A Semidefinite Relaxation-based Algorithm for Robust Attitude Estimation

Shakil Ahmed, Eric C. Kerrigan and Imad M. Jaimoukha

Abstract

This paper presents a tractable method for solving a robust attitude estimation problem, based on a weighted least squares approach with nonlinear constraints. Attitude estimation requires information of a few vector quantities, each obtained from both a sensor and a mathematical model. By considering the modeling errors, measurement noise, sensor biases and offsets as infinity-norm bounded uncertainties, we formulate a robust optimization problem, which is non-convex with nonlinear cost and constraints. The robust min-max problem is approximated with a non-convex minimization problem using an upper bound. A new regularization scheme is also proposed to improve the robust performance. We then use semidefinite relaxation to convert the suboptimal problem with quadratic cost and constraints into a tractable semidefinite program with a linear objective function and linear matrix inequality constraints. We also show how to extract the solution of the suboptimal robust estimation problem from the solution of the semidefinite relaxation. Further, a mathematical proof supported by numerical results is presented stating the gap between the suboptimal problem and its relaxation is zero under a given condition, which is mostly true in real life scenarios. The usefulness of the proposed algorithm in the presence of uncertainties is evaluated with the help of examples.

Index Terms

Estimation, uncertainty, robustness, optimization methods, min-max techniques, relaxation methods

EDICS Category: SSP-PARE

S. Ahmed and Imad M. Jaimoukha are with the Department of Electrical and Electronic Engineering, Imperial College London, UK (e-mail: shakil.ahmed08, i.jaimouka@imperial.ac.uk).

Eric C. Kerrigan is with the Department of Electrical and Electronic Engineering and Department of Aeronautics, Imperial College London, UK (e-mail: e.kerrigan@imperial.ac.uk).

A Semidefinite Relaxation-based Algorithm for Robust Attitude Estimation

I. Introduction

ATTITUDE estimation using vector signals has been widely used in many application areas, such as satellites [1], [2], aerospace, marine and automotive systems [3]. The attitude estimate is obtained by solving an optimization problem based on a weighted least squares approach with nonlinear constraints, known as the Wahba problem [4] in the literature. This mathematical formulation is also closely related to some other problems in various fields, such as independent component analysis (ICA) in signal processing and statistics, pose estimation in image processing [5] and the orthogonal matrix Procrustes problem in mathematics [6]. This type of attitude estimation can be called static, as it does not dependent on the dynamics of the system, hence it could be useful for systems with highly nonlinear dynamics. In such systems, due to high nonlinearities, the dynamic filtering approaches suffer from divergence issues due to lack of good a priori state estimates [2]. The attitude determined using the static approach could be used to obtain a reliable state initialization for filters, thus reducing the likelihood of divergence.

To compute the attitude of an object, two coordinate frames are needed. One, which is fixed to the body of the object, is called the body frame, while the second is called the reference frame. Formally, the attitude of an object is defined as a coordinate transformation that transforms reference coordinates into the body coordinates [7]. This transformation is obtained through a proper orthogonal transformation matrix $C \in \mathbb{R}^{3\times3}$, having the constraint $C^TC = I_3$ for orthogonality, and $\det(C) = +1$ to preserve the frame orientation in a rotation. This includes the set of all rotation matrices in a *special orthogonal* group of rigid rotations in $\mathbb{R}^3$, denoted by $SO(3)$ [5]. Many solutions of this constrained least squares problem can be found in the literature, mostly developed for satellite applications [1], [8]–[11]. Most of these algorithms are based on a quaternion transformation [12], which transforms the Wahba problem into an eigenvalue problem [1].

Some examples of the vector signals normally used in static attitude determination includes the earth magnetic field, sun and star direction, position vector, etc. Information of these vectors is required both in the body and the reference frames in order to determine the attitude. Normally the body frame vectors are measured by some sensor installed on the object, while the same vector information in the reference frame is obtained from some mathematical model. For both sensor measurements or mathematical models, an error is always present. This error is mainly due to noise, biases, offsets and modeling inaccuracies. The existing algorithms do not directly address robustness of the estimated attitude against uncertainties, although a sensitivity analysis is generally presented. One can find a lot of work on robust linear least square problems in literature, such as [13]–[15], however there is much less discussion on robustness in the attitude estimation problem. Some algorithms, such as [16], [17], consider uncertainties in the input measurement, but use a stochastic framework based on minimum variance recursive estimation. Similarly, some related discussions can be found in [18], [19]. However, in these discussions, modeling errors are generally not considered, which could be significant; for example, in the case of the earth magnetic field, which is one of the most common sensors used for attitude estimation in many applications such as satellites and aircrafts, errors between sophisticated models and the actual field can be around 20% [20], [21]. The use of simple models, such as the low order IGRF model [21], which are normally preferred due to lower computational cost, result in a less accurate earth magnetic field, leading to errors in the attitude estimate. Attitude error is further increased due to sensor noise and installation issues. In this work, all such errors are considered as ∞-norm bounded uncertainties.

In this paper, which is mainly based on our previous work on robust static attitude determination [22], [23], the main contribution is to formulate a robust attitude estimation problem considering norm bounded uncertainties. The formulated robust optimization problem is approximated by a minimization problem using an upper bound on the maximization term of the original min-max problem. The approximate formulation is non-convex with a quadratic objective function and constraints (a QCQP). Further, we introduce a new regularization term to improve the robust performance. We propose a tractable method for solving this non-convex QCQP using semidefinite relaxation. The relaxed formulation is convex with a linear objective and linear matrix inequality constraints, which can be solved efficiently in polynomial time [24] using any semidefinite program (SDP) solver. It is also shown how to extract the robust attitude from the SDR solution. Further, we study the optimality properties of the SDR solution and theoretically show that there is no gap between the approximate problem and its semidefinite relaxation under a given condition.

Notation: Vectors are represented by small and matrices by capital letters. For a vector x, its 2-norm is $\|x\|_2 := \sqrt{x^T x}$, while the infinity-norm is $\|x\|_\infty := \max_{1\le i\le n} |x_i|$. The cross product of vectors x and y is represented as $x \times y$. For matrices A and B of the same dimension, $C = A \times B$ is the columnwise cross product i.e. if a is the i^{th} column of A and b is the i^{th} column of B, then $a \times b$ will be the i^{th} column of C. We will also use the symbol $\times$ for the Cartesian product. For a matrix, $A \succeq 0$ means that A is positive semidefinite. I_n denotes the identity matrix of size n, while $0_{n\times m}$ represents a matrix of n rows and m columns with all zero entries. Operator $\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ represents a matrix of size $n \times n$, having only diagonal elements $\lambda_1, \lambda_2, \ldots, \lambda_n$. Operator $\mathrm{tr}(A)$ is the trace and $\det(A)$ is the determinant of a matrix A, $\mathscr{N}(A)$ is the null space of A and $\dim(\cdot)$ represents dimension.

II. Classical Attitude Estimation

The classical static attitude estimation is based on minimizing a weighted least square cost, first proposed by [4] for satellite applications, given as:

$$\begin{aligned} \min_{C} \quad & \frac{1}{2}\sum_{i=1}^{n} w_i \left\| b_i - C r_i \right\|_2^2 \\ \text{subject to} \quad & C^T C = I_3,\ \det(C) = +1, \end{aligned} \tag{1}$$

where $b_i \in \mathbb{R}^3$ represents the i^{th} measurement in the body frame for $i = 1,\ldots,n$, n being the total number of sensors, $r_i \in \mathbb{R}^3$ is the corresponding vector in the reference frame obtained from some model, $w_i \in \mathbb{R}$ are non-negative weights. One common approach used to solve (1) is to convert it into an equivalent maximization problem. Let $B := \begin{bmatrix} b_1 & b_2 & \ldots & b_n \end{bmatrix}$, $R := \begin{bmatrix} r_1 & r_2 & \ldots & r_n \end{bmatrix}$, $W := \text{diag}(w_1, w_2, \ldots, w_n)$, where $B, R \in \mathbb{R}^{3\times n}$ and $W \in \mathbb{R}^{n\times n}$. Using this compact notation and expanding the cost function used in (1), we get

$$\begin{aligned} \frac{1}{2}\sum_{i=1}^{n} w_i \left\| b_i - C r_i \right\|_2^2 = \quad & \frac{1}{2}\text{tr}(WB^T B + WR^T C^T C R) - \text{tr}(WB^T C R) \\ = \quad & \frac{1}{2}\text{tr}(WB^T B + WR^T R) - \text{tr}(WB^T C R). \end{aligned}$$

Using the constraint $C^T C = I_3$ and neglecting the constant term, which has no effect on the solution of the optimization problem, an equivalent maximization problem is

$$\begin{aligned} \max_{C} \quad & \text{tr}(WB^T C R) \\ \text{subject to} \quad & C^T C = I_3,\ \det(C) = +1. \end{aligned} \tag{2}$$

To solve this maximization problem, Davenport's q-method [1], [12] is commonly used, which transforms the optimization variable from matrix C to quaternion $q := \begin{bmatrix} \mathbf{q}^{\mathbf{T}} & q_4 \end{bmatrix}^T \in \mathbb{R}^4$, thus reducing the number of optimization variables. It also avoids the constraint $\det(C) = +1$ of (1), being inherent in its definition. However, the main benefit is the transformation of the optimization problem into an eigenvalue problem. Two steps of the q-method are given now.

Step 1: Find an equivalent formulation of (2) in terms of a quaternion. This new formulation, first reported in [12], states that the maximization of (2) is equivalent to the following problem (see Appendix A for derivation):

$$\max_{q} \quad \left\{ q^T \mathbf{K}(B,R) q \mid q^T q = 1 \right\}, \tag{3}$$

where $\mathbf{K} : \mathbb{R}^{3\times n} \times \mathbb{R}^{3\times n} \to \mathbb{R}^{4\times 4}$ is defined as

$$\mathbf{K}(B,R) := \begin{bmatrix} (\mathbf{B}(B,R))^T + \mathbf{B}(B,R) - \text{tr}(\mathbf{B}(B,R)) I_3 & \mathbf{z}(B,R) \\ (\mathbf{z}(B,R))^T & \text{tr}(\mathbf{B}(B,R)) \end{bmatrix}, \tag{4}$$

where $\mathbf{B}:\mathbb{R}^{3\times n}\times\mathbb{R}^{3\times n}\rightarrow\mathbb{R}^{3\times 3}$ and $\mathbf{z}:\mathbb{R}^{3\times n}\times\mathbb{R}^{3\times n}\rightarrow\mathbb{R}^{3}$ are defined as $\mathbf{B}(B,R):=BWR^T$ and $\mathbf{z}(B,R):=(B\times R)W$. Moreover, $\mathbf{K}(B,R)$ is a symmetric and indefinite matrix, indicating that the objective function in (3) is neither concave nor convex.

Step 2: The maximization problem (3) can easily be converted into an eigenvalue problem. For this we add the constraint $q^Tq=1$ using a Lagrange multiplier λ in (3) as

$$f(q,\lambda)=q^T\mathbf{K}(B,R)q-\lambda(q^Tq-1). \tag{5}$$

To obtain a stationary point, we solve $\partial f/\partial q=0$ and $\partial f/\partial\lambda=0$ and obtain an expression that has the same form as the eigenvalue problem i.e. $\mathbf{K}(B,R)q=\lambda q$, where λ represents eigenvalues of $\mathbf{K}(B,R)$. Four eigenvectors of $\mathbf{K}(B,R)$ are possible solutions of this equation; however, the eigenvector corresponding to the maximum eigenvalue will solve (3) [1], i.e. $\mathbf{K}(B,R)q_{\text{opt}}=\lambda_{\max}q_{\text{opt}}$, where q_{opt} is the solution to (3) and $\lambda_{\max}$ is the maximum eigenvalue of $\mathbf{K}(B,R)$. Most of the work on static attitude estimation is based on this result and many efficient algorithms have been proposed, such as QUEST [1], ESOQ1 [9], ESOQ2 [10], mainly for satellite applications.

III. Robust Problem Description

To formulate a robust attitude estimation problem, we represent an uncertain measurement vector in the body frame with $\bar{b}_i\in\mathsf{B}(b_i)$ and an uncertain reference vector with $\bar{r}_i\in\mathsf{R}(r_i)$, $i=1,\dots,n$, where $\mathsf{B}(b_i)$, $\mathsf{R}(r_i)\subseteq\mathbb{R}^3$ are bounded uncertainty sets. To find the best uncertainty immunized transformation matrix for attitude, we define a robust problem as

$$\begin{aligned}&\min_{C}\ \max_{\substack{\bar{b}_i\in\mathsf{B}(b_i),\bar{r}_i\in\mathsf{R}(r_i),\\ i=1,\dots,n}} && \frac{1}{2}\sum_{i=1}^{n}w_i\left\|\bar{b}_i-C\bar{r}_i\right\|_2^2\\ &\text{subject to} && C^TC=I_3,\ \det(C)=+1.\end{aligned} \tag{6}$$

In order to take advantage of using a quaternion to simplify the optimization problem, as a first step, we reformulate (6) introducing the quaternion q using the same approach used to derive (3) [22]. Let $\bar{B}:=\begin{bmatrix}\bar{b}_1 & \bar{b}_2 & \dots & \bar{b}_n\end{bmatrix}$ and $\bar{R}:=\begin{bmatrix}\bar{r}_1 & \bar{r}_2 & \dots & \bar{r}_n\end{bmatrix}$. Using this stacked notation, we can write the cost function of (6) in terms of q as

$$J(q,\bar{B},\bar{R}):=\left\{\frac{1}{2}\operatorname{tr}(W\bar{B}^T\bar{B}+W\bar{R}^T\bar{R})-q^T\mathbf{K}(\bar{B},\bar{R})q\right\}. \tag{7}$$

The robust attitude determination problem is then defined as

$$\begin{aligned}&\hat{q}^*(B,R):=\arg\min_{q}\ \max_{\bar{B}\in\mathscr{B}(B),\bar{R}\in\mathscr{R}(R)} && J(q,\bar{B},\bar{R})\\ &\text{subject to} && q^Tq=1,\end{aligned} \tag{8}$$

where $\mathscr{B}(B) := \mathsf{B}(b_1) \times \mathsf{B}(b_2) \times \cdots \times \mathsf{B}(b_n)$ and $\mathscr{R}(R) := \mathsf{R}(r_1) \times \mathsf{R}(r_2) \times \cdots \times \mathsf{R}(r_n)$. Like the matrix $\mathbf{K}(B,R)$ in (3), the matrix $\mathbf{K}(\bar{B},\bar{R})$ is also symmetric and indefinite.

A. Uncertainty Model

Uncertainties in the input vectors are of a diverse nature. These vectors are obtained from sensors and mathematical models. Sensor errors are generally attributed to measurement noise, having a stochastic interpretation, and biases and misalignments, which are fixed values. Modeling inaccuracies have generally no clear interpretation. An uncertainty model, which can fully capture all these uncertainties will be fairly complex and can make the problem intractable. Keeping in view the tractability, we consider the following affine parameterization of the uncertainty sets $\mathsf{B}(b)$ and $\mathsf{R}(r)$ [25].

Let $\beta,\rho \in \mathbb{R}^3$ be vectors of perturbation variables for the uncertainty parameterization and $\gamma_b, \gamma_r \in \mathbb{R}$ be uncertainty bounds for each vector in the body and reference frame, respectively. This type of uncertainty is called an interval uncertainty and the corresponding perturbation set represents a box [25]. The interval uncertainty model is a suitable candidate for such a type of mixed uncertainty and can sufficiently capture most realistic errors. This model is especially useful for vector quantities with bounded uncertainties. To elaborate this point, assume that in the vector quantities in $\mathbb{R}^3$, all mentioned uncertainties will introduce an error in the true value. If the maximum error introduced in each axis be bounded by $\pm\gamma$, then we can say that the true value will lie in an interval of size 2γ around the measurement. This interval in each axis will form a box in $\mathbb{R}^3$ with each side of length 2γ. The size of this interval i.e. the bound γ for each measurement or model vector, should be chosen carefully, as unnecessarily large values may result in a large residual. The choice of bounds depends on the specific sensor or mathematical model used. Generally, sensor noise is known in a stochastic sense, e.g. standard deviation or variance, while modelling errors are given based on previous experimentation or analysis. However, biases and offsets need to be separately determined for each installed sensor . Overall, the chosen bound should sufficiently capture all these errors. Further, we normalize each perturbation vector in the body and reference frame with the corresponding uncertainty bound and denote it as $\delta_b := \beta/\gamma_b$ and $\delta_r := \rho/\gamma_r$. Using these normalized perturbation vectors, we describe the uncertainty sets in the body and reference frame as

$$
\begin{aligned}
\mathsf{B}(b) &= \left\{ b + \sum_{l=1}^{3} \delta_{bl}\tilde{b}_l \mid \|\delta_b\|_\infty \leq 1 \right\}, \\
\mathsf{R}(r) &= \left\{ r + \sum_{l=1}^{3} \delta_{rl}\tilde{r}_l \mid \|\delta_r\|_\infty \leq 1 \right\},
\end{aligned} \tag{9}
$$

where $\delta_b := \begin{bmatrix} \delta_{b1} & \delta_{b2} & \delta_{b3} \end{bmatrix}^T$, $\delta_r := \begin{bmatrix} \delta_{r1} & \delta_{r2} & \delta_{r3} \end{bmatrix}^T$, $\tilde{b}_l := \gamma_b e_l$ and $\tilde{r}_l := \gamma_r e_l$ are fixed vectors for a given problem settings with e_l being the l^{th} standard basis vector in $\mathbb{R}^3$.

B. An Approximation in the Robust Formulation

Using (9) we present the following result.

Theorem 1. The formulation given in (8) is equivalent to

$$\begin{aligned} \hat{q}^* := \arg\min_q \quad & \left\{ -q^T \mathbf{K}(B,R) q + \max_{\|\delta\|_\infty \le 1} \left(\mathbf{p}(q,B,R)^T \delta + \delta^T \mathbf{Q}(q) \delta \right) \right\} \\ \text{subject to} \quad & q^T q = 1, \end{aligned} \tag{10}$$

where $\delta := \begin{bmatrix} \delta_{b1}^T & \delta_{r1}^T & \delta_{b2}^T & \delta_{r2}^T & \dots & \delta_{bn}^T & \delta_{rn}^T \end{bmatrix}^T$, $\mathbf{p} := \mathbb{R}^4 \times \mathbb{R}^{3\times n} \times \mathbb{R}^{3\times n} \to \mathbb{R}^{6n}$ is

$$\mathbf{p}(q,B,R) := \begin{bmatrix} w_1 \gamma_{b1} (b_{11} - q^T K_{r1}^1 q) \\ w_1 \gamma_{b1} (b_{12} - q^T K_{r1}^2 q) \\ w_1 \gamma_{b1} (b_{13} - q^T K_{r1}^3 q) \\ w_1 \gamma_{r1} (r_{11} - q^T K_{b1}^1 q) \\ w_1 \gamma_{r1} (r_{12} - q^T K_{b1}^2 q) \\ w_1 \gamma_{r1} (r_{13} - q^T K_{b1}^3 q) \\ \vdots \\ w_n \gamma_{rn} (r_{n1} - q^T K_{bn}^1 q) \\ w_n \gamma_{rn} (r_{n2} - q^T K_{bn}^2 q) \\ w_n \gamma_{rn} (r_{n3} - q^T K_{bn}^3 q) \end{bmatrix}, \tag{11}$$

where b_{ij} and r_{ij} are the j^{th} elements of the i^{th} vector. The definition of matrices K_{ri}^j and K_{bi}^j is given in Appendix B. The matrix $\mathbf{Q} := \mathbb{R}^4 \to \mathbb{R}^{6n\times 6n}$ is given as

$$\mathbf{Q}(q) := \begin{bmatrix} \frac{1}{2} w_1 \gamma_{b1}^2 I_3 & -\frac{1}{2} w_1 \gamma_{b1} \gamma_{r1} C & \dots & 0_{3\times3} & 0_{3\times3} \\ -\frac{1}{2} w_1 \gamma_{b1} \gamma_{r1} C^T & \frac{1}{2} w_1 \gamma_{r1}^2 I_3 & \dots & 0_{3\times3} & 0_{3\times3} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_{3\times3} & 0_{3\times3} & \dots & \frac{1}{2} w_1 \gamma_{bn}^2 I_3 & -\frac{1}{2} w_1 \gamma_{bn} \gamma_{rn} C \\ 0_{3\times3} & 0_{3\times3} & \dots & -\frac{1}{2} w_1 \gamma_{bn} \gamma_{rn} C^T & \frac{1}{2} w_1 \gamma_{rn}^2 I_3 \end{bmatrix}, \tag{12}$$

where the transformation matrix C is a function of q.

Proof: Using (9), the first term of (7) is written as

$$\frac{1}{2}\left(\mathrm{tr}(W\bar{B}^T\bar{B})+\mathrm{tr}(W\bar{R}^T\bar{R})\right)=\frac{1}{2}\left(\mathrm{tr}(WB^TB)+\mathrm{tr}(WR^TR)\right)+$$
$$\mathrm{tr}(WB^T\Delta_b)+\mathrm{tr}(WR^T\Delta_r)+\frac{1}{2}\left(\mathrm{tr}(W\Delta_b^T\Delta_b)+\mathrm{tr}(W\Delta_r^T\Delta_r)\right), \tag{13}$$

where $\Delta_b=\begin{bmatrix}\gamma_{b1}\delta_{b1} & \gamma_{b2}\delta_{b2} & \ldots & \gamma_{bn}\delta_{bn}\end{bmatrix}$ and $\Delta_r=\begin{bmatrix}\gamma_{r1}\delta_{r1} & \gamma_{r2}\delta_{r2} & \ldots & \gamma_{rn}\delta_{rn}\end{bmatrix}$. To simplify the second term, we first write

$$\mathbf{K}(\bar{B},\bar{R})=\mathbf{K}(B,R)+\mathbf{K}(B,\Delta_r)+\mathbf{K}(\Delta_b,R)+\mathbf{K}(\Delta_b,\Delta_r), \tag{14}$$

where $\mathbf{K}$ follows its usual definition (4) with

$$\mathbf{B}(\bar{B},\bar{R})=\mathbf{B}(B,R)+\mathbf{B}(B,\Delta_r)+\mathbf{B}(\Delta_b,R)+\mathbf{B}(\Delta_b,\Delta_r),$$
$$=BWR^T+BW\Delta_r^T+\Delta_bWR^T+\Delta_bW\Delta_r^T,$$

$$\mathbf{z}(\bar{B},\bar{R})=\mathbf{z}(B,R)+\mathbf{z}(B,\Delta_r)+\mathbf{z}(\Delta_b,R)+\mathbf{z}(\Delta_b,\Delta_r),$$
$$=(B\times R)W+(B\times\Delta_r)W+(\Delta_b\times R)W+(\Delta_b\times\Delta_r)W.$$

We first simplify and rearrange (13) and (14) and then write the expressions as a function of δ. Now separating the terms, which are linear or quadratic in δ and using the transformation matrix in terms of q [7], i.e.

$$C=\begin{bmatrix} q_1^2-q_2^2-q_3^2+q_4^2 & 2(q_1q_2+q_3q_4) & 2(q_1q_3-q_2q_4)\\ 2(q_1q_2-q_3q_4) & -q_1^2+q_2^2-q_3^2+q_4^2 & 2(q_2q_3+q_1q_4)\\ 2(q_1q_3+q_2q_4) & 2(q_2q_3-q_1q_4) & -q_1^2-q_2^2+q_3^2+q_4^2\end{bmatrix},$$

we can write the required expression. ■

It can be observed from (10) that the robust problem approaches the nominal problem if no uncertainty in the input vectors is considered. Finding the optimal solution of the formulated robust problem is difficult, because of the following two main reasons. Firstly, due to the matrix $\mathbf{Q}(q)$ being positive semidefinite, the maximization term in (10) is non-concave in δ, hence making it difficult to find a unique optimal maximum, and secondly, because of the matrix $K(B,R)$ being indefinite, the objective function is non-convex in q. To develop a tractable method for solving this problem, as a first step, we determine an upper bound on the maximum of $\mathbf{p}(q,B,R)^T\delta+\delta^T\mathbf{Q}(q)\delta$ over δ. The result is given in the following lemma, however the dependence of $\mathbf{p}(q,B,R)$ and $\mathbf{Q}(q)$ on B,R and q has been omitted for notational simplification.

Lemma 1. An upper bound on the maximization term appearing in (10) is

$$0\leq\max_{\|\delta\|_\infty\leq 1}\left(\mathbf{p}^T\delta+\delta^T\mathbf{Q}\delta\right)\leq\ \|\mathbf{p}\|_1+6n\lambda_{\max}(\mathbf{Q}) \tag{15}$$

Proof: We start with the following inequality

$$\max_{\|\delta\|_\infty \le 1} (\mathbf{p}^T \delta + \delta^T \mathbf{Q} \delta) \le \max_{\|\delta\|_\infty \le 1} \mathbf{p}^T \delta + \max_{\|\delta\|_\infty \le 1} \delta^T \mathbf{Q} \delta \tag{16}$$

Using the Hölder dual norm [26], the first term on the right hand side of (16) is given as

$$\max_{\|\delta\|_\infty \le 1} \mathbf{p}^T \delta = \|\mathbf{p}\|_1. \tag{17}$$

For the second term appearing in (16), since $\mathbf{Q}$ is a symmetric matrix, we can write the maximum eigenvalue of $\mathbf{Q}$ as [26]

$$\lambda_{\max}(\mathbf{Q}) = \sup_{\|\delta\|_2 \le 1} \delta^T \mathbf{Q} \delta \ . \tag{18}$$

Hence, we first replace the ∞-norm in the second term on the right hand side of (16) with the 2-norm using the inequality $\|\delta\|_2 \le \sqrt{6n}\|\delta\|_\infty$ for $\delta \in \mathbb{R}^{6n}$ [27]. We can write

$$\begin{aligned} \max_{\|\delta\|_\infty \le 1} \delta^T \mathbf{Q} \delta \le & \max_{\|\delta\|_2 \le \sqrt{6n}} \delta^T \mathbf{Q} \delta \\ \le & \quad 6n \lambda_{\max}(\mathbf{Q}), \end{aligned}$$

Using (17) and (19), we can write (15). ∎

Lemma 2. The maximum eigenvalue of the block diagonal matrix $\mathbf{Q}(q)$ does not depend on q.

Proof: To find the eigenvalues of the block diagonal matrix $\mathbf{Q}(q) = \text{diag}(\mathbf{Q}_1, \mathbf{Q}_2, \ldots, \mathbf{Q}_n)$, we need to solve n equations, i.e. $\det(\mathbf{Q}_i - \lambda I_6) = 0$, $i = 1, \ldots, n$. Consider the $i = 1$ case, where we can write

$$\det(\mathbf{Q}_1 - \lambda I_6) = \det \begin{bmatrix} \lambda_1(\lambda_2 - a) & 0 & 0 \\ 0 & \lambda_3(\lambda_4 - a) & \\ 0 & 0 & \lambda_5(\lambda_6 - a) \end{bmatrix} = 0,$$

where $a := \frac{1}{2} w_1(\gamma_{b1}^2 + \gamma_{r1}^2)$. The above equation implies that $\lambda_1 = \lambda_3 = \lambda_5 = 0$, and $\lambda_2 = \lambda_4 = \lambda_6 = \frac{1}{2} w_1(\gamma_{b1}^2 + \gamma_{r1}^2)$. Similarly we can find eigenvalues for $\mathbf{Q}_i$, $i = 2, \ldots, n$. Finally, $\lambda_{max}(\mathbf{Q}(q)) = \max_i \frac{1}{2}(w_i \gamma_{bi}^2 + w_i \gamma_{ri}^2)$. However, it is evident that the maximum eigenvalue is independent of q. ∎

C. Comparison of the Analytical Upper Bound

This section discusses the tightness of the upper bound given in (15). Since the maximization term in (10) is convex in δ, it is hard to find the optimum. For such problems, fairly tight bounds can be obtained using semidefinite relaxation [28]. We will find an upper bound using semidefinite relaxation and compare it with the analytical bound. For a given B, R and q, we define $\mathsf{p} := \mathbf{p}(q, B, R)$ and $\mathsf{Q} := \mathbf{Q}(q)$.

Consider the maximization term in (8). Suppose $\overline{\gamma}$ is an upper bound on this term satisfying the constraint $\|\delta\|_\infty \leq 1$, which can also be written as $-e \leq \delta \leq e$, where $e \in \mathbb{R}^{6n}$ is a vector of ones. Let $D \in \mathbb{R}^{6n\times 6n}$ be a diagonal matrix, then the following identity is always true:

$$\delta^T \mathsf{Q}\delta + \mathsf{p}^T\delta - \overline{\gamma} = -(e+\delta)^T D(e-\delta) - \begin{bmatrix}\delta^T & 1\end{bmatrix} \begin{bmatrix} D-\mathsf{Q} & -\frac{\mathsf{p}}{2} \\ -\frac{\mathsf{p}}{2} & \overline{\gamma} - e^T De\end{bmatrix} \begin{bmatrix}\delta \\ 1\end{bmatrix}.$$

In this expression, we know from the constraint $\|\delta\|_\infty \leq 1$, that both $(e+\delta) \geq 0$ and $(e-\delta) \geq 0$. Now the minimum value of $\overline{\gamma}$ represents an upper bound on the maximization term if the diagonal matrix $D \succeq 0$, and matrix $\mathscr{F}(D,\overline{\gamma}) := \begin{bmatrix} D-\mathsf{Q} & -\frac{\mathsf{p}}{2} \\ -\frac{\mathsf{p}}{2} & \overline{\gamma} - e^T De\end{bmatrix} \succeq 0$, i.e.

$$\max_{\|\delta\|_\infty \leq 1} (\delta^T \mathsf{Q}\delta + \mathsf{p}^T\delta) \leq \min_{D,\gamma} \{\overline{\gamma} \mid D \succeq 0, \mathscr{F}(D,\overline{\gamma}) \succeq 0\}, \tag{19}$$

which is a semidefinite relaxation (SDR) of the maximization term of (8) for a given value of q, B and R. A comparison of the analytical bound and the bound obtained using the SDR is given in Section V, showing a small relative error between the two. Thus, use of the analytical bound gives computational benefits, but at the cost of a loss in accuracy of the true solution of (10), although the analysis shows that the loss is small assuming that the gap between (10) and its SDR is small.

At this stage, one might think of (19) as being tighter, instead of (15) used to simplify (8). Note that (19) is based on the assumption that q is known. However, if q is unknown, which is the case in actual problems, we may not get much computational benefits, because the terms of the matrix inequality $\mathscr{F}(D,\overline{\gamma}) \succeq 0$ are nonlinear in q. In this work, however, we used (15) to obtain a tractable, but suboptimal solution of (10).

D. Addition of a Regularization Term

Use of the analytical upper bound in the robust problem (8) introduces an approximation, although it was motivated for computational benefits and the fact that a unique solution of the inner loop maximization may not be guaranteed. However, this approximation may degrade the algorithm performance in terms of robustness. To improve performance, we introduce a new type of regularization in the objective function. The basic idea in using this regularization is given now.

The optimization variable in (8), i.e the quaternion q which represents a coordinate transformation as a consequence of the Euler theorem of rotation [7], is defined as

$$q = \begin{bmatrix} \mathbf{q} \\ q_4 \end{bmatrix} := \begin{bmatrix} \hat{\mathbf{e}} \sin(\alpha/2) \\ \cos(\alpha/2) \end{bmatrix}, \tag{20}$$

where $\hat{\mathbf{e}} \in \mathbb{R}^3$ is the axis of rotation and α is the angle of rotation. We propose minimizing an additional term $-\eta q_4^2$ along with the primary objective function. This regularization term is similar in concept to the Tikhonov regularization in linear least squares problems [26]. In such problems, the regularization term is normally a norm of the solution vector. While minimizing the regularized cost, the added term enforces a trade-off between the primary objective and the norm of the solution vector. However, in our case, since the norm of the solution vector is 1 ($q^T q = 1$), the introduced regularization term is a function of q_4, which corresponds to the angle of rotation for a given quaternion. Hence, the added term ηq_4^2 enforces finding a q, which minimizes a weighted combination of both objectives. In the added term, $\eta > 0$ is a tuning parameter. A large value of η will make the optimal solution of the regularized problem stiff to perturbations with a large residual in the nominal case. Simulations have shown that in the considered environment, $\eta = 0.5$ gives good results with a smaller residual and a reasonably large robust performance margin compared to the nominal solution.

Using the subsequent results and discussion, we now present the final simplified formulation of the robust problem in a form suitable for semidefinite relaxation.

Corollary 1. The max term in (10), being convex, may not always give a true worst case. Replacing the max term using Lemmas 1 and 2, along with the regularization term, (10) is approximated with the following maximization problem

$$\begin{aligned} (q^*, u^*) = \arg\max_{q,u} \quad & q^T \mathbf{K}_r(B,R) q - u^T e \\ \text{subject to} \quad & q^T q = 1, \\ & -u \leq \mathbf{p}(q,B,R) \leq u, \end{aligned} \tag{21}$$

where $u := \begin{bmatrix} u_1 & u_2 & \dots & u_{6n} \end{bmatrix}^T \geq 0$, $\mathbf{K}_r(B,R) := \mathbf{K}(B,R) + \eta S$ and $S = \mathrm{diag}(0,0,0,1)$. Here we neglect all constant terms having no effect on the argument of the optimization problem.

Proof: In (10), we replace the max term with the upper bound given in Lemma 1 and neglect the term involving $6n\lambda_{\max}(\mathbf{Q}(q))$, as it does not depend on q according to Lemma 2 and will not effect the solution. We can represent the regularization term as $\eta q_4^2 = \eta q^T S q$. Finally, using the fact that a set of $12n+1$ linear inequalities $-u_j \leq x_j \leq u_j$, $\sum_j u_j \leq 1$, $j = 1, \dots, 6n$ represent the nonlinear inequality $\sum_j |x_j| \leq 1$ [25, Definition 1.3.1] and expressing it as a maximization problem, we can write (21). All constant terms in the expression are neglected, however for an exact upper bound, these terms need to be added in the bound obtained from SDR. ■

IV. Semidefinite Relaxation for the Robust Estimation Problem

In this section we apply semidefinite relaxation on (21) [23]. Suppose $\bar{\gamma}$ is an upper bound for the objective function of (21). Using a similar approach, as used in deriving (19), we obtain the following expression, such that the right hand side is equal to the left hand side. Again we will drop the dependence for notational simplification, except where necessary:

$$\begin{aligned} q^T \mathbf{K}_r q - u^T e - \bar{\gamma} = &-\mu_1(1 - q^T q) - \mu_2(u_1 - p_1) - \mu_3(u_1 + p_1) \\ &- \mu_4(u_2 - p_2) - \mu_5(u_2 + p_2) - \cdots - \mu_{12n}(u_{6n} - p_{6n}) \\ &- \mu_{12n+1}(u_{6n} + p_{6n}) - x^T \mathscr{L}(\mu, B, R) x, \end{aligned}$$

where $x := \begin{bmatrix} q^T & u^T & 1 \end{bmatrix}^T$, $\mu := \begin{bmatrix} \mu_1 & \mu_2 & \dots & \mu_{12n+1} \end{bmatrix}^T$, $\mathbf{p}(q,B,R) := \begin{bmatrix} p_1 & p_2 & \dots & p_{6n} \end{bmatrix}^T$,

$$\mathscr{L}(\mu,B,R) := \begin{bmatrix} \mathscr{L}_{1,1}(\mu,B,R) & 0_{4\times 1} & \dots & 0_{4\times 1} & 0_{4\times 1} \\ 0_{1\times 4} & 0 & \dots & 0 & \frac{1-\mu_2-\mu_3}{2} \\ 0_{1\times 4} & 0 & \dots & 0 & \frac{1-\mu_4-\mu_6}{2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_{1\times 4} & 0 & \dots & 0 & \frac{1-\mu_{12n}-\mu_{12n+1}}{2} \\ 0_{1\times 4} & \frac{1-\mu_2-\mu_3}{2} & \dots & \frac{1-\mu_{12n}-\mu_{12n+1}}{2} & \ell_{j,j}(\mu,B,R) \end{bmatrix}$$

$$\begin{aligned} \mathscr{L}_{1,1}(\mu,B,R) := &\mu_1 I_4 - (\mu_2 - \mu_3) w_1 \gamma_{b1} K_{r1}^1 - (\mu_4 - \mu_5) w_1 \gamma_{b1} K_{r1}^2 \\ &- (\mu_6 - \mu_7) w_1 \gamma_{b1} K_{r1}^3 - (\mu_8 - \mu_9) w_1 \gamma_{r1} K_{b1}^1 - \dots \\ &- (\mu_{12n} - \mu_{12n+1}) w_n \gamma_{r1} K_{bn}^3 - \mathbf{K}_r(B,R), \\ \ell_{j,j}(\mu,B,R) := &\bar{\gamma} - \mu_1 + \sum_{l=1}^{6n} (\mu_{2l} - \mu_{2l+1}) \mathbf{c}_l(B,R), \end{aligned}$$

where j is the size of x and $\mathbf{c}(B,R) := \begin{bmatrix} w_1 \gamma_{b1} b_1^T & w_1 \gamma_{r1} r_1^T & \dots & w_n \gamma_{bn} b_n^T & w_n \gamma_{rn} r_n^T \end{bmatrix}^T$. Now if the right hand side is either zero or negative, we can say that $\bar{\gamma}$ is an upper bound on the cost of (21). Using this relaxation, we write an optimization problem to find the minimum value of this upper bound ensuring the right hand side is either zero or negative, given as

$$(\bar{\gamma}^*, \mu^*) := \arg\min_{\bar{\gamma},\mu} \{\bar{\gamma} \mid \mathscr{L}(\mu,B,R) \succeq 0,\ \mu_i \geq 0,\ i = 2,3,\dots,12n+1\}. \tag{22}$$

Note that few diagonal entries of the matrix $\mathscr{L}(\mu,B,R)$ are zero. For this matrix to be positive semidefinite, we can force the corresponding non-diagonal terms to zero. This will result in a reduced set of optimization variables and will also avoid numerical issues arising due to the zero diagonal entries.

Theorem 2. Using a reduced set of optimization variables $\mu_r := \begin{bmatrix} \mu_1 & \mu_2 & \mu_4 & \dots & \mu_{12n} \end{bmatrix}^T$, an equivalent formulation of (22) is

$$\begin{aligned} \mu_r^* = \arg\min_{\mu_r} \quad & \mu_1 - \sum_{l=1}^{6n} (2\mu_{2l} - 1)\, \mathbf{c}_l(B,R) \\ \text{subject to} \quad & 0 \le \mu_i \le 1,\, i = 2,4,\dots,12n, \\ & \mathscr{L}_{1,1}(\mu_r, B, R) \succeq 0, \end{aligned} \tag{23}$$

where $\mathscr{L}_{1,1}(\mu_r, B, R)$ is given by

$$\begin{aligned} \mathscr{L}_{1,1}(\mu_r, B, R) := {} & \mu_1 I_4 - 2\mu_2 w_1 \gamma_{b1} K_{r1}^1 - 2\mu_4 w_1 \gamma_{b1} K_{r1}^2 - 2\mu_6 w_1 \gamma_{b1} K_{r1}^3 \\ & - 2\mu_8 w_1 \gamma_{r1} K_{b1}^1 - \dots - 2\mu_{12n} w_n \gamma_{rn} K_{bn}^3 + \dots + w_1 \gamma_{b1} K_{r1}^1 + w_1 \gamma_{b1} K_{r1}^2 \\ & + w_1 \gamma_{b1} K_{r1}^3 + w_1 \gamma_{r1} K_{b1}^1 + w_{12n} \gamma_{rn} K_{bn}^3 - \mathbf{K}_r(B,R). \end{aligned}$$

Proof: Note that in (22) the symmetric matrix $\mathscr{L}(\mu, B, R)$ has zero diagonal elements. For $\mathscr{L}(\mu, B, R)$ to be positive semidefinite, as required in (22), all row/column elements corresponding to zero diagonal entries must also be zero [27, Thm 4.2.6], i.e. $1 - \mu_2 - \mu_3 = 0, 1 - \mu_4 - \mu_5 = 0, 1 - \mu_6 - \mu_7 = 0$ and so on. Using this property, we can force these elements to zero by eliminating $\mu_3, \mu_5, \dots, \mu_{12n+1}$ from (22) with additional constraints $1 - \mu_2 \ge 0, 1 - \mu_4 \ge 0, \dots, 1 - \mu_{12n} \ge 0$. Moreover, the minimum value of $\bar{\gamma}$ satisfying the constraint $\mathscr{L}(\mu_r, B, R) \succeq 0$ results in $\ell_{j,j}(\mu_r, B, R) = 0$, giving

$$\bar{\gamma} = \mu_1 - \sum_{l=1}^{6n} (2\mu_{2l} - 1)\, \mathbf{c}_l(B,R). \tag{24}$$

So with these modifications, instead of $\mathscr{L}(\mu, B, R) \succeq 0$, we only need $\mathscr{L}_{1,1}(\mu_r, B, R) \succeq 0$, hence can write (23) using a reduced number of optimization variables, which is equivalent to solving (22) for the minimum upper bound on (21). ■

V. Finding the Robust Quaternion (q^*)

Although the solution of the semidefinite program (23) gives a minimum upper bound on the robust estimation problem (21), our main interest is to find a q^* that could maximize the cost (21). Now the question arises, can we find q^* using the solution μ_r^* of (23)? Suppose μ_r^* results in a zero value of the right hand side of (22), then $\bar{\gamma}^*$, i.e. the minimum value of cost (23), is equal to the maximum cost of (21), and the corresponding q will be the required q^*.

In this regard, as a first step, we establish whether there exists a q that can make $q^T \mathscr{L}_{1,1}^* q = 0$, where $\mathscr{L}_{1,1}^* := \mathscr{L}_{1,1}(\mu_r^*, B, R)$. If such a q exists, it will further ensure $x^T \mathscr{L}^* x = 0$, where $\mathscr{L}^* := \mathscr{L}(\mu^*, B, R)$ and μ^* can be obtained from μ_r^*.

Lemma 3. Let μ_r^* be a minimizer for the SDR problem (23), then $\lambda_{\min}(\mathscr{L}_{1,1}^*) = 0$.

Proof: Using μ_r, the objective function of (23) can be written as $J := \mu_1 - d$, where d is the sum of all remaining terms. Now whatever the sign of d is, the cost J is minimum when μ_1 is minimum. However, at the same time we need $\mathscr{L}_{1,1}(\mu_r, B, R) \succeq 0$. We can also write $\mathscr{L}_{1,1}(\mu_r, B, R) = \mu_1 I_4 - K_\mu$, where K_μ is the sum of all other terms in the expression. This is a symmetric matrix with real eigenvalues $\lambda_1, \ldots, \lambda_4$, and $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4$. Then, $\mu_1 I_4 - K_\mu$ will have eigenvalues $\mu_1 - \lambda_1, \mu_1 - \lambda_2, \mu_1 - \lambda_3, \mu_1 - \lambda_4$. Now, $\mu_1 = \lambda_1$ is the smallest possible value that can make $\mathscr{L}_{1,1}(\mu_r, B, R) \succeq 0$. This optimal value of μ_1, i.e. μ_1^*, will ensure $\lambda_{\min}(\mathscr{L}_{1,1}^*) = 0$. ■

Remark 1. As stated in Lemma 3, the matrix $\mathscr{L}_{1,1}^*$ has at least one eigenvalue equal to zero. Suppose there is only one eigenvalue equal to zero and $\tilde{q}$ is an eigenvector of $\mathscr{L}_{1,1}^*$ corresponding to the zero eigenvalue, then this $\tilde{q}$ will result in both $\tilde{q}^T \mathscr{L}_{1,1}^* = 0$ and $\tilde{q}^T \mathscr{L}_{1,1}^* \tilde{q} = 0$, because $\tilde{q}$ will belong to the null space of $\mathscr{L}_{1,1}^*$. From this we can deduce that $\tilde{x}^T \mathscr{L}^* \tilde{x} = 0$, where $\tilde{x} := \begin{bmatrix} \tilde{q}^T & \tilde{u}^T & 1 \end{bmatrix}^T$, although we have no knowledge of $\tilde{u}$ at this stage. This is possible because all elements of matrix $\mathscr{L}^*$ are zero, except sub-matrix $\mathscr{L}_{1,1}^*$.

The vector $\tilde{q}$ can be a candidate for the robust quaternion. Now, if there is no gap between the cost of (21) and (23), then $\tilde{q}$ will be the required robust optimal solution of (21).

A. Relaxation Gap

To quantify the gap between the approximate robust problem (21) and its relaxation (23), we present the following result.

Lemma 4. Let μ_r^* be a minimizer for the SDR problem (23), such that $h = \dim \mathscr{N}\left(\mathscr{L}_{1,1}(\mu_r^*)\right) \geq 1$ and

$$\mathscr{L}_{1,1}(\mu_r^*) = \begin{bmatrix} V & V_+ \end{bmatrix} \begin{bmatrix} 0_h & 0 \\ 0 & \Lambda_+ \end{bmatrix} \begin{bmatrix} V \\ V_+ \end{bmatrix} \tag{25}$$

be a spectral decomposition of $\mathscr{L}_{1,1}(\mu_r^*)$ for some orthogonal $\begin{bmatrix} V & V_+ \end{bmatrix}$ and $\Lambda_+ \succ 0$. Consider the optimal cost of (23) to be $J(\mu_r^*)$. Let $z = \begin{bmatrix} z_1 & z_2 & z_4 & \ldots & z_{12n} \end{bmatrix}^T \geq 0$, where $z_i \in \mathbb{R}, i = 1, 2, 4, \ldots, 12n$, then there does not exist a z such that

1) $J(\mu_r^* - z) = J(\mu_r^*)$, i.e. $J(z) = 0$
2) $\mu_{2i}^* \geq z_{2i}$, $i = 1, \ldots, 6n$,
3) $1 - \mu_{2i}^* + z_{2i} \geq 0$, $i = 1, \ldots, 6n$,
4) $V^T \mathscr{L}_0(z) V < 0$, where $\mathscr{L}_0(z) = \mathscr{L}_{1,1}(\mu_r^* - z) - \mathscr{L}_{1,1}(\mu_r^*)$.

Proof: Suppose such a z exists. We choose a small value of $\varepsilon > 0$ such that $\mu_r^* - \varepsilon z$ is another solution to (23), satisfying all above points. We evaluate $\mathscr{L}_{1,1}(\mu_r^* - \varepsilon z) = \mathscr{L}_{1,1}(\mu_r^*) - \varepsilon \mathscr{L}_0(z)$, and write

$$\begin{bmatrix} V^T \\ V_+^T \end{bmatrix} \mathscr{L}_{1,1}(\mu_r^* - \varepsilon z) \begin{bmatrix} V & V_+ \end{bmatrix} = \begin{bmatrix} V^T \\ V_+^T \end{bmatrix} \Big(\mathscr{L}_{1,1}(\mu_r^*) - \varepsilon \mathscr{L}_0(z) \Big) \begin{bmatrix} V & V_+ \end{bmatrix}.$$

Using (25), we can write

$$\begin{bmatrix} V^T \\ V_+^T \end{bmatrix} \mathscr{L}_{1,1}(\mu_r^* - \varepsilon z) \begin{bmatrix} V & V_+ \end{bmatrix} = \begin{bmatrix} -\varepsilon V^T \mathscr{L}_0(z) V & -\varepsilon V^T \mathscr{L}_0(z) V_+ \\ -\varepsilon V_+^T \mathscr{L}_0(z) V & \Lambda_+ - \varepsilon V_+^T \mathscr{L}_0(z) V_+ \end{bmatrix}.$$

Now, from point 4, we know that $-V^T \mathscr{L}_0(z) V > 0$ and

$$-V^T \mathscr{L}_0(z) V - \varepsilon V^T \mathscr{L}_0(z) V_+ \Big(\Lambda_+ - \varepsilon V_+^T \mathscr{L}_0(z) V_+ \Big)^{-1} V_+^T \mathscr{L}_0(z) V > 0,$$

because $\Lambda_+ > 0$ and we can choose $\varepsilon > 0$ such that $\Lambda_+ - \varepsilon V_+^T \mathscr{L}_0(z) V_+ > 0$ and the above is true. Using the Schur complement condition for positive definiteness, the above implies that $\mathscr{L}_{1,1}(\mu_r^* - \varepsilon z) \succ 0$. However, this contradicts with the requirement for $\mu_r^* - \varepsilon z$ to be another solution according to Lemma 3. ■

Next, we present our main result regarding the gap between the SDR and (21) and will also relate the vector $\tilde{q}$ determined in Remark 1 and q^*, i.e. the solution of (21).

Theorem 3. For the $h = 1$ case, the vector $\tilde{q}$, which makes $\tilde{q}^T \mathscr{L}_{1,1}^* \tilde{q} = 0$ will ensure no relaxation gap between the approximate problem (21) and its semidefinite relaxation (23), making $\tilde{q} = q^*$.

Proof: For no gap, we need to prove each term on the right hand side of (22) is zero. We use $\tilde{q}$ obtained from Remark 1, satisfying $\tilde{q}^T \tilde{q} = 1$ and $\mathscr{L}_{1,1}(\mu_r^*) \tilde{q} = 0$.

1) Satisfying $\tilde{q}^T \tilde{q} = 1$ implies $\mu_1 (1 - \tilde{q}^T \tilde{q}) = 0$.
2) Satisfying $\mathscr{L}_{1,1}(\mu_r^*) \tilde{q} = 0$ implies $\tilde{x}^T \mathscr{L}(\mu^*) \tilde{x} = 0$.
3) To prove that remaining terms are zero, we first show that
   a) if $\mu_{2i} \neq 0$, then $p_i \geq 0$, $i = 1, \ldots, 6n$.
   b) if $\mu_{2i+1} \neq 0$, then $p_i \leq 0$, $i = 1, \ldots, 6n$.

   To prove (a), first we write the optimal cost function of (23) in terms of p_i. For this, pre and post multiplying both sides of (24) by $\tilde{q}^T$ and $\tilde{q}$, and using the fact that $\tilde{q}^T \mathscr{L}_{1,1}(\mu_r^*) \tilde{q} = 0$, we can write

$$\begin{aligned} 0 = \mu_1 &- (2\mu_2 - 1) w_1 \gamma_{b1} \tilde{q}^T K_{r1}^1 \tilde{q} - (2\mu_4 - 1) w_1 \gamma_{b1} \tilde{q}^T K_{r1}^2 \tilde{q} \\ &- (2\mu_6 - 1) w_1 \gamma_{b1} \tilde{q}^T K_{r1}^3 \tilde{q} - (2\mu_8 - 1) w_1 \gamma_{r1} \tilde{q}^T K_{b1}^1 \tilde{q} - \ldots \\ &- (2\mu_{12n} - 1) w_n \gamma_{r1} \tilde{q}^T K_{bn}^3 \tilde{q} - \tilde{q}^T \mathbf{K}_r(B, R) \tilde{q}. \end{aligned}$$

Finally, subtracting (26) from (24), we can write

$$\bar{\gamma} = \tilde{q}^T \mathbf{K}_r(B,R)\tilde{q} - (2\mu_2^* - 1)p_1 - (2\mu_4^* - 1)p_2 - \ldots$$
$$- (2\mu_{12n}^* - 1)p_{6n}.$$

Now, consider the case $i = 1$, i.e. if $p_1 < 0$, we need to prove $\mu_2^* = 0$. Let us contradict by assuming that $\mu_2^* > 0$. Then there exist a $z_2 > 0$ such that $\mu_2^* \geq z_2$. We assume $z_4, \ldots, z_{12n}$ to be zero. Using Lemma 4 (point 1), we have $z_1 = 2z_2c_1$, i.e. the new z satisfies points 1-3 of Lemma 4. Then, using $q^T \mathscr{L}_{1,1}(\mu_r^*)q = 0$ we have

$$q^T \mathscr{L}_{1,1}(\mu_r^* - z)q = \tilde{q}^T(-z_1 I_4 + 2z_2 w_1 \gamma_{b1} \tilde{q}^T K_{r1}^1)\tilde{q},$$
$$q^T \mathscr{L}_0(z)q = 2z_2 p_1.$$

Here, as $p_1 < 0$ and $z_2 > 0$, we have $\mathscr{L}_0 \prec 0$, which is against Lemma 4. Hence we conclude that such a z_2 is not possible and $\mu_2^* = 0$. Using a similar approach, we can obtain such results for all values of i, proving part (a). Similarly, for part (b), we need to show that if $p_i > 0$, then $\mu_{2i+1} = 0$ or in reduced variable settings $1 - \mu_{2i} = 0$, using the condition $\mu_{2i} + \mu_{2i+1} = 1$ and the constraint $1 - \mu_{2i} \geq 0$. Now, following a similar approach as part (a), we can write for $i = 1$

$$q^T \mathscr{L}_0(z)q = -2(1 - z_2)p_1. \quad (26)$$

Since $p_1 > 0$ and $1 - z_2 > 0$, hence $\mathscr{L}_0 \prec 0$, which is not possible from Lemma 4, proving part (b).

Finally, we prove there exists a $u \geq 0$, such that the remaining terms in (22) are zero. Since $\mu_{2i} + \mu_{2i+1} = 1$ and $\mu_{2i}, \mu_{2i+1} \geq 0$ and $\|\mathbf{p}\|_1 = u$, there are three possibilities:

i) $\mu_{2i} = 1$, $\mu_{2i+1} = 0$: From (a), we know that in this case $p_i \geq 0$ and we define $u_i = p_i$.

ii) $\mu_{2i} = 0$, $\mu_{2i+1} = 1$: From (b), we know that in this case $p_i \leq 0$ and we define $u_i = -p_i$.

iii) $\mu_{2i} \neq 0$, $\mu_{2i+1} \neq 0$: From (a) and (b), we know that in this case $p_i = 0$ and we define $u_i = 0$.

■

It has been observed in the numerical simulations that $h > 1$ is rare. However, if such a case occurs, more than one solution is possible i.e. the eigenvectors corresponding to the zero eigenvalues. For such solutions, a zero relaxation gap cannot be guaranteed. However, the gap will be small, because any of the solutions will result in some of the terms on the right hand side of (22) to be zero.

VI. SIMULATION RESULTS

We consider the attitude determination for a low cost CubeSat [29], a pico-satellite moving in a circular orbit at an average altitude of 650 km above earth surface. To find attitude, we assumed the use of two measurements, namely the earth magnetic field and the sun vector. For the earth magnetic field, normally two magnetometers are installed, one inside the satellite for the post-launch tumbling phase, while the second is installed on an extended boom, which is deployed once the satellite has achieved an equilibrium. The sun vector is sensed by a pair of sun sensors installed on the satellite. Both of these measurements are in the body frame. For the earth magnetic field in the reference frame, we used the 1^{st} order IGRF model [21], while the reference sun vector is obtained using a simplified sun model based on the sun ephemeris [30]. Both sensor measurements and reference vectors are not accurate. For example, sensor measurements are affected by noise and misalignments. Especially in the post-launch tumbling phase, the measurement errors further increase due to the use of an internal magnetometer, which interacts with the magnetic field generated by the surrounding electronics. Similarly, the reference vectors are also not accurate, because they are obtained from mathematical models, normally based on low-order approximations for computational benefits. In this work, we consider all such errors as ∞-norm bounded uncertainties, and for simulations we set an uncertainty bound of 30% of the norm of the vectors in the body and the reference frame.

A. Tightness of the Analytical Upper Bound

A comparison of the analytical upper bound (15) with the bound obtained from the semidefinite relaxation (19) is presented. We used two pairs of unit vectors, one in the body and the other in the reference frame, given as:

$$\begin{aligned} b_1 &= \begin{bmatrix} -0.542 & -0.316 & 0.779 \end{bmatrix}^T, \quad & r_1 &= \begin{bmatrix} -0.529 & -0.335 & 0.78 \end{bmatrix}^T, \\ b_2 &= \begin{bmatrix} -0.673 & 0.02 & 0.739 \end{bmatrix}^T, \quad & r_2 &= \begin{bmatrix} -0.666 & 0.00037 & 0.746 \end{bmatrix}^T. \end{aligned} \tag{27}$$

A uniformly distributed bounded random error is introduced in the vectors for each simulation. A comparison of both bounds and their relative error for 100 simulations is given in Figure 1. The plot shows that the relative error is less than 2% on average and less than 5% in the worst cases. This analysis reveals that the price paid for using the analytical bound is not much, provided the SDR bound is close to the actual value.

B. Performance Comparison for One Time Instant Data

The effect of uncertainty on the robust and non-robust solutions is presented for a given set of data for one time instant. A number of tests were performed by adding uncertainty in the input vectors within the set bounds. The set of test vectors is given as:

$$\begin{aligned} b_1 &= \begin{bmatrix} -0.776 & -0.46 & 0.43 \end{bmatrix}^T, \quad & r_1 &= \begin{bmatrix} -0.54 & -0.326 & 0.775 \end{bmatrix}^T, \\ b_2 &= \begin{bmatrix} -0.927 & 0.01 & 0.374 \end{bmatrix}^T, \quad & r_2 &= \begin{bmatrix} -0.673 & 0.000133 & 0.74 \end{bmatrix}^T. \end{aligned} \tag{28}$$

The non-robust solution is obtained satisfying (3), while the robust solution satisfies (21). Figure 2 presents a histogram of the distribution of the cost of (6) for different cases of added uncertainty. The x-axis represents the

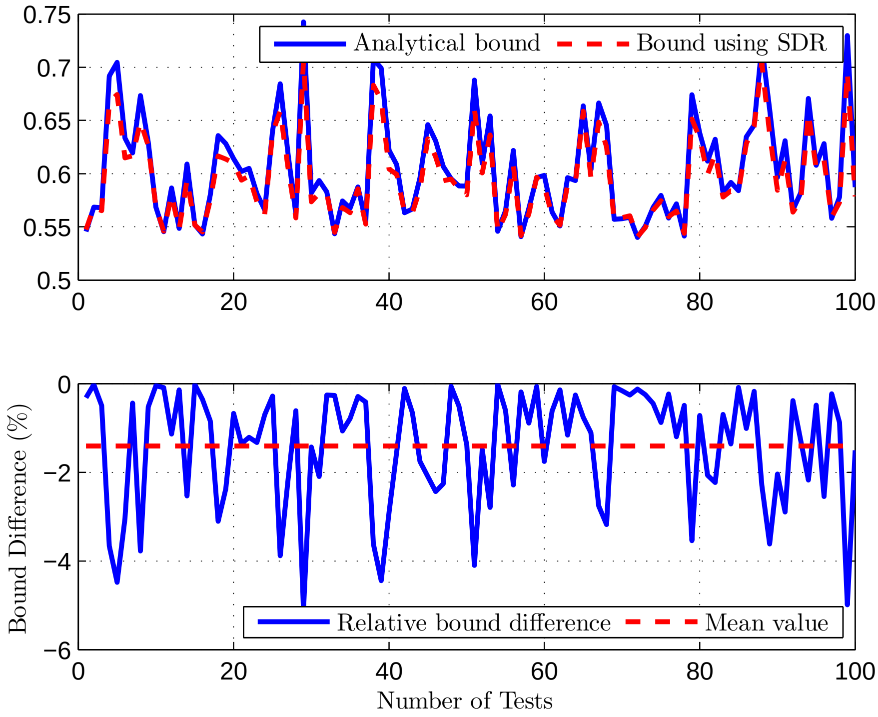

Fig. 1. Comparison of the analytical bound (15) and the bound obtained from SDR (19). Bound difference is given as relative error in percent.

cost value and the y-axis shows the number of tests. The spread of the cost for bounded uncertainties using the nominal solution is much more than the robust solution, showing the usefulness of the robust approach.

C. Effect of the Regularization Term

This section analyzes the effect of the regularization term added in the cost function of (21). Figure 3 shows the effect of variation of the tuning parameter η on the robust performance. In this analysis, we varied first two components of each input vector parameterized with a single variable varying in the range -1 to 1. On the y-axis, we plot the cost value J. It can be observed that the solution without regularization shows robustness against uncertainties compared to the nominal solution; however, in some cases the benefit is not significant. The regularization term in both of these situations improves robust performance. Moreover, η gives the user an option to choose the robustness margins.

D. Comparison of the Quaternion Obtained from the Approximate Problem and the SDR

We present a quantitative comparison of the optimal quaternion obtained from (21) using MATLAB's `fmincon` (with interior-point algorithm, tolerance of 10^{-12} and an initial guess of eigenvector of K corresponding to the largest eigenvalue) and the solution of (23) using `mincx` (with the same tolerance). We used the perturbed vector data given in (28). A comparison is given in Table I. Note that q^* is obtained using Remark 1. The error between the two quaternions is negligible.

TABLE I
QUATERNION OBTAINED FROM (21) AND (23) FOR THE VECTOR SET (28)

q^*	$\tilde{q}$	$\lvert q^* - \tilde{q}\rvert$
0.0761303170	0.0761303011	1.59928936×10^{-8}
0.0444603409	0.0444603345	6.39020509×10^{-9}
-0.0305429683	-0.0305429452	-2.31056824×10^{-8}
0.9956377755	0.9956377777	-2.21610053×10^{-9}

E. Robust Performance Comparison for In-Orbit Simulation Data

This section compares the robust and non-robust approaches in the presence of uncertainties, using in-orbit data obtained from a nonlinear simulation of the satellite initialized with roll, pitch and yaw body rates of 0.5, 0.5 and 0.1 deg/s and roll, pitch and yaw angles of 10, 0, 0 deg, respectively. The ideal data was corrupted by adding uniformly distributed random errors in the range of $\pm\gamma_{bi}$ and $\pm\gamma_{ri}$ in the corresponding vectors. We present attitude determination results for 25 minutes of flight data obtained with a sample time of 1 second. We solved the robust problem formulated in (21) using the nonlinear optimization solver `fmincon` of MATLAB, while the problem formulated using semidefinite relaxation in (23) was solved using the Robust Control toolbox command `mincx`. Figure 4 shows the benefit of the robust over the non-robust approach in the presence of uncertainties, where the non-robust approach gives large errors in the attitude, while the robust approach gives much better performance, limiting the maximum error to a smaller band.

F. Analysis of Sections IV and V: Theoretical Results

Lastly, an analysis is presented to support different theoretical results presented in Section IV and V. Figures 5 and 6 support Theorem 3. Figure 5 shows the relaxation gap between the robust problem and its semidefinite relaxation i.e. $q^T K_r q - u^T e - \overline{\gamma}$. It can be observed that the gap is zero for all time instances. Figure 6 shows the difference between the quaternion obtained from the two solutions. For the SDR case, the quaternions are obtained using Remark 1. It can be observed that both the relaxation gap and quaternion error is almost zero for all time instances.

VII. CONCLUSION

A robust attitude estimation problem was formulated for ∞-norm bounded uncertainties in the measurement and model vectors. The robust min-max optimization problem was transformed into a suboptimal minimization problem with non-convex quadratic cost and constraints. An additional regularization term was proposed to improve robust performance. Semidefinite relaxation was used to transform this non-convex QCQP into a semidefinite program with a linear cost and linear matrix inequality constraints. It was also shown how to extract the robust attitude

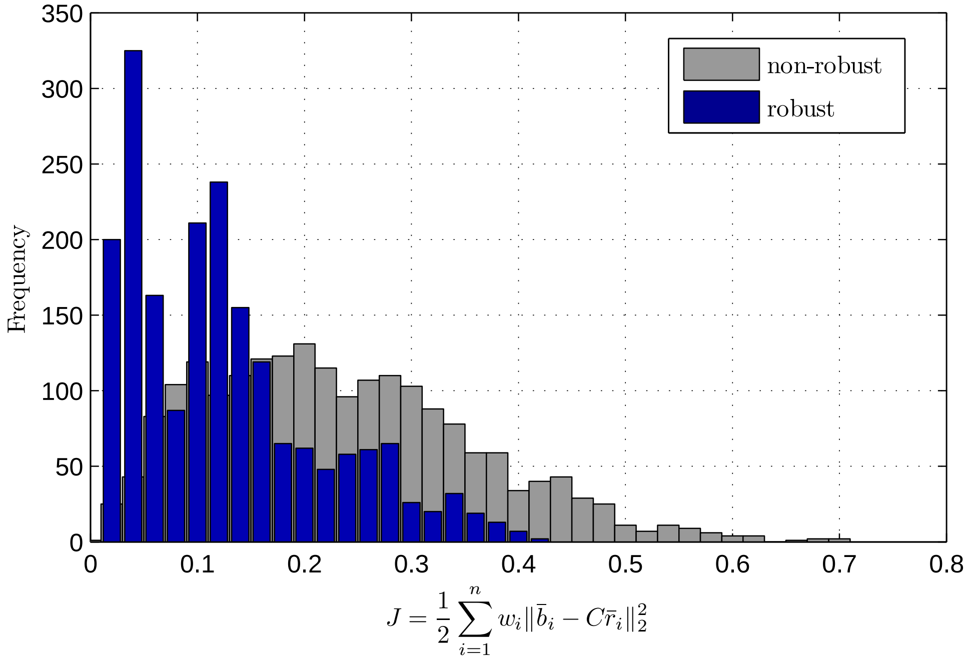

Fig. 2. Histogram showing distribution of the cost of the robust and non-robust designs. The plot shows data for 2000 runs. In each run, uniformly distributed bounded random error is added in the test vectors.

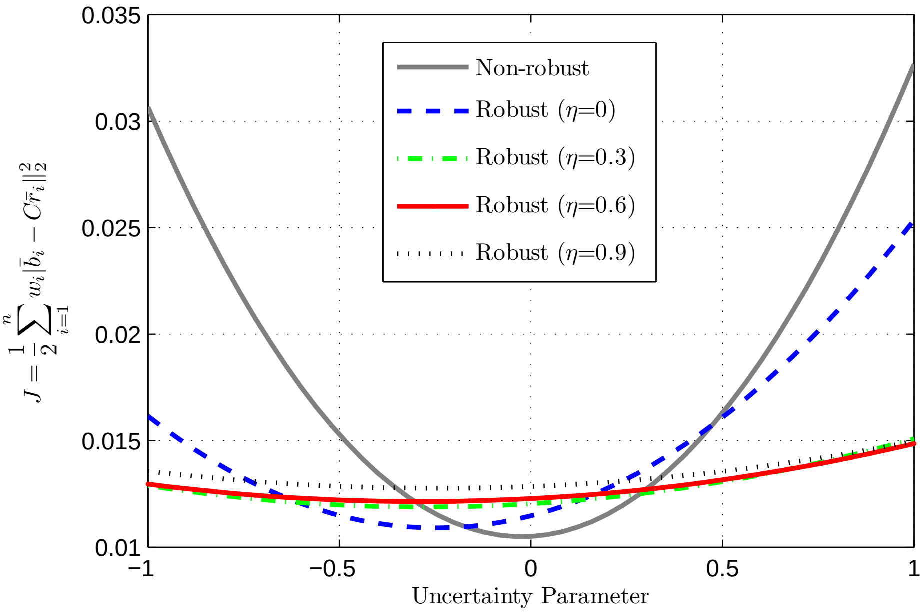

Fig. 3. Effect of the tuning parameter η on robust performance.

from the solution of the relaxed problem. Further, we also showed that the gap between the formulation (21) and its relaxation (23) is zero for $h = 1$ case, showing that the extracted quaternion is the solution to the nonlinear optimization problem (21). The simulation results showed that the robust approach has significant benefit over the nominal approach when inputs have bounded uncertainty. The benefit is maximum in the worst case scenarios, but at the cost of an increased residual for nominal cases.

Some issues need further investigation. Firstly, the relaxation gap for the case when $h > 1$ needs to be explored.

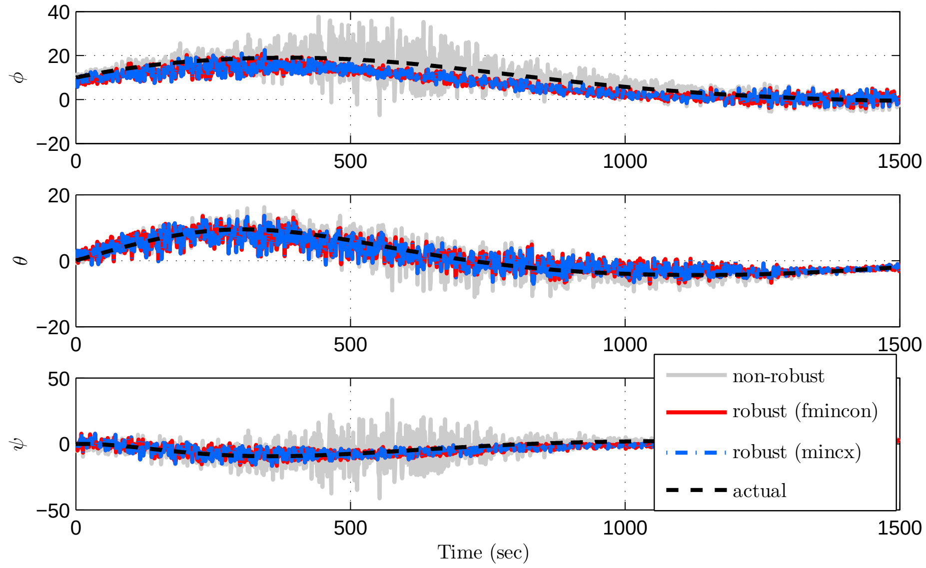

Fig. 4. A comparison of the attitude angles obtained using the robust and non-robust algorithms. The dotted line shows the original data without errors while the other two cases include errors within the chosen uncertainty bound.

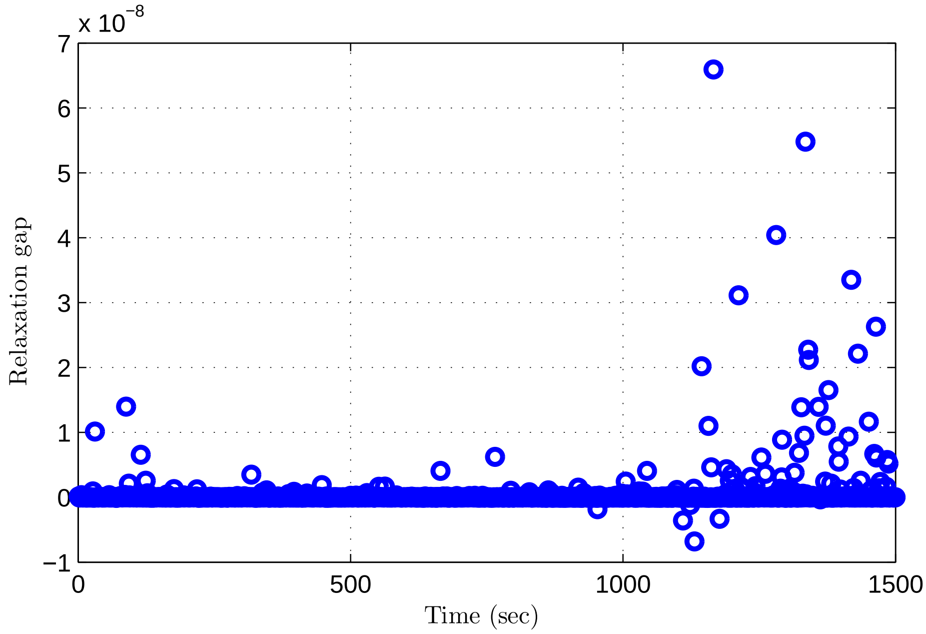

Fig. 5. Gap between the maximum cost of (21) and its upper bound $\bar{\gamma}^*$ obtained using the solution of (23).

Secondly, a procedure to find the optimal tuning parameter η in the regularization term need to be determined. Lastly, a computational complexity analysis of the proposed algorithm is needed to quantify the suitability for an online application.

Appendix A
Davenport Transformation

To derive the Davenport transformation, consider the cost function given in (2) i.e. $\mathrm{tr}(WB^TCR)$. We use two properties of the trace. Firstly trace is invariant under cyclic permutations, and secondly, $\mathrm{tr}\left(\sum_i A_i\right) = \sum_i \mathrm{tr}(A_i) \forall A \in$

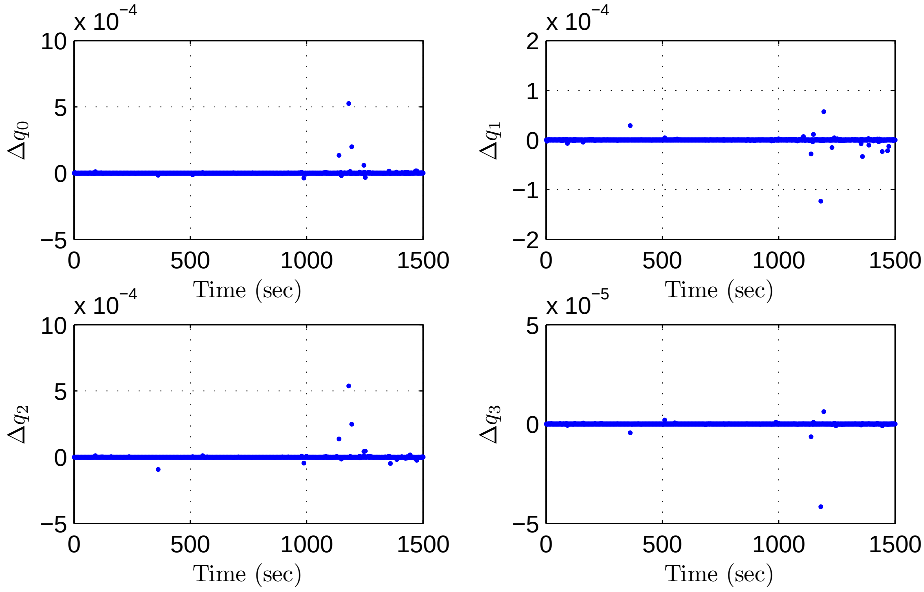

Fig. 6. Error between the quaternions obtained by solving (21) and its SDR.

$\mathbb{R}^{n\times n}$. Using these properties we can write

$$\mathrm{tr}(WB^TCR) = \mathrm{tr}(CRWB^T) = \mathrm{tr}(C\mathbf{B}^T(B,R)), \tag{29}$$

where $\mathbf{B}^T(B,R) = (BWR^T)^T = RWB^T$. Now we represent C using the quaternion $q := \begin{bmatrix}\mathbf{q}^T & q_4\end{bmatrix}^T$, where $\mathbf{q} := \begin{bmatrix}q_1 & q_2 & q_3\end{bmatrix}^T$, written as [1]

$$C = (q_4^2 - \mathbf{q}^T\mathbf{q})I + 2\mathbf{q}\mathbf{q}^T + 2q_4Q. \tag{30}$$

Here Q is the skew symmetric matrix, given as

$$Q = \begin{bmatrix} 0 & q_3 & -q_2 \\ -q_3 & 0 & q_1 \\ q_2 & -q_1 & 0 \end{bmatrix}. \tag{31}$$

Substituting (30) in (29), we get

$$\mathrm{tr}(WB^TCR) = (q_4^2 - \mathbf{q}^T\mathbf{q})\,\mathrm{tr}(\mathbf{B}^T(B,R)) + 2\,\mathrm{tr}(\mathbf{q}\mathbf{q}^T\mathbf{B}^T(B,R)) + 2q_4\mathrm{tr}(Q\mathbf{B}^T(B,R)). \tag{32}$$

Here, the second right hand side term can be written as

$$2\,\mathrm{tr}(\mathbf{q}\mathbf{q}^T\mathbf{B}^T(B,R)) = \mathbf{q}^T(\mathbf{B}^T(B,R) + \mathbf{B}(B,R))\mathbf{q}. \tag{33}$$

The last term can be written as

$$2q_4\,\mathrm{tr}(Q\mathbf{B}^T(B,R)) = q_4(\mathbf{q}^T\mathbf{z}(B,R) + \mathbf{z}^T(B,R)\mathbf{q}), \tag{34}$$

where $\mathbf{z}^T(B,R) = (B\times R)W$. Substituting (33) and (34) in (32) and dropping arguments of $\mathbf{B}(B,R)$ and $\mathbf{z}(B,R)$ for simplification, we get

$$\mathrm{tr}(WB^TCR) = \begin{bmatrix}\mathbf{q}\\ q_4\end{bmatrix}^T \begin{bmatrix}\mathbf{B}^T+\mathbf{B}-\mathrm{tr}(\mathbf{B})I & \mathbf{z}\\ \mathbf{z}^T & \mathrm{tr}(\mathbf{B})\end{bmatrix}\begin{bmatrix}\mathbf{q}\\ q_4\end{bmatrix} = q^T\mathbf{K}(B,R)q, \tag{35}$$

which is the required form, where $\mathbf{K}(B,R)$ is a symmetric, indefinite and traceless matrix defined in (4).

Appendix B

Definition of few Matrices

$$K^1_{r_i} = \begin{bmatrix} r_{i1} & r_{i2} & r_{i3} & 0\\ r_{i2} & -r_{i1} & 0 & -r_{i3}\\ r_{i3} & 0 & -r_{i1} & r_{i2}\\ 0 & -r_{i3} & r_{i2} & r_{i1}\end{bmatrix}, K^2_{r_i} = \begin{bmatrix} -r_{i2} & r_{i1} & 0 & r_{i3}\\ r_{i1} & r_{i2} & r_{i3} & 0\\ 0 & r_{i3} & -r_{i2} & -r_{i1}\\ r_{i3} & 0 & -r_{i1} & r_{i2}\end{bmatrix}, K^3_{r_i} = \begin{bmatrix} -r_{i3} & 0 & r_{i1} & -r_{i2}\\ 0 & -r_{i3} & r_{i2} & r_{i1}\\ r_{i1} & r_{i2} & r_{i3} & 0\\ -r_{i2} & r_{i1} & 0 & r_{i3}\end{bmatrix},$$

$$K^1_{b_i} = \begin{bmatrix} b_{i1} & b_{i2} & b_{i3} & 0\\ b_{i2} & -b_{i1} & 0 & r_{i3}\\ b_{i3} & 0 & -b_{i1} & -b_{i2}\\ 0 & b_{i3} & -b_{i2} & b_{i1}\end{bmatrix}, K^2_{b_i} = \begin{bmatrix} -b_{i2} & b_{i1} & 0 & -b_{i3}\\ b_{i1} & b_{i2} & b_{i3} & 0\\ 0 & b_{i3} & -r_{i2} & r_{i1}\\ -b_{i3} & 0 & b_{i1} & b_{i2}\end{bmatrix}, K^3_{b_i} = \begin{bmatrix} -b_{i3} & 0 & b_{i1} & b_{i2}\\ 0 & -b_{i3} & b_{i2} & -b_{i1}\\ b_{i1} & b_{i2} & b_{i3} & 0\\ b_{i2} & -b_{i1} & 0 & b_{i3}\end{bmatrix}.$$

References

[1] M. D. Shuster and S. Oh, "Three-axis attitude determination from vector observations," *J. Guid. Control*, vol. 4, no. 1, pp. 70–77, 1981.

[2] J. L. Crassidis, F. L. Markley, and Y. Cheng, "Survey of nonlinear attitude estimation methods," *J. Guid. Contr. Dynam.*, vol. 30, no. 1, pp. 12–28, 2007.

[3] D. Murdin, "Standby attitude estimation," in *Position Location and Navigation Symposium*, ser. IEEE, no. 0-7803-5872-4, San Diego, CA , USA, 13-16 March 2000, pp. 429–436.

[4] G. Wahba, "A least square estimate of spacecraft attitude," *SIAM Review*, vol. 7, no. 3, p. 409, 1965.

[5] P. A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. USA: Princeton University Press, 2008.

[6] T. Viklands, "Algorithms for the weighted orthogonal procrustes problem and other least squares problems," Ph.D. dissertation, Department of Computing Science, Umea University, Sweden, 2006.

[7] M. J. Sidi, *Spacecraft Dynamics and Control*, 2nd ed. UK: Cambridge University Press, 2000.

[8] F. L. Markley, "Attitude determination using vector observations and the singular value decomposition," *Journal of Astronautical Sciences*, vol. 36, no. 3, pp. 245–258, 1988.

[9] D. Mortari, "ESOQ: A closed form solution to the Wahba problem," *Journal of the Astronautical Sciences*, vol. 45, no. 2, pp. 195–204, 1997.

[10] ——, "Second estimator of the optimal quaternion," *J. Guid. Contr. Dynam.*, vol. 23, no. 5, pp. 885–888, 2000.

[11] F. L. Markley, "Optimal attitude matrix from two vector measurements," *J. Guid. Contr. Dynam.*, vol. 31, no. 3, pp. 765–768, 2008.

[12] J. E. Keat, "Analysis of least-squares attitude determination routine DOAOP," Tech. Rep. NASA/CSC/TM-77/6034, 1977.

[13] S. Chandrasekaran, G. H. Golub, M. Gu, and A. H. Sayed, "Parameter estimation in the presence of bounded data uncertainties," *IEEE Signal Processing Letters*, vol. 4, no. 7, 1997.

[14] Y. C. Eldar, A. Ben-Tal, and A. Nemirovski, "Robust mean-squared error estimation in the presence of model uncertainties," *IEEE Transactions on Signal Processing*, vol. 53, no. 1, pp. 168–181, 2005.

[15] M. Pilanci, O. Arikan, and M. C. Pinar, "Structured least squares problems and robust estimators," *IEEE Transactions on Signal Processing*, vol. 58, no. 5, pp. 2453–2465, 2010.

[16] I. Bar-Itzhack and Y. Oshman, "Attitude determination from vector observations: Quaternion estimation," *IEEE Transaction on Aerospace and Electronic Systems*, vol. AES-21, no. 1, pp. 128–136, 1985.

[17] D. Chourkroun, H. Weiss, I. Bar-Itzhack, and Y. Oshman, "Direction cosine matrix estimation from vector observations using a matrix Kalman filter," *IEEE Transaction on Aerospace and Electronic Systems*, vol. 46, no. 1, pp. 61–79, 2010.

[18] M. D. Shuster, "The generalized Wahba problem," *Journal of Astronautical Sciences*, vol. 54, no. 2, pp. 245–259, 2006.

[19] D. Mortari and M. Majji, "Multiplicative measurement model and single-point attitude estimate," *AAS Journal*, no. 263, pp. 51–69, 2008.

[20] S. McLean, S. Macmillan, S. Maus, V. Lesur, A. Thomson, and D. Dater, "The US/UK world magnetic model for 2005-2010," Tech. Rep. NOAA Technical Report NESDIS/NGDC-1, December, 2004.

[21] C. M. Roithmayr, "Contributions of spherical harmonics to magnetic and gravitational fields," Tech. Rep. NASA/TM-2004-213007, 2004.

[22] S. Ahmed and E. C. Kerrigan, "Robust static attitude determination using robust optimization," in *IFAC World Conference*, Milano, Italy, August 2011.

[23] S. Ahmed, E. C. Kerrigan, and I. M. Jaimoukha, "Semidefinite relaxation of a robust static attitude determination problem," in *IEEE Conference on Control and Decision*, Orlando, Florida, USA, December 2011.

[24] S. A. M. Y. Y. Luo Z.Q., Ma W.K. and Z. S., "Semidefinite relaxation of quadratic optimization problems," *SP Magazine, Special Issue on Convex Opt. for SP*, pp. 505–515, May 2010.

[25] A. Ben-Tal, L. E. Ghaoui, and A. Nemirovski, *Robust Optimization*. USA: Princeton University Press, 2009.

[26] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge Univ. Press, 2004.

[27] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. USA: The Johns Hopkins University Press, 1996.

[28] U. Malik, I. M. Jaimoukha, G. D. Halikias, and S. K. Gungah, "On the gap between the quadratic inter programming problem and its semidefinite relaxation," *Mathematical Programming*, vol. 107, pp. 505–515, 2006.

[29] H. Heidt, J. P. Suari, A. S. Moore, S. Nakasuka, and R. J. Twiggs, "Cubesat: A new generation of pico-satellite for education and industry low cost space experimentation," in *In 15th Anual/USU Conference on Small Satellite*. SSC01-VIIIb-5, 2001.

[30] J. R. Wertz, *Spacecraft Attitude Determination and Control*. Holland: D. Reidel Publishing Company, 1978.

Shakil Ahmed received the B.Sc. degree in electrical engineering from UET Lahore, Pakistan, in 1996. From 1996 to 1998, he worked in Siemens Pakistan. Since 1998 he has been with the government of Pakistan. He received the M.Sc. degree in electrical engineering from UET Taxila, Pakistan, in 2002 and in Aerospace Dynamics from Cranfield University, U.K., in 2004. Currently, he is pursuing his PhD at Imperial College London. His research interests include robust estimation and control of dynamical systems, robust optimization and its applications.

Eric C. Kerrigan received the B.Sc.(Eng) degree in electrical engineering from the University of Cape Town, South Africa, in 1996, and the Ph.D. degree from the University of Cambridge, in 2001. From 2001 to 2005, he was a Research Fellow at the Department of Engineering, University of Cambridge. He is currently a Senior Lecturer at Imperial College London. His research interests include optimal and robust control of constrained systems, predictive control, optimization and its applications.

Imad M. Jaimoukha received his BSc in Electrical Engineering from the University of Southampton, England in 1983, an MSc and a PhD in Control Systems from Imperial College London in 1986 and 1990. He worked as a Research Fellow at the Center for Process Systems Engineering at Imperial College London from 1990 till 1994. Since 1994 he has been with the Department of Electrical and Electronic Engineering, Imperial College London. His research interests include robust control, system approximation, and global optimization.