Journal of Guidance, Control, and Dynamics
Vol. 39, No. 1, January 2016

Convex Solution to a Joint Attitude and Spin-Rate Estimation Problem

James Saunderson,* Pablo A. Parrilo,† and Alan S. Willsky‡
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

DOI: 10.2514/1.G001107

This paper considers the problem of jointly estimating the attitude and spin rate of a spinning spacecraft. Psiaki has formulated a family of optimization problems that generalize the classical least-squares attitude-estimation problem, known as Wahba's problem, to the case of a spinning spacecraft. In the special case where the rotation axis and the fundamental sampling period are fixed and known, but the spin rate is unknown (such as for nutation-damped spin-stabilized spacecraft), Psiaki's problem can be reformulated exactly as a type of tractable convex optimization problem called a semidefinite optimization problem. This reformulation allows the problem to be globally solved using standard numerical routines for semidefinite optimization.

I. Introduction

Spacecraft§ attitude estimation is a fundamental problem, arising, for instance, as a natural subproblem whenever attitude control is required. Because spacecraft dynamics are nonlinear, a typical and successful approach to attitude estimation is to employ variants of the extended Kalman filter (EKF) [1]. As with any method based on linearization of nonlinear dynamics, EKF-based approaches can fail to converge given poor initial estimates and can become unstable in the presence of large disturbances [2]. Many truly nonlinear attitude-estimation methods have also been proposed (see [3] for a survey). An important example is the static least-squares attitude-estimation problem known as Wahba's problem [4]. In Wahba's problem, we are simultaneously given a batch of vector measurements (from sun sensors, star trackers, etc.) in the body frame and corresponding reference directions in an inertial frame. The aim is to find the rotation matrix (i.e., direction cosine matrix) that minimizes the sum of the squared errors between the transformed reference directions and the observed vector measurements. Wahba's problem, as stated, applies most naturally to a static spacecraft. Nevertheless, it has also found use as a subroutine in various recursive estimation algorithms including those that estimate the full dynamical state of the spacecraft (see, e.g., [2,5]).

Recently, Psiaki has posed a number of generalizations of Wahba's problem to the case of a spinning spacecraft [6]. These problems aim to simultaneously estimate the initial attitude and spin rate (or, more generally, initial angular momentum) of the spacecraft from vector measurements, without the need for gyroscope measurements. These generalizations are particularly suited to spin-stabilized spacecraft without gyroscopes. We describe Wahba's problem and Psiaki's generalizations formally in Sec. III.

In this paper, we focus on the simplest of Psiaki's generalizations of Wahba's problem. We refer to this problem as Psiaki's first problem. In this problem, we assume the spacecraft is spinning at a constant unknown angular velocity around a known (stable) inertia axis. This setting is relevant for nutation-damped spin-stabilized spacecraft [6]. The aim is to estimate the initial attitude and the unknown spin rate given a sequence of noisy vector measurements obtained at equispaced sampling instants, together with corresponding reference directions. Wahba's problem arises as the special case where the spin rate is zero.

Our main contribution is to show that, when the sampling period is constant, Psiaki's first problem can be reformulated exactly as a semidefinite optimization problem (see Theorem 3). Semidefinite optimization problems (described in Sec. IV) are a family of convex optimization problems that generalize linear programming and can be solved globally with provable efficiency guarantees using standard software. Reformulating Psiaki's first problem as a semidefinite optimization problem means that it, like Wahba's problem, can be solved efficiently and globally, to high precision, using numerical methods.

The remainder of the paper is organized as follows. In Sec. II, we summarize notation not defined elsewhere in the paper. In Sec. III, we first describe Psiaki's generalizations of Wahba's problem for spinning spacecraft. We then show how to write Psiaki's first problem as an instance of a family of problems we call *trigonometric Wahba problems*; see Eq. (6). We conclude the section with a summary of prior work on Psiaki's problems. In Sec. IV, we briefly describe semidefinite optimization problems in general before presenting our semidefinite optimization-based reformulation of trigonometric Wahba problems, and in particular of Psiaki's first problem. We defer the proofs to the Appendix. In Sec. V, we describe the results of a simple numerical experiment solving Psiaki's first problem using different numbers of measurements at different noise levels. In Sec. VI, we discuss possible future research related to the work in this paper.

II. Notation

We briefly summarize notation used throughout the body of the paper. Additional notation that is used only in the Appendix is introduced separately there.

A. Spaces

Denote by $\mathbb{R}^{n\times n}$ the space of $n\times n$ real matrices. If $X\in\mathbb{R}^{n\times n}$, let X^T be its transpose. Let $\mathcal{S}^n$ be the space of $n\times n$ symmetric matrices (i.e., matrices for which $X=X^T$). Let $\mathcal{S}^n_+$ denote the set of $n\times n$ symmetric positive semidefinite matrices (i.e., $X\in\mathcal{S}^n_+$ if and only if $u^TXu\geq 0$ for all $u\in\mathbb{R}^n$). If $X\in\mathcal{S}^n_+$, we write $X\succcurlyeq 0$ when the dimension is clear from the context.

B. Inner Products

If $x,y\in\mathbb{R}^n$, then $\langle x,y\rangle=\sum_{i=1}^n x_iy_i$. If $x\in\mathbb{R}^n$, then define $\|x\|=\langle x,x\rangle^{1/2}=(\sum_{i=1}^n x_i^2)^{1/2}$ to be the usual Euclidean norm. If $X,Y\in\mathbb{R}^{n\times n}$, then define an inner product by $\langle X,Y\rangle=\mathrm{tr}(X^TY)=\sum_{i,j=1}^n X_{ij}Y_{ij}$.

Received 10 October 2014; revision received 4 April 2015; accepted for publication 10 April 2015; published online 3 July 2015.

*Graduate Student, Department of Electrical Engineering and Computer Science; jamess@mit.edu.

†Professor, Department of Electrical Engineering and Computer Science; parrilo@mit.edu.

‡Professor, Department of Electrical Engineering and Computer Science; willsky@mit.edu.

§A prelimiary version of this work appears as part of [31].

C. Convex Hull

Given a subset $S \subset \mathbb{R}^n$ (which need not be finite), the convex hull of S is the set of all possible convex combinations of elements of S. Formally,

$$\mathrm{conv}(S) = \Big\{ z \in \mathbb{R}^n\colon \text{there exist a positive interger } r, \text{ points } x_1, x_2, \ldots, x_r \in S, \text{ and scalars } \lambda_1, \lambda_2, \ldots, \lambda_r \geq 0 \text{ such that } \sum_{i=1}^{r} \lambda_i = 1 \text{ and } z = \sum_{i=1}^{r} \lambda_i x_i \Big\}$$

D. Block Matrices

If $T_0, T_1, \ldots, T_N$ are $d \times d$ matrices, with T_0 being symmetric, define the corresponding $d(N+1) \times d(N+1)$ symmetric block Toeplitz matrix by

$$\mathrm{Toeplitz}(T_0, T_1, \ldots, T_N) = \begin{bmatrix} T_0 & T_1 & T_2 & \cdots & T_N \\ T_1^T & T_0 & T_1 & \ddots & \vdots \\ T_2^T & T_1^T & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & T_1 \\ T_N^T & \cdots & \cdots & T_1^T & T_0 \end{bmatrix} \tag{1}$$

Similarly, if $S_1, S_2, \ldots, S_{2N+1}$ are symmetric $d \times d$ matrices, define the corresponding $d(N+1) \times d(N+1)$ block Hankel matrix by

$$\mathrm{Hankel}(S_1, S_2, \ldots, S_{2N+1}) = \begin{bmatrix} S_1 & S_2 & \cdots & S_N & S_{N+1} \\ S_2 & & & S_{N+1} & S_{N+2} \\ \vdots & & \iddots & & \vdots \\ S_N & S_{N+1} & & & S_{2N} \\ S_{N+1} & S_{N+2} & \cdots & S_{2N} & S_{2N+1} \end{bmatrix} \tag{2}$$

E. Unit Quaternion Parameterization of Rotations

Let $SO(3)$ denote the set of rotation (or direction-cosine) matrices. Let $\mathbb{H}$ denote the unit quaternions. Throughout, we think of the unit quaternions geometrically as the unit sphere in $\mathbb{R}^4$ (i.e., $\mathbb{H} = \{q \in \mathbb{R}^4 \colon \|q\| = 1\}$). Any element of $SO(3)$ can be expressed as $\mathcal{A}(qq^T)$, where $q \in \mathbb{H}$, and $\mathcal{A}\colon \mathcal{S}^4 \to \mathbb{R}^{3\times3}$ is the linear map defined (following the convention in [3]) by

$$\mathcal{A}(Z) \coloneqq \begin{bmatrix} Z_{11} - Z_{22} - Z_{33} + Z_{44} & 2Z_{12} + 2Z_{34} & 2Z_{13} - 2Z_{24} \\ 2Z_{12} - 2Z_{34} & -Z_{11} + Z_{22} - Z_{33} + Z_{44} & 2Z_{23} + 2Z_{14} \\ 2Z_{13} + 2Z_{24} & 2Z_{23} - 2Z_{14} & -Z_{11} - Z_{22} + Z_{33} + Z_{44} \end{bmatrix} \tag{3}$$

The adjoint of $\mathcal{A}$ (with respect to the inner product on matrices) is $A^*\colon \mathbb{R}^{3\times3} \to \mathcal{S}^4$ defined by

$$\mathcal{A}^*(Y) \coloneqq \begin{bmatrix} Y_{11} - Y_{22} - Y_{33} & Y_{12} + Y_{21} & Y_{13} + Y_{31} & Y_{23} - Y_{32} \\ Y_{12} + Y_{21} & -Y_{11} + Y_{22} - Y_{33} & Y_{23} + Y_{32} & -Y_{13} + Y_{31} \\ Y_{13} + Y_{31} & Y_{23} + Y_{32} & -Y_{11} - Y_{22} + Y_{33} & Y_{12} - Y_{21} \\ Y_{23} - Y_{32} & -Y_{13} + Y_{31} & Y_{12} - Y_{21} & Y_{11} + Y_{22} + Y_{33} \end{bmatrix} \tag{4}$$

In other words, for any $Z \in \mathcal{S}^4$ and any $Y \in \mathbb{R}^{3\times3}$, we have the identity

$$\langle \mathcal{A}(Z), Y \rangle = \langle Z, \mathcal{A}^*(Y) \rangle \tag{5}$$

III. Psiaki's Generalizations of Wahba's Problem for Spinning Spacecraft

In this section, we describe Wahba's problem [4] and Psiaki's generalizations to the case of a spinning spacecraft [6]. For reasons discussed in Sec. III.B, we subsequently focus on the simplest of Psiaki's problems: jointly estimating the attitude and spin rate of a spacecraft spinning around a stable inertia axis at a constant unknown rate. In this case, we show how to reformulate the resulting optimization problem in the general form

$$\max_{\substack{Q \in SO(3) \\ \omega' \in [-\pi, \pi)}} \langle A_0, Q \rangle + \sum_{n=1}^{N} [\langle A_n, \cos(\omega' n) Q \rangle + \langle B_n, \sin(\omega' n) Q \rangle] \tag{6}$$

for appropriate collections of 3×3 matrices $(A_n)_{n=0}^N$ and $(B_n)_{n=1}^N$. Throughout, we call problems in the form of Eq. (6) trigonometric Wahba problems. In Sec. IV, we show how to reformulate trigonometric Wahba problems as semidefinite optimization problems.

Our reformulation only works when each of the quantities n in the expressions $\cos(\omega' n)$ and $\sin(\omega' n)$ are integers, as is the case in Eq. (6). Essentially, this is because $\cos(\omega' n)$ and $\sin(\omega' n)$ can be expressed as polynomials in $\cos(\omega')$ and $\sin(\omega')$ only in these cases. This is the basic source of the restriction that the sampling period be constant, mentioned in the Abstract, in Sec. I, and later in this section.

A. Wahba's Problem

We briefly describe Wahba's least-squares attitude-estimation problem posed in [4]. Initial solutions were published one year later in [7]. Subsequently, Davenport's q method was developed (see, e.g., [8]) and has become the standard solution to Wahba's problem.

1. *Vector Measurements*

Suppose we are given a batch of noisy unit vector measurements $y_0, y_1, \ldots, y_N$ in the body frame (obtained from star trackers, sun sensors, magnetometers, etc.) of corresponding unit reference directions $x_0, x_1, \ldots, x_N$ in the inertial frame.

2. *Least-Squares Objective*

Wahba's problem is to find the rotation matrix $Q \in SO(3)$ that transforms the reference directions to best fit the measured vector measurements in the weighted least-squares sense by solving

$$\min_{Q\in SO(3)} \sum_{n=0}^{N} \frac{\kappa_n}{2} \|y_n - Qx_n\|^2 \tag{7}$$

where $\kappa_0, \kappa_1, \ldots, \kappa_N$ are nonnegative scalar weights that one would take to be larger for measurements with smaller noise variance. Because $\|Qx\|^2 = \|x\|^2$ for all $x \in \mathbb{R}^3$, we can expand the squares and see that this optimization problem is equivalent to

$$\max_{Q\in SO(3)} \left\langle \sum_{n=0}^{N} \kappa_n y_n x_n^T, Q \right\rangle \tag{8}$$

where we have dropped an additive constant of

$$\sum_{n=0}^{N} \frac{\kappa_n}{2} (\|y_n\|^2 + \|x_n\|^2)$$

B. Psiaki's Generalizations

We now describe Psiaki's generalizations of Wahba's problem and show how Wahba's problem arises as a special case.

1. *Rigid-Body (Euler) Equations*

Let $Q(t_0) \in SO(3)$ denote the initial attitude of the spacecraft, $\Omega(t_0) \in \mathbb{R}^3$ the initial body angular velocity, and $I_1 \geq I_2 \geq I_3$ the principal moments of inertia. Assuming that the spacecraft undergoes torque-free motion about its center of mass, then for $t \geq t_0$, the attitude $Q(t)$ and the body angular velocity $\Omega(t) := [\,\omega_1(t) \quad \omega_2(t) \quad \omega_3(t)\,]^T$ satisfy the rigid-body equations:

$$\begin{aligned} I_1\dot{\omega}_1(t) &= (I_2 - I_3)\omega_2(t)\omega_3(t) \\ I_2\dot{\omega}_2(t) &= (I_3 - I_1)\omega_3(t)\omega_1(t) \quad \text{and} \\ I_3\dot{\omega}_3(t) &= (I_1 - I_2)\omega_1(t)\omega_2(t) \\ \dot{Q}(t) &= \begin{bmatrix} 0 & \omega_3(t) & -\omega_2(t) \\ -\omega_3(t) & 0 & \omega_1(t) \\ \omega_2(t) & -\omega_1(t) & 0 \end{bmatrix} Q(t) \end{aligned} \tag{9}$$

Note that, for every $t \geq t_0$ and every $\Omega(t_0)$, we have that $Q(t) = \Phi(t - t_0; \Omega(t_0))Q(t_0)$ for some map Φ taking values in $SO(3)$. In particular, $Q(t)$ is always linear in the initial attitude $Q(t_0)$.

2. *Vector Measurements*

Let $t_0, t_1, \ldots, t_N$ be a finite set of sampling instants. Assume at sample instant t_n that we are given a noisy unit vector measurement y_n in the spacecraft body frame of a corresponding reference direction x_n in the inertial frame.

3. *Least-Squares Objective*

Following Wahba's least-squares-based objective, Psiaki suggests solving the following weighted least-squares problem to estimate the initial attitude and body angular velocity of the spacecraft, given only the vector measurements $(y_n)_{n=0}^N$ and the reference directions $(x_n)_{n=0}^N$:

$$\min_{Q(t_0),\Omega(t_0)} \sum_{n=0}^{N} \frac{\kappa_n}{2} \|y_n - Q(t_n)x_n\|_2^2 \tag{10}$$

subject to $Q(t)$ satisfying Eq. (9) with initial conditions $Q(t_0)$ and $\Omega(t_0)$. Just as for Wahba's problem, the κ_n are nonnegative scalars.

4. *Dependence on* $\Omega(t_0)$

In general, the dependence of $Q(t)$ on the initial body angular velocity $\Omega(t_0)$ is quite complicated. The relationship between $Q(t)$ and $\Omega(t_0)$ simplifies under additional assumptions on $\Omega(t_0)$ and the inertia tensor of the spacecraft. We now summarize these simplified problems and name them for later reference.

Wahba's problem: If $\Omega(t_0) = 0$, then $Q(t) = Q(t_0)$ for all $t \geq t_0$, and so the spacecraft is stationary. Adding this as a constraint, we recover Wahba's original formulation [Eq. (7)].

Psiaki's first problem: Suppose $\Omega(t_0)$ is aligned with the major inertia axis, and (without loss of generality) this is the first axis direction in body coordinates. Then, $\Omega(t_0) = [\,\omega \quad 0 \quad 0\,]^T$, and so the dynamical constraints [Eq. (9)] reduce to

$$Q(t) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\omega t) & \sin(\omega t) \\ 0 & -\sin(\omega t) & \cos(\omega t) \end{bmatrix} Q(t_0) \tag{11}$$

where ω is the spin rate (in radians per second). In this case, the spacecraft is spinning with an unknown constant angular velocity ω around a known axis (fixed in body coordinates). Minimizing the least-squares objective [Eq. (10)] subject to the constraints [Eq. (11)] is the first generalization of Wahba's problem posed in [6] and is relevant for a nutation damped spin-stabilized spacecraft.

Psiaki's second problem: If $\Omega(t_0)$ is unconstrained and no additional assumptions are made about the moments of inertia of the spacecraft, we obtain the second generalization of Wahba's problem posed in [6]. In this setting, the dependence of $Q(t)$ on $\Omega(t_0)$ is more complicated. This case is discussed further in [9] (see Sec. III.C).

In each case, Psiaki's formulations involve solving nonconvex optimization problems of the form in Eq. (10) subject to dynamical constraints.

5. *Focus of the Paper*

For the remainder of the paper, we focus on Psiaki's first problem because, in this case, $Q(t)$ only depends on the initial body angular velocity through $\cos(\omega t)$ and $\sin(\omega t)$. In addition to focusing on Psiaki's first problem, we also assume that the sampling instants $t_0, t_1, \ldots, t_N$ are equally spaced. As such, we assume that there is some τ such that $t_n = n\tau$ for $n = 0, 1, \ldots, N$.

This paper does not address Psiaki's more general second problem, where the dependence of $Q(t)$ on $\Omega(t_0)$ is significantly more complicated. It would be very interesting if the techniques we develop can be extended to this more general situation.

6. *Aliasing*

Because we only observe ω via vector measurements at time instants that are integer multiples of τ, from the data alone we cannot distinguish between spin rates at different integer multiples of $2\pi/\tau$ due to aliasing. Hence, we assume that $\omega \in [-\pi/\tau, \pi/\tau)$, so that it is possible to determine the unknown spin rate from the data. (We could, alternatively, fix some a in radians per second and assume $\omega \in [a, a + 2\pi/\tau)$.) In a Bayesian formulation of the problem, we could interpret this as encoding prior information on the spin rate.

7. *Reformulation*

We now reformulate Psiaki's first problem as a trigonometric Wahba problem. Because $\|Q(t)x_n\|^2 = \|x_n\|^2$ for all t and n, observe that with $t_n = n\tau$ the optimization problem [Eq. (10)] can be rewritten as

$$\min_{\substack{Q(0)\in SO(3) \\ \omega\in[-\pi/\tau,\pi/\tau)}} \sum_{n=0}^{N} \frac{\kappa_n}{2} [\|y_n\|^2 - 2\langle y_n, Q(n\tau)x_n\rangle + \|x_n\|^2] \tag{12}$$

$$\text{s.t. } Q(n\tau) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(n\tau\omega) & \sin(n\tau\omega) \\ 0 & -\sin(n\tau\omega) & \cos(n\tau\omega) \end{bmatrix} Q(0) \tag{13}$$

Putting $\omega' = \tau\omega$, we see that this is equivalent, as an optimization problem, to

$$\max_{\substack{Q\in SO(3)\\ \omega'\in[-\pi,\pi)}} \langle A_0, Q\rangle + \sum_{n=1}^{N}[\langle A_n, \cos(n\omega')Q\rangle + \langle B_n, \sin(n\omega')Q\rangle] \quad (14)$$

where

$$A_0 = \kappa_0 y_0 x_0^T + \begin{bmatrix} 1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\left(\sum_{n=1}^{N}\kappa_n y_n x_n^T\right) \quad (15)$$

and for $n = 1, 2, \ldots, N$,

$$A_n = \kappa_n\begin{bmatrix} 0 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} y_n x_n^T \quad \text{and} \quad B_n = \kappa_n\begin{bmatrix} 0 & 0 & 0\\ 0 & 0 & 1\\ 0 & -1 & 0\end{bmatrix} y_n x_n^T \quad (16)$$

We have now expressed Psiaki's first problem in the general form described in Eq. (6).

C. Prior Work and Alternative Solution Methods for Psiaki's Problems

In this section, we summarize previous approaches to Psiaki's generalizations of Wahba's problem for spinning spacecraft. We then briefly discuss a simple discretization-based approach, implicit in the work of Psiaki and Hinks [9], for solving Psiaki's problems globally.

Psiaki's original paper [6] describes a method to globally solve Psiaki's first problem using two noise-free vector measurements (sampled at distinct times) and information from a third measurement to resolve a sign ambiguity. In this situation, the problem reduces to finding all the solutions of the corresponding nonlinear equations satisfied by the initial attitude and spin rate. This method seems sensitive to measurement noise and is unable to exploit additional measurements to mitigate the effects of noise. (The possible advantages of incorporating multiple measurements are considered in Sec. V.)

In subsequent work [10], Hinks and Psiaki describe an approach to Psiaki's second problem under the assumption that the spacecraft is axially symmetric and exactly three noise-free vector measurements are used. In this case, it is again possible to find an initial body angular velocity $\Omega(t_0)$ and an initial attitude that are consistent with the measurements by solving a set of nonlinear equations. They suggest different formulations of these equations and apply Newton's method (with possibly many different initializations) to obtain a solution to the equations. Again, this approach is likely to be useful when there is very little noise.

In later work [9], Psiaki and Hinks describe a method to find local optima of Psiaki's first and second problems (with no additional assumptions) by a novel optimization scheme. The procedure has many details; we just sketch the main ideas here. The first main idea is that, for fixed $\Omega(t_0)$, each point $Q(t_0), Q(t_1), \ldots, Q(t_N)$ on the trajectory is linear in $Q(t_0)$. Hence, if we can compute the trajectory $(Q(t_n))_{n=0}^N$ for fixed $\Omega(t_0)$, we can minimize the objective function of Eq. (10) over $Q(t_0)$ for fixed $\Omega(t_0)$ by solving an instance of Wahba's problem. To obtain the trajectory $(Q(t_n))_{n=0}^N$ for fixed $\Omega(t_0)$, Psiaki and Hinks suggest numerically solving the rigid-body equations. They then employ a trust-region method to locally optimize over $\Omega(t_0)$, taking into account (to second order) the effect that $\Omega(t_0)$ has on the optimal $Q(t_0)$. This approach finds local minima for $\Omega(t_0)$ and $Q(t_0)$. By using many different initializations for $\Omega(t_0)$, it can find many different local minima for $\Omega(t_0)$ and $Q(t_0)$ and return one with the lowest cost. This method makes very few assumptions and can incorporate many measurements and so should behave well in the presence of measurement noise.

A simpler, but much more naive, strategy would be to discretize the space of $\Omega(t_0)$, solve (in parallel) the corresponding instance of Wahba's problem for each value of $\Omega(t_0)$, then output the pair $(\Omega(t_0), Q(t_0))$ with the smallest cost. This is a reasonable strategy for Psiaki's first problem because aliasing issues mean there is always an optimal ω in the interval $[-\pi/\tau, \pi/\tau)$. To find a value of ω that is at most ϵ from the optimal value requires solving $O(1/\epsilon)$ instances of Wahba's problem. This is very efficient unless particularly accurate solutions for ω are required (i.e., ϵ is very small).

IV. Semidefinite Optimization Reformulations

The main aim of this section is to describe how to reformulate trigonometric Wahba problems, and hence Psiaki's first problem (which is a special case), as semidefinite optimization problems. Before doing so, we briefly explain what semidefinite optimization problems are and what we mean by a semidefinite reformulation of an optimization problem. We illustrate this in Sec. IV.B by giving a semidefinite reformulation of Wahba's problem that can be thought of as a more flexible description of the q method [8]. In Sec. IV.C, we give a semidefinite reformulation of trigonometric Wahba problems, before giving in Sec. IV.D pseudocode illustrating how to solve the semidefinite optimization problems we formulate using generic semidefinite optimization solvers. We briefly discuss algorithms for these semidefinite optimization problems, as well as their worst-case complexity, in Sec. IV.E.

A. Semidefinite Optimization

Semidefinite optimization problems are convex optimization problems of the form

$$\max_x \langle c, x\rangle \text{ s.t. } A_0 + \sum_{i=1}^{n} A_i x_i \succcurlyeq 0 \quad (17)$$

where $x \in \mathbb{R}^n$ is a vector of decision variables, $c \in \mathbb{R}^n$ represents a linear cost functional, and the matrices $A_0, A_1, \ldots, A_n$ are symmetric $m \times m$ matrices. Recall that $X \succcurlyeq 0$ means that the symmetric matrix X is positive semidefinite. An expression of the form

$$A(x) = A_0 + \sum_{i=1}^{n} A_i x_i \succcurlyeq 0$$

is often called a linear matrix inequality because it is linear in the decision variable x. The form of the optimization problem in Eq. (17) is a "standard form" for semidefinite optimization problems. Any optimization problem with a linear cost function, linear equality constraints, and linear matrix inequality constraints can be reformulated in this standard form in a completely algorithmic way (see, e.g., [11]). Modern optimization parsers usually take care of such transformations automatically [12].

Semidefinite optimization problems can be solved to any desired accuracy in time polynomial in n and m using standard software based on interior point methods [11]. The semidefinite optimization problems that arise in this paper have additional structure that can be exploited to obtain even more efficient algorithms. (See Sec. IV.E for further discussion of this point.) For much more information about semidefinite optimization, including duality theory, numerical algorithms, and applications, see for example [11].

Semidefinite Reformulations. Many different optimization problems arising in a variety of contexts, including some optimization problems for which the natural formulation is not convex, can be reformulated as semidefinite optimization problems. Given an optimization problem, by a semidefinite reformulation we mean a semidefinite optimization problem such that 1) there is an efficient procedure to take the data of the original problem and construct the data of the semidefinite optimization problem; 2) the optimal value of the semidefinite optimization problem and the original optimization problem are the same; and 3) there is an efficient procedure to take an optimal solution to the semidefinite optimization problem and produce an optimal solution to the original optimization problem.

The following basic fact of convex analysis plays an important role in many such reformulations. It is a paraphrasing of [13] (Theorems 32.2 and 18.3).

Theorem 1: Suppose that $c \in \mathbb{R}^n$, and S is compact. Then,

$$\text{cost}_\star := \max_{x \in S} \langle c, x \rangle = \max_{x \in \text{conv}(S)} \langle c, x \rangle \tag{18}$$

i.e., the optimal cost is $\text{cost}_\star$ whether we optimize the linear functional defined by c over S or over its convex hull. Moreover,

$$\text{conv}\{x \in S : \langle c, x \rangle = \text{cost}_\star\} = \{x \in \text{conv}(S) : \langle c, x \rangle = \text{cost}_\star\}$$

i.e., the set of optimal points for the linear optimization problem over $\text{conv}(S)$ is the convex hull of the set of optimal points for the linear optimization problem over S.

B. Wahba's Problem

We illustrate the basic idea of semidefinite reformulations with the example of solving Wahba's problem. We note that there are much better ways to solve Wahba's problem. The advantage of the semidefinite reformulation is that it can be extended to more complicated situations, such as Psiaki's first problem. The reformulation presented in this section appears (in a more general context) in [14] and is generalized to the analogous problem where $SO(3)$ is replaced with $SO(n)$ for any $n \geq 2$ in [15]. (See also [16] where a semidefinite relaxation of Wahba's problem is described, as well as conditions under which it is exact.)

Wahba's problem fits into the general form [Eq. (6)] where $A_0 = \sum_{n=0}^{N} \kappa_n y(n\tau) x(n\tau)^T$ [cf. Eq. (8)], and all the other terms vanish. Using the quaternion parameterization of $SO(3)$, Wahba's problem can be expressed as

$$\max_{Q \in SO(3)} \langle A_0, Q \rangle = \max_{q \in \mathbb{H}} \langle A_0, \mathcal{A}(qq^T) \rangle = \max_{q \in \mathbb{H}} \langle \mathcal{A}^*(A_0), qq^T \rangle \tag{19}$$

We now explain how to reformulate Eq. (19) as a semidefinite optimization problem following a general pattern that we use again in Sec. IV.C.

1) Rewrite the problem as the optimization of a linear functional over some set. In this case,

$$\max_{Z} \langle \mathcal{A}^*(A_0), Z \rangle \text{ s.t. } Z \in \{qq^T : q \in \mathbb{H}\}$$

2) Replace the constraint set with the convex hull of the constraint set. In this case,

$$\max_{Z} \langle \mathcal{A}^*(A_0), Z \rangle \text{ s.t. } Z \in \text{conv}\{qq^T : q \in \mathbb{H}\} \tag{20}$$

This optimization problem has the same optimal value as the original nonconvex problem because the cost function is linear (see Theorem 1).

3) Describe the convex hull of the constraint set as the feasible region of a semidefinite optimization problem (if possible). In this case, such a description is well known (see, e.g., [17], Theorem 3) and given by

$$\text{conv}\{qq^T : q \in \mathbb{H}\} = \{Z \in \mathcal{S}^4 : Z \succcurlyeq 0, \text{tr}(Z) = 1\}$$

(This holds because, if $Z \succcurlyeq 0$ and $\text{tr}(Z) = 1$, then any eigendecomposition

$$Z = \sum_{i=1}^{4} \lambda_i q_i q_i^T$$

expresses Z as a convex combination of matrices of the form qq^T with $\|q\| = 1$.)

The resulting semidefinite reformulation of Wahba's problem is

$$\max_{Z} \langle \mathcal{A}^*(A_0), Z \rangle \text{ s.t. } \text{tr}(Z) = 1, \qquad Z \succcurlyeq 0 \tag{21}$$

1. Extracting an Optimal Point

Under mild assumptions (such as having access to at least two generic vector measurements), Wahba's problem has a unique solution $Q_\star$. There are two corresponding optimal unit quaternions $\pm q_\star$. These both satisfy $Q_\star = \mathcal{A}(q_\star q_\star^T)$. Hence, $q_\star q_\star^T$ is the unique optimum of Eq. (20). By the second statement of Theorem 1, $Z_\star = q_\star q_\star^T$ is the unique optimum of the semidefinite optimization problem [Eq. (21)]. Overall, we can recover the solution to Wahba's problem from the solution of the semidefinite reformulation by taking $\mathcal{A}(Z_\star)$.

2. Relationship with the q Method

The value of the semidefinite optimization problem [Eq. (21)] is the largest eigenvalue of the Davenport matrix $\mathcal{A}^*(A_0)$. This can already be seen from Eq. (19) and the fact that $\max_{q \in \mathbb{H}} \langle \mathcal{A}^*(A_0), qq^T \rangle = \max_{q \in \mathbb{H}} q^T \mathcal{A}^*(A_0) q = \lambda_{\max}(\mathcal{A}^*(A_0))$. If q is an eigenvector corresponding to the largest eigenvalue of $\mathcal{A}^*(A_0)$, then $Z = qq^T$ is an optimal solution of the semidefinite reformulation [Eq. (21)]. As such, our reformulation is closely related to the q method for solving Wahba's problem problem [8].

3. Discussion

Note that the transformations in the first and second steps before are merely formal and can be applied to essentially any optimization problem. The third step is nontrivial. In general, it is not well understood which sets S have the property that $\text{conv}(S)$ can be described as the feasible region of a semidefinite optimization problem; this is an area of active research (see, for example, [18]). One view of this paper is that it shows how to express the convex hulls of the nonconvex constraint sets appearing in certain joint spin rate and attitude-estimation problems as the feasible regions of semidefinite optimization problems.

C. Trigonometric Wahba Problems

We now show how to give semidefinite reformulations of trigonometric Wahba problems, defined in Eq. (6). By specializing to the case where $(A_n)_{n=0}^N$ and $(B_n)_{n=1}^N$ are given by Eqs. (15) and (16), we obtain a semidefinite reformulation of Psiaki's first problem. We explicitly number the three steps in obtaining a semidefinite reformulation that were illustrated in Sec. IV.B.

1. Linear Objective Function

As in the case of Wahba's problem, we use the parameterization of $SO(3)$ in terms of unit quaternions to rewrite trigonometric Wahba problems as

$$\max_{\substack{q \in \mathbb{H} \\ \omega' \in [-\pi, \pi)}} \langle \mathcal{A}^*(A_0), qq^T \rangle + \sum_{n=1}^{N} [\langle \mathcal{A}^*(A_n), \cos(n\omega') qq^T \rangle + \langle \mathcal{A}^*(B_n), \sin(n\omega') qq^T \rangle] \tag{22}$$

We can view this problem as the maximization of a linear functional over the set

$$\mathcal{M}_N := \{(qq^T, qq^T \cos(\omega'), qq^T \sin(\omega'), \ldots, qq^T \cos(N\omega'), qq^T \sin(N\omega')) \in (\mathcal{S}^4)^{2N+1} : q \in \mathbb{H}, \omega' \in [-\pi, \pi)\} \tag{23}$$

2. Convexification

As such, the convexified version of Eq. (22) is the following optimization problem where the decision variables are the $2N + 1$ symmetric matrices $X_0, X_1, Y_1, \ldots, X_N, Y_N$:

$$\max_{(X_n)_{n=0}^N,(Y_n)_{n=1}^N} \langle \mathcal{A}^*(A_0), X_0\rangle + \sum_{n=1}^N [\langle \mathcal{A}^*(A_n), X_n\rangle + \langle \mathcal{A}^*(B_n), Y_n\rangle]$$
$$\text{subject to } (X_0, X_1, Y_1, \ldots, X_N, Y_N) \in \operatorname{conv}(\mathcal{M}_N) \tag{24}$$

This problem is certainly convex and has the same optimal value as Eqs. (6) and (22). It may not be immediately clear that the constraint set $\operatorname{conv}(\mathcal{M}_N)$ has a succinct representation in terms of the feasible region of a semidefinite optimization problem. In fact, $\operatorname{conv}(\mathcal{M}_N)$ does have such a representation, and we now turn our attention to describing it.

3. Linear Matrix Inequality Description of $\operatorname{conv}(\mathcal{M}_N)$

We now describe $\operatorname{conv}(\mathcal{M}_N)$ in terms of a linear matrix inequality, making use of the block matrix notation defined in Sec. III. We establish the correctness of this description in the Appendix, by combining standard results with a novel symmetry reduction argument.

Proposition 2:

$$\operatorname{conv}(\mathcal{M}_N) = \{(X_0, X_1, Y_1, \ldots, X_N, Y_N) \in (\mathcal{S}^4)^{2N+1}: \operatorname{tr}(X_0) = 1,$$
$$\operatorname{Toeplitz}(X_0, X_1, \ldots, X_N) + \operatorname{Hankel}(Y_N, Y_{N-1}, \ldots, Y_1, 0, -Y_1, \ldots,$$
$$-Y_{N-1}, -Y_N) \succcurlyeq 0\} \tag{25}$$

Proof: We provide a proof in the Appendix.

4. Semidefinite Reformulation in the General Case

Now that we have a semidefinite description of $\operatorname{conv}(\mathcal{M}_N)$, we can give a semidefinite reformulation for all trigonometric Wahba problems. The following theorem explicitly describes this reformulation, which is obtained by replacing $\operatorname{conv}(\mathcal{M}_N)$ in Eq. (24) with its semidefinite description from proposition 2.

Theorem 3: Let $A_0, A_1, \ldots, A_N, B_1, \ldots, B_N \in \mathbb{R}^{3\times 3}$. Then, the trigonometric Wahba problem

$$\max_{\substack{Q\in SO(3)\\ \omega'\in[-\pi,\pi)}} \langle A_0, Q\rangle + \sum_{n=1}^N [\langle A_n, \cos(\omega' n)Q\rangle + \langle B_n, \sin(\omega' n)Q\rangle] \tag{26}$$

and the semidefinite optimization problem

$$\max_{(X_n)_{n=0}^N,(Y_n)_{n=1}^N} \langle \mathcal{A}^*(A_0), X_0\rangle + \sum_{n=1}^N [\langle \mathcal{A}^*(A_n), X_n\rangle + \langle \mathcal{A}^*(B_n), Y_n\rangle]$$
$$\text{s.t. } \operatorname{Toeplitz}(X_0, X_1, \ldots, X_N)$$
$$+ \operatorname{Hankel}(Y_N, Y_{N-1}, \ldots, Y_1, 0, -Y_1, \ldots, -Y_{N-1}, -Y_N) \succcurlyeq 0$$
$$\operatorname{tr}(X_0) = 1 \tag{27}$$

have the same optimal value. The set of optimal points of the semidefinite reformulation is

$$\operatorname{conv}\{(qq^T, qq^T\cos(\omega'), qq^T\sin(\omega'), \ldots, qq^T\cos(N\omega'),$$
$$qq^T\sin(N\omega')) : (\omega', \mathcal{A}(qq^T)) \text{ is an optimal point for (26)}\} \tag{28}$$

Proof: The equivalence of Eqs. (26) and (27) follows directly from the equivalence of Eqs. (22) and (24) and the linear matrix inequality description of $\operatorname{conv}(\mathcal{M}_N)$ in proposition 2. The description of the set of optimal points is directly from the second statement of Theorem 1.

5. Extracting an Optimal Solution

If $N \geq 2$, we expect a generic trigonometric Wahba problem to have a unique optimal point $(\omega'_\star, Q_\star)$ [6]. In that case, the semidefinite reformulation [Eq. (27)] has a unique optimal point denoted $(X_{0\star}, X_{1\star}, Y_{1\star}, \ldots, X_{N\star}, Y_{N\star})$, from which we can recover $(\omega'_\star, Q_\star)$ via

$$Q_\star = \mathcal{A}(X_{0\star}),\ \cos(\omega'_\star) = \operatorname{tr}(X_{1\star}) \text{ and } \sin(\omega'_\star) = \operatorname{tr}(Y_{1\star}) \tag{29}$$

D. Pseudocode

In this section, we describe code to implement our semidefinite optimization-based formulations [Eq. (27)] of trigonometric Wahba problems. Our motivation for doing this is to show that it is quite straightforward to use standard numerical routines to solve the semidefinite optimization problems that appears in this paper.

The code is expressed in a parsing language called YALMIP [12] that runs under MATLAB. Internally, YALMIP reformulates the human-readable description of the optimization problem we specify into a standard format, then calls a numerical solver for semidefinite optimization problems (we used MOSEK [19] version 7 for these experiments) to solve the optimization problem.

In what follows, we assume we have functions:

1) `A_map` implementing the linear map $\mathcal{A}$ taking a 4×4 symmetric matrix and returning a 3×3 matrix according to Eq. (3).

2) `block_toeplitz` implementing the linear map $(X_0, X_1, \ldots, X_N) \mapsto \operatorname{Toeplitz}(X_0, X_1, \ldots, X_N)$ taking a $4\times 4\times (N+1)$ array and returning a $4(N+1)\times 4(N+1)$ matrix according to Eq. (1).

3) `block_hankel` implementing the linear map $(Y_1, Y_2, \ldots, Y_N) \mapsto \operatorname{Hankel}(-Y_N, \ldots, -Y_1, 0, Y_1, \ldots, Y_N)$ taking a $4\times 4\times N$ array and returning a $4(N+1)\times 4(N+1)$ matrix according to Eq. (2).

We declare variables in YALMIP using the `sdpvar` command.

```
1: X = sdpvar(4,4,N+1,'symmetric');
2: Y = sdpvar(4,4,N,'symmetric');
```

For example, `Y` is a $4\times 4\times N$ array of variables with each slice `Y(:,:,n)` being a symmetric matrix. We specify constraints by constructing an array of constraints expressed in a very natural way. We express the two constraints in Eq. (27) by

```
3: K = [trace(X(:,:,1)) == 1,
block_toeplitz(X) + block_hankel(Y) >= 0];
```

where we have indexed from 1 following MATLAB's conventions. Note that in YALMIP this latter inequality is automatically interpreted in the positive semidefinite sense because the matrix on the left-hand side is structurally symmetric.

Suppose that the variables `A` and `B` are, respectively, $4\times 4\times (N+1)$ and $4\times 4\times N$ arrays, with `A(:,:,n+1)` being $\mathcal{A}^*(A_n)$ and `B(:,:,n)` being $\mathcal{A}^*(B_n)$. Then, we can solve the semidefinite optimization problem [Eq. (27)] with the single line:

```
4: solvesdp(K,-(A(:)'*X(:) + B(:)'*Y(:)));
```

which calls a numerical solver with the constraint set `K` and the cost function `-(A(:)'*X(:) + B(:)'*Y(:))` (with the minus sign because minimization is the default). Assuming that there is a unique solution to the nonconvex problem, we can extract the optimal rotation matrix Q and optimal ω with

```
5: Q_opt = A_map(double(X(:,:,1)));
6: omega_opt = atan2(trace(double(Y(:,:,1))),
   trace(double(X(:,:,2))));
```

E. Numerical Algorithms and Complexity

The semidefinite reformulation of Psiaki's first problem [Eq. (27)] involves a linear matrix inequality with matrices of size $4N+4$ that have a very specific structure. This structure can be exploited to develop numerical algorithms for its solution that are much faster than the generic interior-point algorithms we used for our experiments. Indeed, semidefinite optimization problems with a similar structure arise in problems related to the Kalman–Yakubovich–Popov lemma in robust control and in that context numerous specialized algorithms have been developed for their solution (see, e.g., [20–22]). For instance, the algorithms in [21] when applied to this problem produce solutions within ϵ of being optimal in $O(N^{3.5}\log(1/\epsilon))$ operations (if we focus on the N and ϵ dependence). Because the semidefinite reformulation [Eq. (27)] always has low-rank solutions, it is likely that methods with significantly better dependence on N can be developed based on the ideas in [23]. Furthermore, significant gains can be made by producing optimized low-level code for a particular family of convex optimization problems. An excellent example of this is the code-generation software CVXGEN, which focuses on linear and convex quadratic programs [24].

V. Numerical Experiments

In this section, we describe the results of a simple numerical experiment to illustrate solving Psiaki's first problem using semidefinite optimization. The main point of the experiment is to see how the estimates obtained from Psiaki's first problem improve as more measurements are used.

The smallest number of measurements we use is three (corresponding to $N = 2$); the largest is 11 (corresponding to $N = 10$). We do not use two measurements because, in that case, the solution to Psiaki's first problem is not unique. Indeed, the analytical two-vector solution of [6] requires the use of some information from a third vector to resolve the ambiguity. The estimation errors we obtain for three measurements are a lower bound on the errors that would be obtained from the two-vector solution of [6].

We use the same parameters as in Psiaki's truth-model simulation in [6]; the true spin period is 45.32 s (and so the true spin rate is $\omega = 0.1386$ rad/s), the sampling period is $\tau = 7.7611$ seconds per sample, and the initial attitude is $Q(0) = I$. The attitude dynamics are described by Eq. (11). We fix three reference directions $x_0 = (-1/\sqrt{2}, 1\sqrt{2}, 0)$, $x_1 = (-1/\sqrt{2}, 0, 1\sqrt{2})$, and $x_2 = (0, 1/\sqrt{2}, 1\sqrt{2})$ and define $x_3 = x_6 = x_9 = x_0$, $x_4 = x_7 = x_{10} = x_1$, and $x_5 = x_8 = x_2$ (i.e., we cycle through the three reference directions).

We repeat the following experiment $T = 1000$ times. For each $\sigma \in \{0.001, 0.005, 0.01, 0.05\}$ and each $N \in \{2, 3, \ldots, 10\}$, do the following.

1) For $n = 0, 1, 2 \ldots, N$, let

$$y_n = \frac{Q(t_n)x_n + w_n}{\|Q(t_n)x_n + w_n\|}$$

where w_n is a Gaussian vector with independent components each with mean zero and standard deviation σ, and the x_n are defined in the previous paragraph;

2) Solve the semidefinite optimization reformulation of Psiaki's first problem [Eq. (27)].

To get a sense of the noise level corresponding to each value of σ, the mean angle (in degrees) between y_n and $Q(t_n)x_n$ over all samples is 0.0716 deg when $\sigma = 0.001$, 0.362 deg when $\sigma = 0.005$, 0.717 deg when $\sigma = 0.01$, and 3.61 deg when $\sigma = 0.05$.

The average angular error (in degrees) between the estimate of the initial attitude and the true initial attitude is indicated for $\sigma = 0.001$ (right) and $\sigma = 0.05$ (left) in Fig. 1. Given an estimate $\hat{Q}$ of the true initial attitude $Q(0) = I$, the angular error θ satisfies $\mathrm{tr}(\hat{Q}^T Q(0)) = 2\cos(\theta) + 1$. Hence, we compute the angular error via $|\cos^{-1}[\mathrm{tr}(\hat{Q}^T Q(0)) - 1)/2]|$. The corresponding average percentage error in the spin-rate estimate $\hat{\omega}$ is computed by taking the mean of $|\hat{\omega} - \omega|/\omega$ (as a percentage) over all trials. This is indicated for $\sigma = 0.001$ (right) and $\sigma = 0.05$ (left) in Fig. 2.

It is clear that, as more vector measurements are used (i.e., as N increases), the estimates improve. The relative improvement is greater for the spin-rate estimates than the estimates of the initial attitude. The plots are remarkably similar (apart from the scales) for $\sigma = 0.001$ and $\sigma = 0.05$ (and for $\sigma = 0.005$ and $\sigma = 0.01$, hence we omitted these). Indeed the relative improvement as N increases is very similar at the four different noise levels we considered. For example, the percentage error in the spin rate is roughly halved when using four measurements ($N = 3$) rather than three measurements ($N = 2$) regardless of whether $\sigma = 0.001, 0.005, 0.01, 0.05$. This may justify using a few more than the minimum number of measurements required for the optimization problem to have a unique optimum.

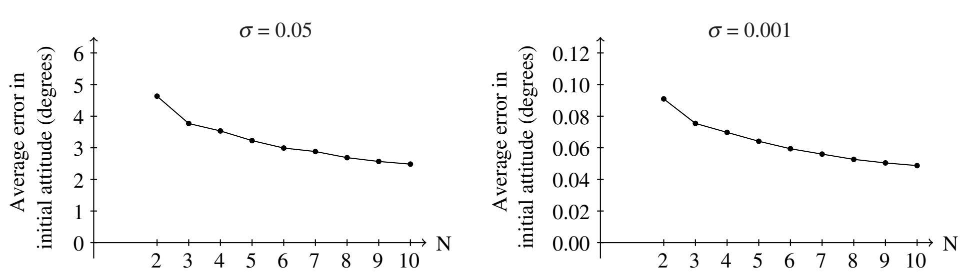

Fig. 1 Angular error (in degrees) in initial attitude estimate for different values of N and σ.

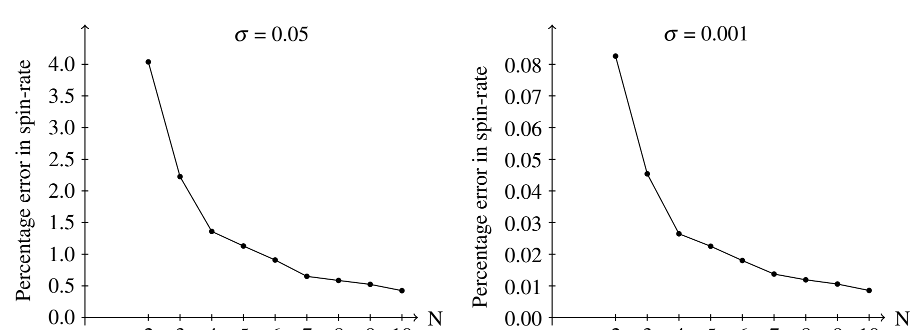

Fig. 2 Percentage error in spin-rate estimate for different values of N and σ.

VI. Future Directions

We briefly comment on possible future research directions based on the work in the present paper.

A. Variants on Psiaki's First Problem

An exact reformulation of Psiaki's first problem as the solution to a semidefinite optimization problem allows us to do more than just solve the original problem as stated. It also makes it possible to take a semidefinite relaxation-based approach to many variants on Psiaki's problem. Using basic techniques from convex optimization (see, e.g., [25]), semidefinite relaxations can be obtained for variations that accommodate additional constraints or that are robust to errors in certain parameters.

A natural variation that could be approached this way would be to obtain semidefinite relaxations of Psiaki's first problem that are robust to uncertainty in certain model parameters. A similar idea has been carried out in detail for Wahba's problem by Ahmed et al. [26]. They extended the semidefinite formulation of Wahba's problem [14] to a variant that is robust to uncertainty in certain parameters, such as the reference directions. As suggested by Ahmed et al. [26], this could be useful when using magnetometer measurements together with a low-order magnetic field model.

B. Psiaki's Second Problem

It would be interesting to try to take a similar approach to the one taken in the present paper to related problems, such as Psiaki's second problem. To do so, we would need to give a semidefinite description (or perhaps a relaxation) of

$$\operatorname*{conv}_{Q(t_0)\in SO(3),\Omega(t_0)}\{(Q(t_0), Q(t_1), \ldots, Q(t_N))\colon Q(t) = \Phi(t-t_0;\Omega(t_0))Q(t_0) \quad \text{for all } t\geq t_0\} \tag{30}$$

Given this, the objective function [Eq. (10)] can be rewritten as the maximization of a linear functional over the convex hull described in Eq. (30).

A more modest goal along similar lines might be to discretize the differential equation [Eq. (9)] and try to compute the convex hull (over all initial conditions $Q(t_0)$, $\Omega(t_0)$) of an appropriately subsampled trajectory of the associated difference equation for the attitude variables. This approach of convexifying a problem based on discretized dynamics would, in a sense, be a convex analogue of the methods proposed for Psiaki's second problem in [9].

VII. Conclusions

It has been shown how Psiaki's generalization of Wahba's problem to the case of a spacecraft spinning around a fixed axis at an unknown rate can be exactly reformulated as a semidefinite optimization problem. Such convex optimization problems can be solved globally using standard methods for semidefinite optimization. As suggested by Psiaki when formulating his generalizations of Wahba's problem, the present solutions to these generalizations of Wahba's problem could be used to initialize standard extended Kalman filter-based methods for attitude estimation.

Appendix: Proofs

In this Appendix, we prove proposition 2. We split the proof into two parts, given by lemmas 5 and 6. Together, these clearly imply proposition 2. In what follows, we extend the notation $\mathrm{Toeplitz}(T_0, T_1, \ldots, T_N)$ defined in Eq. (1) to include the case where $T_1, \ldots, T_N$ are $d\times d$ complex matrices, T_0 is a $d\times d$ Hermitian matrix, and all transposes of real matrices are replaced with conjugate transposes, denoted $A\mapsto A^*$, of complex matrices.

Lemma 5, to follow, is a slight modification of the fact that any Hermitian positive semidefinite block-Toeplitz matrix admits a decomposition as a sum of rank one positive semidefinite block-Toeplitz matrices. This fact may be more familiar in its dual form as the matrix spectral factorization (or Fejér–Riesz) theorem (see, e.g., [27]). This classical result says that any Hermitian matrix-valued function

$$S(e^{i\omega}) = \sum_{n=-N}^{N} S_n e^{in\omega}$$

that is positive semidefinite for all ω has a factorization as $S(e^{i\omega}) = W(e^{i\omega})^* W(e^{i\omega})$, where $W(e^{i\omega})$ has the form

$$W(e^{i\omega}) = \sum_{n=0}^{N} W_n e^{in\omega}$$

This result can also be interpreted as saying that nonnegative functions of the form $(z,\omega)\mapsto z^* S(e^{i\omega})z$, with $S(e^{i\omega})$ as before, are sums-of-squares [28].

Theorem 4 (Tismenetsky [29]): If $\mathrm{Toeplitz}(T_0, T_1, \ldots, T_N) \succ 0$, then there are $u_k\in\mathbb{C}^4$, $\omega_k\in[-\pi,\pi)$, and $\lambda_k > 0$ for $k = 1,2,\ldots,4(N+1)$, such that

$$\mathrm{Toeplitz}(T_0, T_1, \ldots, T_N) = \sum_{k=1}^{4(N+1)} \lambda_k \mathrm{Toeplitz}(u_k u_k^*, u_k u_k^* e^{i\omega_k}, \ldots, u_k u_k^* e^{iN\omega_k})$$

Consequently,

$$T_\ell = \sum_{k=1}^{4(N+1)} \lambda_k u_k u_k^* e^{i\ell\omega_k} \quad \text{for} \quad \ell = 0,1,\ldots,N$$

The following lemma is a slight modification of Theorem 4.

Lemma 5: Let $\mathcal{M}_N$ be defined as in Eq. (23). Then,

$$\mathrm{conv}(\mathcal{M}_N) = \{(X_0, X_1, Y_1, \ldots, X_N, Y_N)\in(\mathcal{S}^4)^{2N+1}\colon \mathrm{tr}(X_0) = 1 \quad \mathrm{Toeplitz}(X_0, X_1 + iY_1, \ldots, X_N + iY_N)\succcurlyeq 0\} \tag{A1}$$

Proof: Let $(qq^T, qq^T\cos(\omega), qq^T\sin(\omega), \ldots, qq^T\cos(N\omega), qq^T\sin(N\omega))\in\mathcal{M}_N$. Then, $\mathrm{tr}(qq^T) = \|q\|^2 = 1$, and it is straightforward to check that

$$\mathrm{Toeplitz}(qq^T, qq^T\cos(\omega) + iqq^T\sin(\omega), \ldots, qq^T\cos(N\omega) + iqq^T\sin(N\omega)) = \begin{bmatrix} q \\ qe^{-i\omega} \\ \vdots \\ qe^{-i\omega N}\end{bmatrix}\begin{bmatrix} q \\ qe^{-i\omega} \\ \vdots \\ qe^{-iN\omega}\end{bmatrix}^* \succcurlyeq 0$$

Hence, $\mathcal{M}_N$ is a subset of the right-hand side of Eq. (A1). Because the right-hand side of Eq. (A1) is convex, it follows that $\mathrm{conv}(\mathcal{M}_N)$ is also a subset of the right-hand side of Eq. (A1).

Now, suppose that $(X_0, X_1, Y_1, \ldots, X_N, Y_N)\in(\mathcal{S}^4)^{2N+1}$ satisfies

$$\mathrm{tr}(X_0) = 1 \quad \text{and} \quad \mathrm{Toeplitz}(X_0, X_1 + iY_1, \ldots, X_N + iY_N) \succ 0$$

Then, by Theorem 4, there are $u_k\in\mathbb{C}^4$, $\omega_k\in[-\pi,\pi)$ and $\lambda_k > 0$ for $k = 1,2,\ldots,(4N+1)$, such that $1 = \mathrm{tr}(X_0) = \sum_{k=1}^{4(N+1)}\lambda_k\|u_k\|^2$ and $X_\ell + iY_\ell = \sum_{k=1}^{4(N+1)}\lambda_k u_k u_k^* e^{i\ell\omega_k}$ for $\ell = 0,1,\ldots,N$. Because $X_\ell^T = X_\ell$ and $Y_\ell^T = Y_\ell$ for all ℓ, it follows by a straightforward calculation that there are $v_k\in\mathbb{R}^4$ and $\lambda_{k'} > 0$ for $k = 1,2,\ldots,(8N+1)$, such that

$$X_\ell + iY_\ell = \frac{1}{2}((X_\ell + iY_\ell) + (X_\ell + iY_\ell)^T)$$
$$= \sum_{k=1}^{4(N+1)} \frac{\lambda_k}{2}((u_k u_k^*) + (u_k u_k^*)^T) e^{i\ell\omega_k}$$
$$= \sum_{k=1}^{8(N+1)} \lambda'_k v_k v_k^T e^{i\ell\omega_k}$$

Defining $\mu_k = \lambda_{k'} \|v_k\|^2 > 0$ and $q_k = v_k/\|v_k\| \in \mathbb{H}$ for $k = 1, 2, \ldots, (8N+1)$, we have that

$$\sum_{k=1}^{8(N+1)} \mu_k = 1$$

and

$$X_\ell = \sum_{k=1}^{8(N+1)} \mu_k q_k q_k^T \cos(\ell\omega_k) \quad \text{for} \quad \ell = 0, 1, \ldots, N \quad \text{and}$$
$$Y_\ell = \sum_{k=1}^{8(N+1)} \mu_k q_k q_k^T \sin(\ell\omega_k)$$

for $\ell = 1, 2, \ldots, N$. This shows that $(X_0, X_1, Y_1, \ldots, X_N, Y_N) \in \mathrm{conv}(\mathcal{M}_N)$. Hence, the relative interior of the right-hand side of Eq. (A1) is a subset of $\mathrm{conv}(\mathcal{M}_N)$. Because $\mathrm{conv}(\mathcal{M}_N)$ is closed, the right-hand side of Eq. (A1) is also a subset of $\mathrm{conv}(\mathcal{M}_N)$, establishing the result.

Lemma 6: If $X_0, X_1, Y_1, \ldots, X_N, Y_N \in \mathcal{S}^d$, then

$$\mathrm{Toeplitz}(X_0, X_1 + iY_1, \ldots, X_N + iY_N) \succcurlyeq 0$$

if and only if

$$\mathrm{Toeplitz}(X_0, X_1, \ldots, X_N) + \mathrm{Hankel}(-Y_N, -Y_{N-1}, \ldots, -Y_1, 0, Y_1, \ldots, Y_{N-1}, Y_N) \succcurlyeq 0$$

Proof: First, observe that $Z = \mathrm{Toeplitz}(X_0, X_1 + iY_1, \ldots, X_N + iY_N) \succcurlyeq 0$ if and only if the real $2d(N+1) \times 2d(N+1)$ symmetric matrix

$$Z_\mathbb{R} = \begin{bmatrix} \Re Z & \Im Z \\ -\Im Z & \Re Z \end{bmatrix}$$

is positive semidefinite [30]. Here, $\Re Z$ and $\Im Z$ are the real and imaginary parts of Z, respectively. Indeed,

$$\Re Z = \begin{bmatrix} X_0 & X_1 & X_2 & \cdots & X_N \\ X_1 & X_0 & X_1 & \ddots & \vdots \\ X_2 & X_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & X_1 \\ X_N & \cdots & \cdots & X_1 & X_0 \end{bmatrix} \quad \text{and}$$

$$\Im Z = \begin{bmatrix} 0 & Y_1 & Y_2 & \cdots & Y_N \\ -Y_1 & 0 & Y_1 & \ddots & \vdots \\ -Y_2 & -Y_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & Y_1 \\ -Y_N & \cdots & \cdots & -Y_1 & 0 \end{bmatrix}$$

where we have used the assumption that the X_i and the Y_i are symmetric.

Let J be the $d(N+1) \times d(N+1)$ matrix with $d \times d$ identity blocks on the secondary (anti-) block diagonal. For instance, when $N = 2$,

$$J = \begin{bmatrix} 0 & 0 & I_d \\ 0 & I_d & 0 \\ I_d & 0 & 0 \end{bmatrix}$$

where I_d denotes the $d \times d$ identity matrix. Note that left multiplication by J reverses the block rows of a block matrix, and right multiplication by $J^T = J$ reverses the block columns. Observe that $J(\Re Z)J = \Re Z$ and $J(\Im Z) + (\Im Z)J = 0$. Let Q denote the orthogonal matrix defined by

$$Q = \frac{1}{\sqrt{2}}\begin{bmatrix} I & -J \\ J & I \end{bmatrix}$$

A straightforward calculation shows that

$$QZ_\mathbb{R}Q^T = \frac{1}{2}\begin{bmatrix} I & -J \\ J & I \end{bmatrix}\begin{bmatrix} \Re Z & \Im Z \\ -\Im Z & \Re Z \end{bmatrix}\begin{bmatrix} I & J \\ -J & I \end{bmatrix}$$
$$= \begin{bmatrix} \Re Z + J\Im Z & 0 \\ 0 & \Re Z + J\Im Z \end{bmatrix}$$

Thus, $Z \succcurlyeq 0$ if and only if $Z_\mathbb{R} \succcurlyeq 0$, which holds if and only if $\Re Z + J\Im Z \succcurlyeq 0$. Finally, we note that

$$\Re Z + J\Im Z = \mathrm{Toeplitz}(X_0, X_1, \ldots, X_N) + \mathrm{Hankel}(-Y_N, \ldots, -Y_1, 0, Y_1, \ldots, Y_N)$$

(because reversing the block rows of a block Toeplitz matrix makes it block Hankel) to complete the proof.

Acknowledgment

This research was funded by the U.S. Air Force Office of Scientific Research under grants FA9550-12-1-0287 and FA9550-11-1-0305.

References

[1] Lefferts, E. J., Markley, F. L., and Shuster, M. D., "Kalman Filtering for Spacecraft Attitude Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 5, No. 5, 1982, pp. 417–429. doi:10.2514/3.56190

[2] Psiaki, M. L., "Attitude-Determination Filtering via Extended Quaternion Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 23, No. 2, 2000, pp. 206–214. doi:10.2514/2.4540

[3] Crassidis, J. L., Markley, F. L., and Cheng, Y., "Survey of Nonlinear Attitude Estimation Methods," *Journal of Guidance, Control, and Dynamics*, Vol. 30, No. 1, 2007, pp. 12–28. doi:10.2514/1.22452

[4] Wahba, G., "A Least Squares Estimate of Satellite Attitude," *SIAM Review*, Vol. 7, No. 3, 1965, pp. 409–409. doi:10.1137/1007077

[5] Gebre-Egziabher, D., Elkaim, G. H., Powell, J. D., and Parkinson, B. W., "A Gyro-Free Quaternion-Based Attitude Determination System Suitable for Implementation Using Low Cost Sensors," *Proceedings of the IEEE Position Location and Navigation Symposium*, IEEE Publ., Piscataway, NJ, 2000, pp. 185–192. doi:10.1109/PLANS.2000.838301

[6] Psiaki, M. L., "Generalized Wahba Problems for Spinning Spacecraft Attitude and Rate Determination," *Journal of the Astronautical Sciences*, Vol. 57, Nos. 1–2, 2009, pp. 73–92. doi:10.1007/BF03321495

[7] Farrell, J. L., Stuelpnagel, J. C., Wessner, R. H., Velman, J. R., and Brook, J. E., "A Least Squares Estimate of Satellite Attitude," *SIAM Review*, Vol. 8, No. 3, 1966, pp. 384–386. doi:10.1137/1008080

[8] Keat, J. E., "Analysis of Least-Squares Attitude Determination Routine DOAOP," Computer Sciences Corp. TR-CSC/TM-77/6034, Falls Church, VA, Feb. 1977.

[9] Psiaki, M. L., and Hinks, J. C., "Numerical Solution of a Generalized Wahba Problem for a Spinning Spacecraft," *Journal of Guidance, Control, and Dynamics*, Vol. 35, No. 3, 2012, pp. 764–773. doi:10.2514/1.56151
[10] Hinks, J. C., and Psiaki, M. L., "Solution Strategies for an Extension of Wahba's Problem to a Spinning Spacecraft," *Journal of Guidance, Control, and Dynamics*, Vol. 34, No. 6, 2011, pp. 1734–1745. doi:10.2514/1.53530
[11] Vandenberghe, L., and Boyd, S., "Semidefinite Programming," *SIAM Review*, Vol. 38, No. 1, 1996, pp. 49–95. doi:10.1137/1038003
[12] Löfberg, J., "YALMIP: A Toolbox for Modeling and Optimization in MATLAB," *IEEE International Symposium on Computer Aided Control Systems Design*, IEEE Publ., Piscataway, NJ, 2004, pp. 284–289.
[13] Rockafellar, R. T., *Convex Analysis*, Vol. 28, Princeton Landmarks in Mathematics, Princeton Univ. Press, Princeton, NJ, 1997.
[14] Sanyal, R., Sottile, F., and Sturmfels, B., "Orbitopes," *Mathematika*, Vol. 57, No. 2, 2011, pp. 275–314. doi:10.1112/S002557931100132X
[15] Saunderson, J., Parrilo, P. A., and Willsky, A. S., "Semidefinite Descriptions of the Convex Hull of Rotation Matrices," *SIAM Journal on Optimization* (to be published).
[16] Forbes, J. R., and de Ruiter, A. H. J., "Linear-Matrix-Inequality-Based Solution to Wahba's Problem," *Journal of Guidance, Control, and Dynamics*, Vol. 38, No. 1, 2015, pp. 147–151. doi:10.2514/1.G000132
[17] Overton, M. L., and Womersley, R. S., "On the Sum of the Largest Eigenvalues of a Symmetric Matrix," *SIAM Journal on Matrix Analysis and Applications*, Vol. 13, No. 1, 1992, pp. 41–45. doi:10.1137/0613006
[18] Blekherman, G., Parrilo, P. A., and Thomas, R. R. (eds.), *Semidefinite Optimization and Convex Algebraic Geometry*, Vol. 13, MOS–SIAM Series on Optimization, Society for Industrial and Applied Mathematics, Philadelphia, 2013.
[19] Andersen, E. D., and Andersenm, K. D., "The MOSEK Interior Point Optimizer for Linear Programming: An Implementation of the Homogeneous Algorithm," *High Performance Optimization*, Springer, New York, 2000, pp. 197–232. doi:10.1007/978-1-4757-3216-0_8
[20] Liu, Z., and Vandenberghe, L., "Low-Rank Structure in Semidefinite Programs Derived from the KYP Lemma," *Proceedings of the 46th IEEE Conference on Decision and Control*, IEEE Publ., Piscataway, NJ, 2007, pp. 5652–5659. doi:10.1109/CDC.2007.4434343
[21] Genin, Y., Hachez, Y., Nesterov, Y., and Van Dooren, P., "Optimization Problems over Positive Pseudopolynomial Matrices," *SIAM Journal on Matrix Analysis and Applications*, Vol. 25, No. 1, 2003, pp. 57–79. doi:10.1137/S0895479803374840
[22] Kao, C.-Y., and Megretski, A., "On the New Barrier Function and Specialized Algorithms for a Class of Semidefinite Programs," *SIAM Journal on Control and Optimization*, Vol. 46, No. 2, 2007, pp. 468–495. doi:10.1137/050623796
[23] Burer, S., and Monteiro, R. D. C., "A Nonlinear Programming Algorithm for Solving Semidefinite Programs via Low-Rank Factorization," *Mathematical Programming*, Vol. 95, No. 2, 2003, pp. 329–357. doi:10.1007/s10107-002-0352-8
[24] Mattingley, J., and Boyd, S., "CVXGEN: A Code Generator for Embedded Convex Optimization," *Optimization and Engineering*, Vol. 13, No. 1, 2012, pp. 1–27. doi:10.1007/s11081-011-9176-9
[25] Boyd, S., and Vandenberghe, L., *Convex Optimization*, Cambridge Univ. Press, New York, 2004.
[26] Ahmed, S., Kerrigan, E. C., and Jaimoukha, I. M., "A Semidefinite Relaxation-Based Algorithm for Robust Attitude Estimation," *IEEE Transactions on Signal Processing*, Vol. 60, No. 8, 2012, pp. 3942–3952. doi:10.1109/TSP.2012.2198820
[27] Rosenblatt, M., "A Multi-Dimensional Prediction Problem," *Arkiv för Matematik*, Vol. 3, No. 5, 1958, pp. 407–424. doi:10.1007/BF02589495
[28] Aylward, E. M., Itani, S. M., and Parrilo, P. A., "Explicit SOS Decompositions of Univariate Polynomial Matrices and the Kalman–Yakubovich–Popov Lemma," *Proceedings of the 46th IEEE Conference on Decision and Control*, IEEE Publ., Piscataway, NJ, 2007, pp. 5660–5665. doi:10.1109/CDC.2007.4435026
[29] Tismenetsky, M., "Matrix Generalizations of a Moment Problem Theorem 1. The Hermitian Case," *SIAM Journal on Matrix Analysis and Applications*, Vol. 14, No. 1, 1993, pp. 92–112. doi:10.1137/0614008
[30] Goemans, M. X., and Williamson, D., "Approximation Algorithms for MAX-3-CUT and Other Problems via Complex Semidefinite Programming," *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing*, ACM, New York, 2001, pp. 443–452. doi:10.1145/380752.380838
[31] Saunderson, J., Parrilo, P. A., and Willsky, A. S., "Semidefinite Relaxations for Optimization Problems Over Rotation Matrices," *Proceedings of 53rd IEEE Conference on Decision and Control*, IEEE Publ., Piscataway, NJ, 2014, pp. 160–166. doi:10.1109/CDC.2014.7039375