cov}[\hat{\boldsymbol{X}}] = \left(\boldsymbol{H}^T \boldsymbol{P}_\Phi^{-1} \boldsymbol{H}\right)^{-1} = \frac{1}{3}\sigma_\theta^2 \left(\boldsymbol{H}^T \boldsymbol{H}\right)^{-1}. \tag{90}$$

**Algorithm Summary**

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

1. Construct the measurement matrix $\bar{\boldsymbol{Q}}$ as in Eq. 49 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. 75: $\hat{\boldsymbol{\omega}} = \hat{\boldsymbol{u}}_2 \otimes \hat{\boldsymbol{u}}_1^{-1}$.
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. 76.
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 $\hat{\Phi}_i$ using Eqs. 78-81.
6. Estimate the angular velocity $\hat{\Omega}$ and its associated covariance using Eqs. 88 and 89.

**QuateRA Analysis**

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

First of all, our algorithm pursues the minimization of the TLS cost function (Eq. 51). In order to use such cost function, we assumed that it wasn't necessary to normalize the measured quaternions into a form whose covariance is the identity matrix. Our assumption was based on the fact that the normalization would just be approximately a vector scaling. This approximation could have dangerous implications, since classical TLS formulation assumes measurements whose covariance are of the type $\sigma^2\boldsymbol{I}$, with some $\sigma \in \mathbb{R}_{>0}$ [29]. This means that our assumption could jeopardize the final results, including the *strong consistency* (asymptotic unbiasedness) property related to TLS problems. However, as we will show in our Monte Carlo analysis, QuateRA is close to unbiased for the noise values under consideration.

An assumption made in QuateRA is that $x \triangleq \left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_1\right)^2 + \left(\bar{\boldsymbol{q}}_i^T \hat{\boldsymbol{u}}_2\right)^2 \approx 1$, which implies a small angle approximation for the quaternion measurement noise of Eq. 29. This approximation is used to reach the final optimization problem (Eq. 71), which has a closed form solution through SVD (no need for any iterative nonlinear programming). The downside of this approximation is that we cannot guarantee that the SVD solution is also the minimizer for the original TLS cost function Eq. 51 under high measurement noise. It is, however, still the maximizer for the cost function of Eq. 68.

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. 69 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.

## QUATERA MONTE CARLO ANALYSIS

This section provides a Monte Carlo analysis of QuateRA, endorsing the statistical properties derived in the previous sections. We perform extensive simulations for multiple values of $n$ (number of measurements) and $\sigma_\theta$ (standard deviation for the angle in the noise quaternion).

In all simulations, we used an angular velocity with magnitude $\Omega = 0.1$rad/s, direction $\vec{\omega} = \frac{1}{\sqrt{3}}\begin{bmatrix}1 & 1 & 1\end{bmatrix}^T$, and measurements are taken once every $0.1$ seconds. 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 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 $\vec{\omega}_\perp \in \mathbb{S}^2$ as an arbitrary unit vector perpendicular to $\vec{\omega}$, i.e., $\vec{\omega}^T\vec{\omega}_\perp = 0$.

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

$$\mu_\perp \triangleq \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} e_{i\perp}, \qquad \sigma_\perp^2 \triangleq \frac{1}{n_{MC}-1} \sum_{i=1}^{n_{MC}} \left(e_{i\perp} - \mu_\perp\right)^2. \tag{91}$$

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 $\sigma_\perp \leq 0.1$. Figure 1 presents the Monte Carlo results for the AOR estimation, indicating that the estimator is asymptotically unbiased (as usual in TLS problems) and that the standard deviations satisfy $\sigma_\perp \leq 0.1$ for a sufficiently large set of measurements. For instance, one would need approximately $n = 12$ measurements if $\sigma_\theta = 1°$, $n = 20$ for $\sigma_\theta = 2°$, $n = 25$ for $\sigma_\theta = 3°$, $n = 30$ for $\sigma_\theta = 4°$, and $n = 35$ for $\sigma_\theta = 5°$.

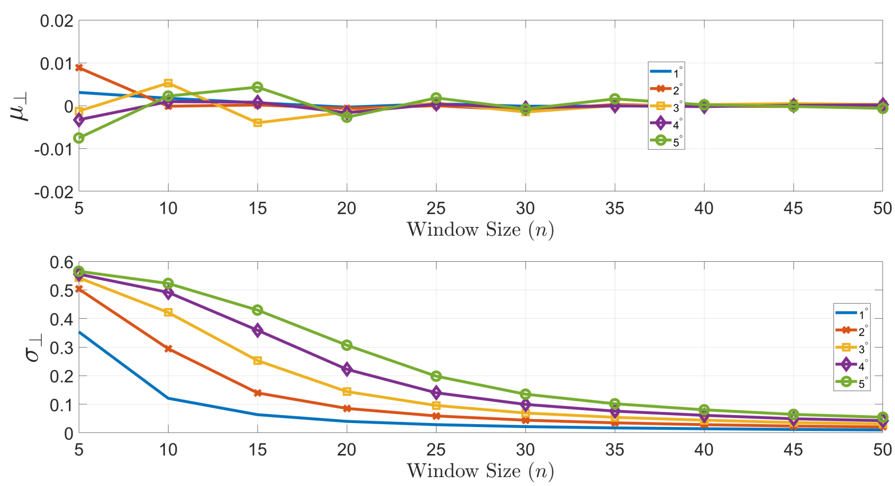


**Figure 1. Sample Mean and Standard Deviation of the projection of the estimated AOR along a direction perpendicular to the true AOR. Results are shown as a function of the number of measurements (x axis) and the standard deviations $\sigma_\theta$ (different plots).**

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

[†]Note that the values chosen for $\sigma_\theta$ are extremely large when compared to Star Tracker technology. The analysis of this section would be quite uninteresting for $\sigma_\theta$ values expected for Star Trackers, since QuateRA's performance would not change much as a function of the number of measurements $n$.

estimated AVM for the $i^{th}$ Monte Carlo execution. We calculate the mean $\mu_\Omega$ and variance $\sigma^2_\Omega$ of $e_{i\Omega}$ as:

$$\mu_\Omega \triangleq \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} e_{i\Omega}, \qquad \sigma^2_\Omega \triangleq \frac{1}{n_{MC}-1} \sum_{i=1}^{n_{MC}} \left(e_{i\Omega} - \mu_\Omega\right)^2. \tag{92}$$

Figure 2 shows that the mean error $\mu_\Omega$ 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 (we determine heuristically that $\sigma_\perp \leq 0.1$ is a reasonable rule of thumb).

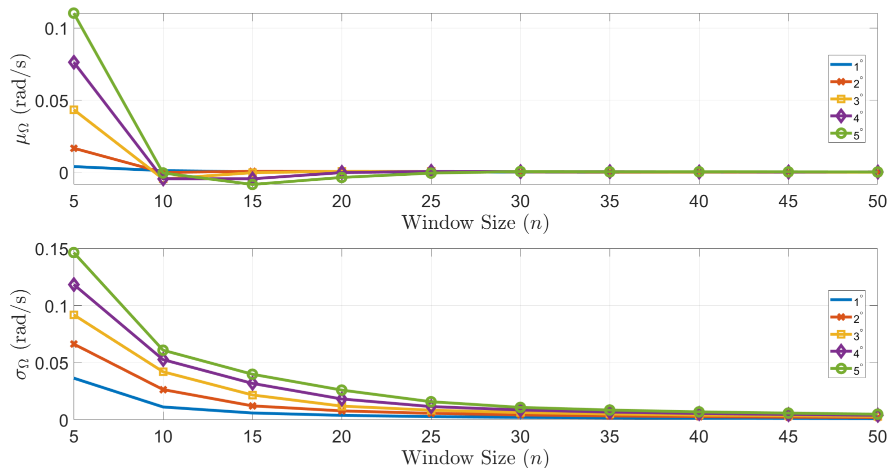


**Figure 2. Sample Mean and Standard Deviation of the estimated AVM error. Results are shown as a function of the number of measurements (x axis) and the standard deviations $\sigma_\theta$ (different plots).**

We denote the average estimated covariance for $\hat{\Omega}_i$ as $\hat{\sigma}^2_\Omega$, where the covariance is estimated using Eq. 90. Figure 3 shows the average percentage error $PE_\Omega$ of the estimated standard deviation $\hat{\sigma}_\Omega$ w.r.t. the sample standard deviation $\sigma_\Omega$: $PE_\Omega \triangleq \frac{\hat{\sigma}_\Omega - \sigma_\Omega}{\sigma_\Omega}$. Figure 3 indicates that the estimated standard deviation is consistent with the sample standard deviation provided that the number of measurements $n$ is large enough. If we desire $|PE_\Omega| < 3\%$, Figure 3 shows that one needs $n > 10$ if $\sigma_\theta = 1°$, $n > 15$ if $\sigma_\theta = 2°$, $n > 20$ if $\sigma_\theta = 3°$, $n > 25$ if $\sigma_\theta = 4°$, and $n > 35$ if $\sigma_\theta = 5°$.

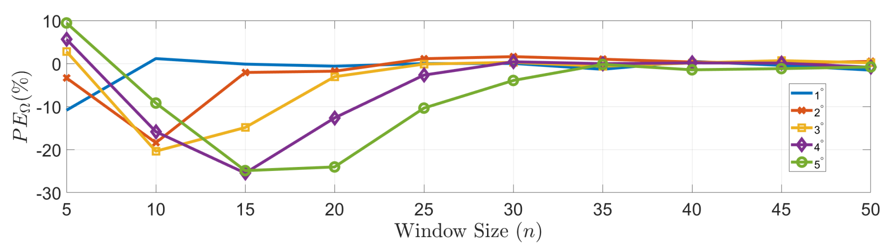


**Figure 3. Percentual error of the average estimated standard deviation for $\hat{\Omega}_i$ w.r.t. the sample standard deviation for $\hat{\Omega}_i$. Results are shown as a function of the number of measurements (x axis) and the standard deviations $\sigma_\theta$ (different plots).**

## 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.

We motivate our solution by contrasting the Least Squares solution with the Total Least Squares one for the trivial problem of estimating translational velocity from cartesian position measurements. The Least Squares (LS) solution is the best estimator for this case (in the Minimum Mean Squared Error sense), but the Total Least Squares (TLS) performs asymptotically similar to LS as the number of measurements increase. We were able to propose a TLS-based estimator for determining constant angular velocity from $n$ orientation measurements, whose performance is acceptable for large sample sets. Since there is no LS solution to the same problem (to the best of our knowledge), we can't verify that the TLS solution performs as well as the LS one as $n$ increases.

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 AVM estimation for sufficiently large sample sets. For a fixed $\sigma_\theta$, one can determine the *meaning* of sufficiently large by stating desired values for $\sigma_\perp$ (the standard deviation of the estimated axis projected along a direction perpendicular to the true angular velocity direction), $\sigma_\Omega$ (standard deviation of the estimated angular velocity magnitude), and expected bounds for $|PE_\Omega|$.

Our earlier contributions have already demonstrated the application of preliminary versions of QuateRA for estimating a non-constant angular velocity. These works made use of adaptive sliding windows to estimate the time-varying Axis of Rotation (AOR), while the Angular Velocity Magnitude (AVM) was either estimated through a lowpass-filtered *dirty derivative* or through a modified MEKF. These works introduced tuning parameters for adapting the size of the sliding window and for tuning the AVM estimator. In contrast, the current work presents a method for estimating the AVM that is free of tuning parameters, and it does produce a covariance estimate for the AVM (provided a sufficiently large sample set). These contributions are relevant for the overall problem of estimating a time-varying AOR without the need for heuristic tuning.

We should highlight that a weakness within QuateRA is that we were not able to determine the covariance 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 Theorem 2, the TLS solution for this problem is not 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. Deriving the AOR covariance would be a meaningful contribution for future work. Alternatively, practitioners can estimate the AOR covariance through Monte Carlo simulations as follows: given a sensor with measurement noise covariance $\sigma_\theta^2$, the covariance of the AOR can be recorded into a lookup table as a function of the number of measurements $n$ (similar to what was done in Figure 1).

Another interesting path for future work would be to expand QuateRA for different measurement weights. Also, we have assumed that the axis of the noise quaternion is distributed in a uniform spherical distribution, whereas this is not always true. 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.

## APPENDICES

### APPENDIX 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: $\mathbb{E}[e] = \mathbf{0}$ 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 [9], if we slice both the cylinder and the sphere at the same height as shown on Fig. 4, then the lateral surface area of the spherical segment ($S_1$) is equal to the lateral surface area of the cylindrical segment ($S_2$).

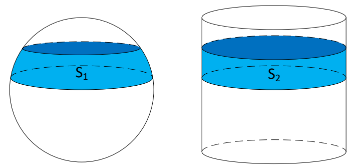


**Figure 4. Illustration of Archimedes' Hat-Box Theorem.**

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 [30] 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:

$$\boldsymbol{e} = \begin{bmatrix}\sqrt{1-z^2}\cos(\phi) & \sqrt{1-z^2}\sin(\phi) & z\end{bmatrix}^T. \tag{93}$$

The transformation of Eq. 93 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 $\phi$ respectively, then:

$$\mathbb{E}[z] = \int_{-1}^{1} x P_z(x)\,dx = \frac{1}{2}\int_{-1}^{1} x\,dx = \frac{1}{4}x^2\Big|_{-1}^{1} = 0, \tag{94}$$

$$\mathbb{E}[z^2] = \int_{-1}^{1} x^2 P_z(x)\,dx = \frac{1}{2}\int_{-1}^{1} x^2\,dx = \frac{1}{6}x^3\Big|_{-1}^{1} = \frac{1}{3}, \tag{95}$$

$$\mathbb{E}[1-z^2] = 1 - \frac{1}{3} = \frac{2}{3}, \qquad \mathbb{E}[\cos\phi] = 0, \qquad \mathbb{E}[\sin\phi] = 0, \qquad \mathbb{E}[\cos\phi\sin\phi] = 0 \tag{96}$$

$$\mathbb{E}[\cos^2\phi] = \int_{-\pi}^{\pi} \cos^2 x P_\phi(x)\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi}\cos^2 x\,dx = \frac{1}{8\pi}(2x + \sin 2x)\Big|_{-\pi}^{\pi} = \frac{1}{2} \tag{97}$$

$$\mathbb{E}[\sin^2\phi] = \int_{-\pi}^{\pi} \sin^2 x P_\phi(x)\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sin^2 x\,dx = \frac{1}{8\pi}(2x - \sin 2x)\Big|_{-\pi}^{\pi} = \frac{1}{2} \tag{98}$$

Therefore, given that $z$ and $\phi$ are independently distributed, we have that:

$$\mathbb{E}[\boldsymbol{e}] = \begin{bmatrix}\mathbb{E}[\sqrt{1-z^2}\cos(\phi)]\\ \mathbb{E}[\sqrt{1-z^2}\sin(\phi)]\\ \mathbb{E}[z]\end{bmatrix} = \begin{bmatrix}\mathbb{E}[\sqrt{1-z^2}]\mathbb{E}[\cos(\phi)]\\ \mathbb{E}[\sqrt{1-z^2}]\mathbb{E}[\sin(\phi)]\\ \mathbb{E}[z]\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}, \tag{99}$$

$$\mathbb{E}[\boldsymbol{e}\boldsymbol{e}^T] = \mathbb{E}\left[\begin{bmatrix}(1-z^2)\cos^2\phi & (1-z^2)\cos\phi\sin\phi & (1-z^2)z\cos\phi\\ (1-z^2)\cos\phi\sin\phi & (1-z^2)\sin^2\phi & (1-z^2)z\sin\phi\\ (1-z^2)z\cos\phi & (1-z^2)z\sin\phi & z^2\end{bmatrix}\right] = \frac{1}{3}\boldsymbol{I}. \tag{100}$$

## REFERENCES


[1] Robert James Adcock. A problem in least squares. *The Analyst*, 5(1):53–54, 1878.

[2] Marcelino Almeida and Maruthi Akella. Discrete adaptive angular velocity estimation - an experimental analysis. In *2016 Space Flight Mechanics Meeting*, 2016.

[3] Marcelino Almeida, Renato Zanetti, Daniele Mortari, and Maruthi Akella. Real-time angular velocity estimation of non-cooperative space objects using camera measurements. In *2018 AAS/AIAA Astrodynamics Specialist Conference*, 2018.

[4] Itzhack Y Bar-Itzhack. Classification of algorithms for angular velocity estimation. *Journal of Guidance, Control, and Dynamics*, 24(2):214–218, 2001.

[5] Itzhack Y Bar-Itzhack, Richard R Harman, and Julie K Thienel. Rigid body rate inference from attitude variation. *Journal of guidance, control, and dynamics*, 30(1):275–281, 2007.

[6] Yaakov Bar-Shalom, X Rong Li, and Thiagalingam Kirubarajan. *Estimation with applications to tracking and navigation: theory algorithms and software*. John Wiley & Sons, 2004.

[7] James R Carr. Orthogonal regression: a teaching perspective. *International Journal of Mathematical Education in Science and Technology*, 43(1):134–143, 2012.

[8] John L Crassidis and Yang Cheng. Error-covariance analysis of the total least-squares problem. *Journal of Guidance, Control, and Dynamics*, 37(4):1053–1063, 2014.

[9] H Cundy and A Rollett. Sphere and cylinder—archimedes' theorem. *Mathematical Models*, pages 172–173, 1989.

[10] Eliezer Gai, Kevin Daly, James Harrison, and Linda Lemos. Star-sensor-based satellite attitude/attitude rate estimator. *Journal of Guidance, Control, and Dynamics*, 8(5):560–565, 1985.

[11] Joanna C Hinks and Mark L Psiaki. Solution strategies for an extension of wahba's problem to a spinning spacecraft. *Journal of Guidance, Control, and Dynamics*, 34(6):1734–1745, 2011.

[12] Sabine Van Huffel and Joos Vandewalle. 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*, 10(3):294–315, 1989.

[13] Ronald H Ketellapper. The relevance of large sample properties of estimators for the errors-in-variables model: A monte carlo study. *Communications in Statistics-Simulation and Computation*, 11(5):625–634, 1982.

[14] Ern J Lefferts, F Landis Markley, and Malcolm D Shuster. Kalman filtering for spacecraft attitude estimation. *Journal of Guidance, Control, and Dynamics*, 5(5):417–429, 1982.

[15] Gerald M Lerner. Three-axis attitude determination. *Spacecraft Attitude Determination and Control*, 73:420–428, 1978.

[16] F Landis Markley. Attitude error representations for kalman filtering. *Journal of guidance, control, and dynamics*, 26(2):311–317, 2003.

[17] F Landis Markley, Yang Cheng, John Lucas Crassidis, and Yaakov Oshman. Averaging quaternions. *Journal of Guidance, Control, and Dynamics*, 30(4):1193–1197, 2007.

[18] F Landis Markley and Daniele Mortari. Quaternion attitude estimation using vector observations. *Journal of the Astronautical Sciences*, 48(2):359–380, 2000.

[19] Brian J. McCartin. A geometric characterization of linear regression. *Statistics: A Journal of Theoretical and Applied Statistics*, 37(2):101–117, 2003.

[20] D. Mortari and M. Akella. Discrete and continuous time adaptive angular velocity estimators. In *Proceedings of the AAS/AIAA Space Flight Mechanics Conference, Williamsburg, VA*, 2015.

[21] Daniele Mortari and Manoranjan Majji. Multiplicative measurement model. *The Journal of the Astronautical Sciences*, 57(1-2):47–60, 2009.

[22] Yaakov Oshman, Franc-para, and ois Dellus. Spacecraft angular velocity estimation using sequential observations of a single directional vector. *Journal of Spacecraft and Rockets*, 40(2):237–247, 2003.

[23] Mark L Psiaki. Backward-smoothing extended kalman filter. *Journal of guidance, control, and dynamics*, 28(5):885–894, 2005.

[24] Mark L Psiaki. Generalized wahba problems for spinning spacecraft attitude and rate determination. *The Journal of the Astronautical Sciences*, 57(1-2):73–92, 2009.

[25] Mark L Psiaki and Joanna C Hinks. Numerical solution of a generalized wahba problem for a spinning spacecraft. *Journal of Guidance, Control, and Dynamics*, 35(3):764–773, 2012.

[26] S Salcudean. A globally convergent angular velocity observer for rigid body motion. *IEEE transactions on Automatic Control*, 36(12):1493–1497, 1991.

[27] James Saunderson, Pablo A Parrilo, and Alan S Willsky. Convex solution to a joint attitude and spin-rate estimation problem. *Journal of Guidance, Control, and Dynamics*, 39(1):118–127, 2015.

[28] Alan Shuchat. Generalized least squares and eigenvalues. *The American Mathematical Monthly*, 92(9):656–659, 1985.

[29] Sabine Van Huffel and Joos Vandewalle. *The total least squares problem: computational aspects and analysis*, volume 9. Siam, 1991.

[30] Eric W Weisstein. Sphere Point Picking. from mathworld–a wolfram web resource. `http://mathworld.wolfram.com/SpherePointPicking.html`, 2018. Accessed: 2018-08-08.