Approximating a matrix as the square of a skew-symmetric matrix, with application to estimating angular velocity from acceleration data

Yang Wan[a], Benjamin E. Grossman-Ponemon[a,b], Haneesh Kesari[a,*]

[a]*School of Engineering, Brown University, Providence, RI 02912, USA*
[b]*New address: Department of Physics and Engineering, John Carroll University, University Heights, OH 44118, USA*

Abstract

In this paper we study the problem of finding the best approximation of a real square matrix by a matrix that can be represented as the square of a real, skew-symmetric matrix. This problem is important in the design of robust numerical algorithms aimed at estimating rigid body kinematics from multiple accelerometer measurements. We give a constructive proof for the existence of a best approximant in the Frobenius norm. We demonstrate the construction with some small examples, and we showcase the practical importance of this work to the problem of determining the angular velocity of a rotating rigid body from its acceleration measurements.

Keywords: Rigid body motion, Skew-symmetric Matrix, Matrix approximation, Angular velocity

1. Introduction

Approximating a given matrix by one with special properties appears in a variety of contexts in applied mathematics and engineering. For example, the problems of approximating real, square matrices in the set of symmetric, positive, semidefinite matrices [1, 2, 3] and in the orthonormal group [4] see applications in computational methods for machine learning and control theory.

Here, we concern ourselves with the following matrix approximation problem. Given a real, square matrix $\mathbf{A}$, can we find a matrix $\mathbf{U} = \mathbf{K}^2$—where $\mathbf{K}$ is a real, skew-symmetric matrix—that best approximates $\mathbf{A}$ among all such matrices? More precisely, if $\mathcal{M}_n$ is the Euclidean space of real, $n \times n$ matrices equipped with the Frobenius inner product, and $\mathcal{K}_n \subset \mathcal{M}_n$ is the subspace of real, skew-symmetric, $n \times n$ matrices, then we solve

$$\operatorname*{argmin}_{\mathbf{U} \in \mathcal{K}_n^{\mathrm{sq}}} \|\mathbf{U} - \mathbf{A}\|_F, \tag{$\mathbb{P}$}$$

where

$$\mathcal{K}_n^{\mathrm{sq}} = \left\{ \mathbf{K}^2 \ : \ \mathbf{K} \in \mathcal{K}_n \right\} \tag{1}$$

and $\|\cdot\|_F$ denotes the norm induced by the Frobenius inner product. As the above minimization depends on the matrix $\mathbf{A} \in \mathcal{M}_n$, we will refer to the problem as Problem $\mathbb{P}(\mathbf{A})$.

The main contribution of this work is an explicit construction for a minimizer of Problem $\mathbb{P}(\mathbf{A})$. We remark that this may not be the sole minimizer of Problem $\mathbb{P}(\mathbf{A})$. We provide the construction in two formats: first as Definition 2, and later as Algorithm 9. Because our minimizer is provided explicitly, we do not need to use optimization techniques such as Newton's Method to solve Problem $\mathbb{P}(\mathbf{A})$ which may only find local minima.

We are not the first to consider this optimization problem. In the case where $n = 3$, Problem $\mathbb{P}(\mathbf{A})$ has been attempted in the literature [5]. However, the construction from this previous work is obtained by investigating the stationary points of Problem $\mathbb{P}(\mathbf{A})$[1]. The construction is not proved to produce a global

---

minimizer; moreover, as presented in [5, Section III.B], the construction does not work for certain input matrices $\mathbf{A}$.[2]

To conclude this section, we elaborate on the immediate practical application of our study below. Afterwards, this work is structured as follows. A formal statement of the mathematical problem is given in §2, wherein our solution to Problem $\mathbb{P}(\mathbf{A})$ is posed as Theorem 3. In preparation for the proof of Theorem 3, we recall a few identities in §3.1. The proof of Theorem 3 then follows in §3.2. We return to the practical application of our study in §4, illustrating Algorithm 9 with some examples in §4.2 and showcasing an experimental demonstration in §5. Concluding remarks are provided in §6.

1.1. Motivating application

Angular velocity is a vector-valued measure of a rigid body's rate of rotation. Its experimental measurement is critical in a number of fields, such as aeronautics [6], astronautics [7, 8, 9, 10], robotics [11, 12], and, most recently, biomechanics [13, 14, 15]. The head's angular velocity during a traumatic event, such as a blunt impact or a fall, is considered as a key parameter for assessing that event's risk of leading to mild traumatic brain injury (mTBI) [14, 15].

The $\sqrt{\text{AO}}$-algorithm [16] is a method that estimates rigid body kinematics (including angular velocity and angular acceleration) using only measurements from accelerometers, rather than from gyroscopes. This algorithm falls into the class of what are termed *gyroscope-free algorithms* [17, 18, 5, 19]. Accelerometers typically have larger bandwidths than gyroscopes when they are of similar size and weight (Table. A.1), and so gyroscope-free algorithms allow for measurement of angular velocity at much higher frequencies since they rely only on accelerometer data.

We briefly outline the template of the $\sqrt{\text{AO}}$-algorithm below. The algorithm's inputs are measurements of the acceleration vectors at four of the rigid body's points at a discrete sequence of time instances. (See last paragraph of §5.2 for elaboration.) The acceleration vectors of the rigid body's points are usually measured in different bases (*accelerometer bases*[3]). The accelerometer bases are attached to the rigid body and move with it (cf. Figs. 1(b.i)–(b.ii)). A plain version of the $\sqrt{\text{AO}}$-algorithm consists of the following two steps. *(i)* At each time instance τ_i, $i = 1, 2, \ldots$, the acceleration data are synthesized to produce the symmetric matrix $\mathbf{B}[\tau_i]$ (see Ref. [16] for details). The computation for $\mathbf{B}[\tau_i]$ also requires information about the relative distances between the rigid body's points, and the relative orientations of the accelerometer bases. *(ii)* The angular velocity matrix of the rigid body at the time instance τ_i, $\overline{\mathbf{W}}[\tau_i] \in \mathcal{K}_n$, is then computed from the roots[4] of the function $r[\mathbf{B}[\tau_i]][\cdot]$, where $r[\mathbf{B}][\cdot] : \mathcal{K}_n \to \mathcal{S}_n$,

$$r[\mathbf{B}][\mathbf{W}] = \mathbf{W}\mathbf{W} - \mathbf{B}, \tag{2}$$

and $\mathbf{B} \in \mathcal{S}_n$, the subspace of $\mathcal{M}_n$ consisting of all real, symmetric, $n \times n$ matrices.

In theory, applying the plain $\sqrt{\text{AO}}$-algorithm to compute angular velocities is straightforward; in practice, however, this proves challenging.

The set $\mathcal{K}_n^{\text{sq}}$ defined in (1) is a proper subset of $\mathcal{S}_n$, though not possessing the same vector space structure. It can be shown that $r[\mathbf{B}][\cdot]$ must have a root when $\mathbf{B} \in \mathcal{K}_n^{\text{sq}}$ and have no roots when $\mathbf{B} \in \mathcal{S}_n \backslash \mathcal{K}_n^{\text{sq}}$. For each τ_i, the matrix $\mathbf{B}[\tau_i]$ always belongs to $\mathcal{S}_n$. When $\mathbf{B}[\tau_i]$ comes from perfect rigid body motion, then $\mathbf{B}[\tau_i] \in \mathcal{K}_n^{\text{sq}}$. Therefore, in theory, $\overline{\mathbf{W}}[\tau_i]$ can always be computed from the roots of $r[\mathbf{B}[\tau_i]][\cdot]$ (cf. [16, §4.2]), i.e., by applying step *(ii)* of the plain $\sqrt{\text{AO}}$-algorithm.

Due to noise in real acceleration data and lack of precision in position and orientation measurements, applying step *(i)* of the plain $\sqrt{\text{AO}}$-algorithm only yields an approximation for $\mathbf{B}[\tau_i]$. Let us call this approximation $\widetilde{\mathbf{B}}[\tau_i]$. Then what we can hope to get from $\widetilde{\mathbf{B}}[\tau_i]$ is only an approximation for the actual angular velocity $\overline{\mathbf{W}}[\tau_i]$. This is not an issue in itself, since most experimental procedures also only yield

---

[2]These cases are accounted for in the computer algorithm presented in [5, Appendix B], though it is unclear how these changes relate to the minimizers of the original optimization problem.

[3]e.g., the orthonormal vector sets $\left({}^{\ell}\boldsymbol{e}_i[\tau]\right)_{i \in (1,2,3)}$, $\ell = 1, 2, 3, 4$, shown in Fig. 1(b.ii)

[4]When multiple roots exist, the physically meaningful root is taken to be the one that is closest to the physically meaningful root from the previous time instance. Thus, in effect, the rigid body's initial angular velocity dictates the choice of the physically meaningful root at each time instance.

approximations for the physical quantity of interest. The issue lies in that $\widetilde{\mathbf{B}}\left[\tau_i\right]$ rarely, if ever, belongs to $\mathcal{K}_n^{\mathrm{sq}}$, and so $r\left[\widetilde{\mathbf{B}}\left[\tau_i\right]\right][\cdot]$ has no roots. Consequently, step *(ii)* of the plain $\sqrt{\mathrm{AO}}$-algorithm does not work.

An alternate strategy for computing $\widetilde{\overline{\mathbf{W}}}\left[\tau_i\right]$ from $\widetilde{\mathbf{B}}\left[\tau_i\right]$ is based on finding best approximants as follows. By construction, $\widetilde{\mathbf{B}}\left[\tau_i\right]$ always belongs to $\mathcal{S}_n$. Find a best approximant for $\widetilde{\mathbf{B}}\left[\tau_i\right]$ in $\mathcal{K}_n^{\mathrm{sq}}$. Let this matrix be $\widehat{\mathbf{B}}\left[\tau_i\right]$. Compute the roots of $r\left[\widehat{\mathbf{B}}\left[\tau_i\right]\right][\cdot]$. This is now straightforward, since $\widehat{\mathbf{B}}\left[\tau_i\right] \in \mathcal{K}_n^{\mathrm{sq}}$. Finally, compute $\widetilde{\overline{\mathbf{W}}}\left[\tau_i\right]$ from the roots of $r\left[\widehat{\mathbf{B}}\left[\tau_i\right]\right][\cdot]$, in the same manner as $\overline{\mathbf{W}}\left[\tau_i\right]$ is computed from the roots of $r\left[\mathbf{B}\left[\tau_i\right]\right][\cdot]$. This alternate strategy is the $\sqrt{\mathrm{AO}}$-algorithm. Note that it is the same as the plain $\sqrt{\mathrm{AO}}$-algorithm except for the additional approximation step, i.e., approximating $\widetilde{\mathbf{B}}\left[\tau_i\right]$ with $\widehat{\mathbf{B}}\left[\tau_i\right]$. Thus, success of the $\sqrt{\mathrm{AO}}$-algorithm is dependent on this approximation step, namely approximating an arbitrary, real, symmetric matrix by one belonging to $\mathcal{K}_n^{\mathrm{sq}}$.

In the sequel, we outline the problem we study, which is a more general version of the problem needed for the $\sqrt{\mathrm{AO}}$-algorithm. In this more general problem, we consider determining a best approximant in $\mathcal{K}_n^{\mathrm{sq}}$ of an arbitrary real, $n \times n$ matrix $\mathbf{A}$ rather than a real, symmetric $n \times n$ matrix $\mathbf{B}$. Our analysis is general with respect to the dimension n, whereas for the $\sqrt{\mathrm{AO}}$-algorithm it suffices to consider only the cases $n = 2$ and $n = 3$.

2. Preliminaries and main result

2.1. Definitions and notation

To make the statement of $\mathbb{P}$ more precise we first introduce some definitions and notation.

Matrix components:. Say the matrix $\mathbf{A} = \left(\left(A_{ij}\right)_{j\in\mathcal{I}}\right)_{i\in\mathcal{I}}$, where $A_{ij} \in \mathbb{R}$, and $\mathcal{I} = (1, \ldots, n)$. We sometimes denote the p-q^{th} component of $\mathbf{A}$, i.e. A_{pq}, as $\mathbf{A}_{\cdot p\cdot q}$.

Diagonal matrices:. An important subset of $\mathcal{S}_n$ is the set of diagonal matrices. We will compactly write these matrices using the operator $\mathsf{diag}_n\left[\cdot, \cdot, \ldots, \cdot\right] : \mathbb{R}^n \to \mathcal{S}_n$,

$$\mathsf{diag}_n\left[a_1, a_2, \ldots, a_n\right] = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}. \tag{3}$$

Set of orthogonal matrices $\mathcal{O}_n$:. A matrix $\mathbf{Q} \in \mathcal{M}_n$ is orthogonal if and only if

$$\mathbf{Q}^\mathsf{T}\mathbf{Q} = \mathbf{Q}\mathbf{Q}^\mathsf{T} = \mathbf{I}_n, \tag{4}$$

where $\mathbf{I}_n = \mathsf{diag}_n\left[1, \ldots, 1\right]$ is the identity matrix. In the previous equation, $\mathbf{Q}^\mathsf{T}$ denotes the transpose of $\mathbf{Q}$, i.e. if $\mathbf{Q} = \left(\left(Q_{ij}\right)_{j\in\mathcal{I}}\right)_{i\in\mathcal{I}}$ then $\mathbf{Q}^\mathsf{T} = \left(\left(Q_{ji}\right)_{j\in\mathcal{I}}\right)_{i\in\mathcal{I}}$. We denote the set of all orthogonal matrices in $\mathcal{M}_n$ as $\mathcal{O}_n$.

Let $\mathbf{q}_k = \left(\left(Q_{ik}\right)\right)_{i\in\mathcal{I}} \in \mathbb{R}^n$ be the kth column of $\mathbf{Q}$. Then it follows that

$$\langle \mathbf{q}_i, \mathbf{q}_j \rangle = \delta_{ij}, \tag{5}$$

for any $i, j \in \mathcal{I}$, where $\langle \cdot, \cdot \rangle$ is the standard inner product on $\mathbb{R}^n$ and δ_{ij} is the Kronecker delta symbol, which equals unity when $i = j$ and zero otherwise.

Frobenius inner product and norm:. In this work $\mathcal{M}_n$ is a Euclidean vector space. Its inner product is $\langle \cdot, \cdot \rangle_F : \mathcal{M}_n \times \mathcal{M}_n \to \mathbb{R}$,

$$\langle \mathbf{X}, \mathbf{Y} \rangle_F = \sum_{i=1}^{n} \sum_{j=1}^{n} X_{ij} Y_{ij}, \tag{6}$$

where X_{ij}, and Y_{ij} are, respectively, the i-j^{th} components of $\mathbf{X}$, and $\mathbf{Y}$. The inner product $\langle \cdot, \cdot \rangle_F$ is called the *Frobenius inner product*. The Frobenius inner product induces a norm on $\mathcal{M}_n$, $\| \cdot \|_F : \mathcal{M}_n \to \mathbb{R}$, defined as

$$\|\mathbf{A}\|_F = \sqrt{\langle \mathbf{A}, \mathbf{A} \rangle_F},$$

which we call the *Frobenius norm*.

2.2. Best approximant of a matrix by the square of a skew-symmetric matrix

Proposition 1. Problem $\mathbb{P}(\mathbf{A})$ admits a solution.

A proof for Proposition 1 can be constructed using *Weierstrass' Extreme Value Theorem* [20, Proposition A.8] and the facts that: *(i)* there exists an isometric isomorphism between $\mathcal{M}_n$ and $\mathbb{R}^{n^2}$, *(ii)* the set $\mathcal{K}_n^{\text{sq}}$ is closed in $\mathcal{M}_n$, and *(iii)* the function which maps $\mathbf{U}$ to $\|\mathbf{U} - \mathbf{A}\|_F$ is continuous and coercive on $\mathcal{K}_n^{\text{sq}}$.

However, Problem $\mathbb{P}(\mathbf{A})$ is more effectively addressed by Theorem 3, which we state later in this section. Not only does Theorem 3 imply Proposition 1 as a corollary, but it also provides a recipe for constructing a solution to Problem $\mathbb{P}(\mathbf{A})$. Therefore, we focus the remainder of this section on Theorem 3 and its proof.

To make the statement of Theorem 3 more compact, we define what we term the *skew-square-spectral approximant* of a matrix $\mathbf{A}$ as follows.

Definition 2 (Skew-square-spectral approximant). Let $\mathbf{A} \in \mathcal{M}_n$, and $\mathbf{B} = \left(\mathbf{A} + \mathbf{A}^\mathsf{T} \right) / 2$. It follows from the *Real Spectral Theorem* [21, 7.29] that $\mathbf{B}$ can be decomposed as

$$\mathbf{B} = \mathbf{N} \mathbf{\Lambda} \mathbf{N}^\mathsf{T}, \tag{7a}$$

where $\mathbf{N} \in \mathcal{O}_n$, and

$$\mathbf{\Lambda} := \mathsf{diag}_n \left[\lambda_1, \lambda_2, \ldots, \lambda_n \right], \tag{7b}$$

where $\lambda_i \in \mathbb{R}$, $i = 1, \ldots, n$, are a non-increasing sequence of real numbers. The *skew-square-spectral approximant* of $\mathbf{A}$ is defined as

$$\mathbf{U}^\star [\mathbf{A}] := \mathbf{N} \mathbf{D}^\star \mathbf{N}^\mathsf{T}, \tag{8a}$$

where $\mathbf{D}^\star$ is a diagonal matrix that depends on $\mathbf{\Lambda}$. Its form is slightly different depending on whether n is even or odd. When n is even let $k = n/2$. In this case

$$\mathbf{D}^\star := \mathsf{diag}_n \left[\mu_1^\star, \mu_1^\star, \ldots, \mu_k^\star, \mu_k^\star \right], \tag{8b}$$

where for $i \in (1, \ldots, k)$

$$\mu_i^\star = \begin{cases} \left(\lambda_{2i-1} + \lambda_{2i} \right) / 2, & \lambda_{2i-1} + \lambda_{2i} \leqslant 0, \\ 0, & \text{otherwise.} \end{cases} \tag{8c}$$

When n is odd let $k = (n-1)/2$. In this case

$$\mathbf{D}^\star := \mathsf{diag}_n \left[0, \mu_1^\star, \mu_1^\star, \ldots, \mu_k^\star, \mu_k^\star \right], \tag{9a}$$

where for $i \in (1, \ldots, k)$

$$\mu_i^\star = \begin{cases} \left(\lambda_{2i} + \lambda_{2i+1} \right) / 2, & \lambda_{2i} + \lambda_{2i+1} \leqslant 0, \\ 0, & \text{otherwise.} \end{cases} \tag{9b}$$

Theorem 3. Given $\mathbf{A} \in \mathcal{M}_n$, a solution to Problem $\mathbb{P}(\mathbf{A})$ is $\mathbf{U}^\star [\mathbf{A}]$.

Recall that problem $\mathbb{P}$ has been stated in the Introduction. In Theorem 3, $\mathbf{U}^\star [\mathbf{A}]$ is the skew-square-spectral approximant of $\mathbf{A}$, which is defined in Definition 2. Definition 2 is also a recipe for constructing $\mathbf{U}^\star [\mathbf{A}]$.

In the sequel, we prove Theorem 3.

3. Proofs

3.1. Supporting results

Before proving Theorem 3 we state four lemmas which we will use in the proof. Lemma 4 and Lemma 5 are standard results, while parts of the proof of Lemma 7 can be found in the literature. For the sake of completeness, we provide complete proofs for these lemmas in Appendix C. The final lemma, Lemma 8, is easily proven from existing results in the literature; however, we were unaware of these results during the preparation of this manuscript. A short proof, using specialized results from the literature, as well as our own longer proof, are also provided in Appendix C.

Lemma 4. For $\mathbf{S} \in \mathcal{S}_n$ and $\mathbf{K} \in \mathcal{K}_n$,

$$\|\mathbf{S} + \mathbf{K}\|_F^2 = \|\mathbf{S}\|_F^2 + \|\mathbf{K}\|_F^2.$$

Lemma 5. Let $\mathbf{A} \in \mathcal{M}_n$ and $\mathbf{Q} \in \mathcal{O}_n$. Then

$$\|\mathbf{A}\mathbf{Q}\|_F = \|\mathbf{A}\|_F = \|\mathbf{Q}\mathbf{A}\|_F.$$

For presenting Lemma 7, we first define the set $\mathcal{V}_n$.

Definition 6. For $n \in \mathbb{N}$

$$\mathcal{V}_n := \left\{ \mathbf{N}\mathbf{D}\mathbf{N}^\mathsf{T} \; : \; \mathbf{N} \in \mathcal{O}_n \text{ and } \mathbf{D} \in \mathcal{D}_n \right\}. \tag{10}$$

The set $\mathcal{D}_n$ depends on whether n is even or odd. When n is even, $\mathcal{D}_n$ is the set of all matrices of the form $\mathsf{diag}_n\left[\mu_1, \mu_1, \ldots, \mu_k, \mu_k\right]$, where $\mu_i \leqslant 0$, $i = 1, \ldots, k$, are a non-increasing sequence of real numbers and $k = n/2$. When n is odd $\mathcal{D}_n$ is the set of all matrices of the form $\mathsf{diag}_n\left[0, \mu_1, \mu_1, \ldots, \mu_k, \mu_k\right]$, with $\mu_i \leqslant 0$, $i = 1, \ldots, k$, as before and $k = (n-1)/2$.

Lemma 7. The sets $\mathcal{V}_n$ and $\mathcal{K}_n^{\mathsf{sq}}$ are the same.

Lemma 8. Let $\mathbf{\Lambda} = \mathsf{diag}_n\left[\lambda_1, \lambda_2, \ldots, \lambda_n\right]$ where $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$, and let $\mathbf{D} = \mathsf{diag}_n\left[\mu_1, \mu_2, \ldots, \mu_n\right]$ where $\mu_1 \geqslant \mu_2 \geqslant \ldots \geqslant \mu_n$. Then,

$$\min_{\mathbf{Q} \in \mathcal{O}_n} \|\mathbf{\Lambda} - \mathbf{Q}\mathbf{D}\mathbf{Q}^\mathsf{T}\|_F = \|\mathbf{\Lambda} - \mathbf{D}\|_F. \tag{11}$$

3.2. Proof of Theorem 3

Theorem 3. Given $\mathbf{A} \in \mathcal{M}_n$, a solution to Problem $\mathbb{P}(\mathbf{A})$ is $\mathbf{U}^\star\left[\mathbf{A}\right]$.

Proof. We consider the case where n is even, i.e., $n = 2k$, for some $k \in \mathbb{N}$. The case where n is odd can be handled similarly. Let $\mathbf{C} = (\mathbf{A} - \mathbf{A}^\mathsf{T})/2$ and $\mathbf{B} = (\mathbf{A} + \mathbf{A}^\mathsf{T})/2$. For any $\mathbf{U} \in \mathcal{K}_n^{\mathsf{sq}}$, by Lemma 4

$$\|\mathbf{A} - \mathbf{U}\|_F^2 = \|\mathbf{B} - \mathbf{U}\|_F^2 + \|\mathbf{C}\|_F^2. \tag{12}$$

Owing to (12), a solution to $\mathbb{P}(\mathbf{A})$ is also a solution to $\mathbb{P}(\mathbf{B})$ and vice versa. Furthermore, from $\mathbf{U}^\star\left[\mathbf{A}\right]$'s definition we note that $\mathbf{U}^\star\left[\mathbf{A}\right] = \mathbf{U}^\star\left[\mathbf{B}\right]$. Thus, it suffices to show that a solution to $\mathbb{P}(\mathbf{B})$ is $\mathbf{U}^\star\left[\mathbf{B}\right]$, i.e.,

$$\|\mathbf{B} - \mathbf{U}^\star\left[\mathbf{B}\right]\|_F^2 \leqslant \|\mathbf{B} - \mathbf{U}\|_F^2, \tag{13}$$

for any $\mathbf{U} \in \mathcal{K}_n^{\mathsf{sq}}$.

1. Since $\mathbf{U} \in \mathcal{K}_n^{\mathsf{sq}}$ it follows from Lemma 7 that $\mathbf{U}$ can be decomposed as

$$\mathbf{U} = \mathbf{M}\mathbf{D}\mathbf{M}^\mathsf{T}, \tag{14}$$

where $\mathbf{D} = \mathsf{diag}_n\left[\mu_1, \mu_1, \ldots, \mu_k, \mu_k\right]$ with $0 \geqslant \mu_1 \geqslant \ldots \geqslant \mu_k$, and $\mathbf{M} \in \mathcal{O}_n$.

2. Applying $\mathbf{U}$'s decomposition given in (14) and $\mathbf{B}$'s decomposition given in (7) we get that

$$\|\mathbf{B} - \mathbf{U}\|_F^2 = \left\|\mathbf{N}\boldsymbol{\Lambda}\mathbf{N}^\mathsf{T} - \mathbf{M}\mathbf{D}\mathbf{M}^\mathsf{T}\right\|_F^2,$$

$$= \left\|\mathbf{N}\left(\boldsymbol{\Lambda} - \mathbf{N}^\mathsf{T}\mathbf{M}\mathbf{D}\mathbf{M}^\mathsf{T}\mathbf{N}\right)\mathbf{N}^\mathsf{T}\right\|_F^2,$$

$$= \left\|\boldsymbol{\Lambda} - \mathbf{N}^\mathsf{T}\mathbf{M}\mathbf{D}\mathbf{M}^\mathsf{T}\mathbf{N}\right\|_F^2, \tag{15a}$$

where the third equality follows from two applications of Lemma 5 to remove the orthogonal matrices $\mathbf{N}$ and $\mathbf{N}^\mathsf{T}$. Defining $\mathbf{Q} = \mathbf{N}^\mathsf{T}\mathbf{M} \in \mathcal{O}_n$,

$$\|\mathbf{B} - \mathbf{U}\|_F^2 = \left\|\boldsymbol{\Lambda} - \mathbf{Q}\mathbf{D}\mathbf{Q}^\mathsf{T}\right\|_F^2, \tag{15b}$$

$$\geqslant \|\boldsymbol{\Lambda} - \mathbf{D}\|_F^2, \tag{15c}$$

where the inequality follows from Lemma 8.
3. Next, using $\mathbf{B}$'s decomposition given in (7) and $\mathbf{U}^\star[\mathbf{B}]$'s decomposition (which is the same as $\mathbf{U}^\star[\mathbf{A}]$'s decomposition) given in (8) we have

$$\|\mathbf{B} - \mathbf{U}^\star[\mathbf{B}]\|_F^2 = \|\boldsymbol{\Lambda} - \mathbf{D}^\star\|_F^2, \tag{16}$$

where we have used similar manipulations to the previous step.
4. Subtracting (16) from (15c) we get

$$\|\mathbf{B} - \mathbf{U}\|_F^2 - \|\mathbf{B} - \mathbf{U}^\star[\mathbf{B}]\|_F^2 \geqslant \|\boldsymbol{\Lambda} - \mathbf{D}\|_F^2 - \|\boldsymbol{\Lambda} - \mathbf{D}^\star\|_F^2,$$

$$= 2\sum_{i=1}^{k} (\mu_i - \mu_i^\star)(\mu_i + \mu_i^\star$$

$$- \lambda_{2i-1} - \lambda_{2i}), \tag{17}$$

where the equality follows direct calculation of the Frobenius norms using the explicit formulas for $\boldsymbol{\Lambda}$, $\mathbf{D}$, and $\mathbf{D}^\star$ in terms of λ_i, μ_i, and $\mu_i^\star$, respectively.
5. We claim each term in the sum in (17) is non-negative, and so (13) holds. Let $1 \leqslant i \leqslant k$.
   (5.a) Suppose that $\lambda_{2i-1} + \lambda_{2i}$ is non-positive. Then, by definition $\mu_i^\star = (\lambda_{2i-1} + \lambda_{2i})/2$, and so the term in the sum simplifies to

$$\left(\mu_i - \frac{\lambda_{2i-1} + \lambda_{2i}}{2}\right)^2,$$

   which is clearly non-negative.
   (5.b) Otherwise, suppose that $\lambda_{2i-1} + \lambda_{2i}$ is positive. Then $\mu_i^\star = 0$ and the term in the sum is

$$\mu_i\left(\mu_i - \lambda_{2i-1} - \lambda_{2i}\right).$$

   Since $\mu_i \leqslant 0$ (see Step 1) and $-\lambda_{2i-1} - \lambda_{2i} < 0$, this is the product of two negative numbers, which is positive.

$\square$

4. Algorithm and examples

4.1. Algorithm

Following Definition 2, we summarize the procedure to compute the best approximant of $\mathbf{A}$ in $\mathcal{K}_n^{\mathrm{sq}}$, i.e., $\mathbf{U}^\star[\mathbf{A}]$ in Algorithm 9.

4.2. Basic examples

We illustrate the application of Algorithm 9 for some simple matrices, which highlight the salient steps and emphasize the potential existence of multiple solutions.

Algorithm 9: Computing a best approximant to a matrix $\mathbf{A} \in \mathcal{M}_n$ in the set $\mathcal{K}_n^{\mathrm{sq}}$

Input: A matrix $\mathbf{A} \in \mathcal{M}_n$, $n \geqslant 1$;

Output: A best approximant to $\mathbf{A}$ in $\mathcal{K}_n^{\mathrm{sq}}$, $\mathbf{U}^\star [\mathbf{A}]$;

Compute the symmetric part of $\mathbf{A}$, $\mathbf{B} = \left(\mathbf{A} + \mathbf{A}^\mathsf{T}\right)/2$;

Compute a real spectral decomposition of

$$\mathbf{B} = \mathbf{N}\mathrm{diag}_n\left[\lambda_1, \lambda_2, \ldots, \lambda_n\right]\mathbf{N}^\mathsf{T}$$

  in which $\mathbf{N} \in \mathcal{O}_n$ and $\lambda_i \in \mathbb{R}$, $i = 1, \ldots, n$, are a non-increasing sequence of real numbers;

$\mathbf{d}^\star \leftarrow \varnothing$;   `/* an empty list */`

if *n is even* **then**

    $k = n/2$;

    for $i = 1$ *to* k **do**

        if $\lambda_{2i-1} + \lambda_{2i} \leqslant 0$ **then**

            $\mu^\star \leftarrow \left(\lambda_{2i-1} + \lambda_{2i}\right)/2$;

        else

            $\mu^\star \leftarrow 0$;

        end

        Append $\mu^\star$ to $\mathbf{d}^\star$ twice;

    end

else

    $k = (n-1)/2$;

    Append 0 to $\mathbf{d}^\star$;

    for $i = 1$ *to* k **do**

        if $\lambda_{2i} + \lambda_{2i+1} \leqslant 0$ **then**

            $\mu^\star \leftarrow \left(\lambda_{2i} + \lambda_{2i+1}\right)/2$;

        else

            $\mu^\star \leftarrow 0$;

        end

        Append $\mu^\star$ to $\mathbf{d}^\star$ twice;

    end

end

$\mathbf{D}^\star \leftarrow \mathrm{diag}_n\left[\mathbf{d}^\star\right]$;

return $\mathbf{U}^\star[\mathbf{A}] = \mathbf{N}\mathbf{D}^\star\mathbf{N}^\mathsf{T}$.

Example 1. Let

$$\mathbf{A} = \begin{pmatrix} -1 & 4 & 2 \\ 2 & -1 & 3 \\ -2 & -3 & -6 \end{pmatrix}. \tag{18}$$

The symmetric part of $\mathbf{A}$ is

$$\mathbf{B} = \left(\mathbf{A} + \mathbf{A}^\mathsf{T}\right)/2 = \begin{pmatrix} -1 & 3 & 0 \\ 3 & -1 & 0 \\ 0 & 0 & -6 \end{pmatrix}. \tag{19}$$

A real spectral decomposition of $\mathbf{B}$ of the form stipulated by Algorithm 9[5] is

$$\mathbf{B} = \mathbf{N}\boldsymbol{\Lambda}\mathbf{N}^\mathsf{T}, \tag{20a}$$

---

[5]It can be shown that this real spectral decomposition for the symmetric part of the given $\mathbf{A}$ is the only one possible for it that has the form stipulated by Algorithm 9.

where

$$\mathbf{N} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{20b}$$

$$\mathbf{\Lambda} = \mathsf{diag}_3 \left[2, -4, -6 \right]. \tag{20c}$$

Next we need to construct the matrix $\mathbf{D}^\star$. As $n = 3$ (and $k = (n-1)/2 = 1$) we know

$$\mathbf{D}^\star = \mathsf{diag}_3 \left[0, \mu_1^\star, \mu_1^\star \right]. \tag{21}$$

Since $\lambda_2 + \lambda_3 = -4 - 6 = -10 < 0$, we get that

$$\mu_1^\star = (-4 - 6)/2 = -5.$$

Thus, we can construct $\mathbf{U}^\star \left[\mathbf{A} \right]$:

$$\mathbf{U}^\star \left[\mathbf{A} \right] = \mathbf{N} \mathbf{D}^\star \mathbf{N}^\mathsf{T} = \begin{pmatrix} -\frac{5}{2} & \frac{5}{2} & 0 \\ \frac{5}{2} & -\frac{5}{2} & 0 \\ 0 & 0 & -5 \end{pmatrix}. \tag{22}$$

Example 2. Our work provides a means of constructing *a* best approximation to a given matrix in the sense of Problem $\mathbb{P}(\mathbf{A})$; however, there is no guarantee of uniqueness. For example, consider the matrix $\mathbf{A} = \mathbf{B} = \mathsf{diag}_3 \left[-1, -1, -1 \right]$. This matrix does not have a unique spectral decomposition of the form stipulated by Algorithm 9. That is, $\mathbf{B}$ can be written as $\mathbf{N}\mathsf{diag}_3 \left[-1, -1, -1 \right] \mathbf{N}^\mathsf{T}$ where $\mathbf{N}$ is any element in $\mathcal{O}_3$. Following Algorithm 9, we set

$$\mathbf{U}^\star \left[\mathbf{A} \right] = \mathbf{N}\mathsf{diag}_3 \left[0, -1, -1 \right] \mathbf{N}^\mathsf{T}, \tag{23}$$

where we reiterate that $\mathbf{N}$ is an arbitrary element of $\mathcal{O}_3$. In other words, all elements of the set

$$\left\{ \mathbf{N} \, \mathsf{diag}_3 \left[0, -1, -1 \right] \mathbf{N}^\mathsf{T} \; : \; \mathbf{N} \in \mathcal{O}_3 \right\} \subset \mathcal{K}_n^{\mathrm{sq}} \tag{24}$$

are best approximants of $\mathbf{A}$.

5. An application in measuring angular velocity

In this section, we demonstrate an experimental application of Theorem 3. We use our result to robustly estimate a rigid body's angular velocity from experimental accelerometer data. Specifically, we apply the $\sqrt{\text{AO}}$-algorithm to the accelerometer data from a rigid body rotation experiment (Fig. 1) and estimate the angular velocity vector of the rigid body in that experiment. Recall that Theorem 3 is an integral part of the $\sqrt{\text{AO}}$-algorithm; it is used for carrying out the critical step of constructing an approximation for the matrix $\widetilde{\mathbf{B}} \left[\tau_i \right]$ in the space $\mathcal{K}_n^{\mathrm{sq}}$ (see §1.1 for details). Without this approximation step, the $\sqrt{\text{AO}}$-algorithm reduces to the plain $\sqrt{\text{AO}}$-algorithm. The $\sqrt{\text{AO}}$-algorithm is a viable means for estimating angular velocity in real-world situations wherein the data contains noise and errors, and, due to the elasticity of materials, the accelerometers are no longer rigidly affixed to one another (so that generally $\widehat{\mathbf{B}} \left[\tau_i \right] \notin \mathcal{K}_n^{\mathrm{sq}}$). In contrast, the plain $\sqrt{\text{AO}}$-algorithm can only calculate the angular velocity in ideal situations[6], wherein the data is free from noise and errors and the motion is that of a perfect rigid body.

---

[6]By "real-world" and "ideal" we are not simply contrasting between experimental and computational situations. Even motion extracted from computational mechanics simulations of rigid body motion can contain numerical noise, especially at large time steps, and it too could only be an approximation of rigid body motion, since specialized time integration schemes are required to perfectly maintain the rigidity constraint in the simulation.

a Apparatus of the rotation experiment

Accelerometer#4, A_4

Rotation direction

Rotation axis, N

A_2

A_3

Sensor holder

A_1

Motor

Chassis

b Abstract copy of the rotation experiment

(i)

(ii)

Figure 1: Rigid body rotation experiment setup. (a) shows a sketch of the rigid body rotation experiment described in §5.1. (b) shows a more abstract version of the experiment. Subfigures (i) and (ii) respectively denote the configurations of the rigid body in our experiment at the initial and a later time instance. All the mathematical symbols in this figure are described in §5.1–5.3. For instance, the arrows marked e_i, $i = 1, 2, 3$, denote the basis vectors, and the blue cuboid marked A_4 denotes the accelerometer #4. Its position vector at the initial and at a later time instance is, respectively, marked in (i) and (ii) as $^4x[0]$ and $^4x[\tau]$. The components of $^4x[0]$ w.r.t. $(e_i)_{i \in I}$, which we denote as $^4\mathbf{x}[0]$, in the punctuated rotation trial (see, e.g., §5.1.2) are $(0.02, 0.03, 0.11)$. Similarly, the components of the other accelerometers' initial position vectors are $^1\mathbf{x}[0] = (-0.08, -0.01, 0.04)$, $^2\mathbf{x}[0] = (0.04, -0.06, 0.01)$, and $^3\mathbf{x}[0] = (0.02, 0.08, -0.05)$. The arrows marked $\acute{e}_i[0]$, $i = 1, 2, 3$, in (i) denote accelerometer #4's measurement directions at the initial time instance. Their components w.r.t. $(e_i)_{i \in I}$ in the punctuated rotation trial are $(-0.68, 0.14, 0.71)$, $(0.08, -0.95, 0.27)$, and $(0.72, 0.24, 0.64)$. Similarly, the components of $(^1e_i[0])_{i \in I}$ are $((0.44, 0.66, 0.61), (-0.87, 0.48, 0.11), (-0.22, -0.58, 0.79))$, of $(^2e_i[0])_{i \in I}$ are $((0.07, 0.97, 0.22), (-0.96, 0, 0.28), (0.27, -0.23, 0.93))$, and of $(^3e_i[0])_{i \in I}$ are $((-0.42, 0.45, -0.78), (-0.90, -0.24, 0.34), (-0.03, 0.85, 0.51))$. Furthermore, the arrows marked $e_i[0]$, $i = 1, 2, 3$, in (i) denote the body basis vectors at the initial time instance. Their respective components in the punctuated rotation trial are $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$. The arrow marked n denotes the rotation axis, and its components are $\mathbf{n} = (-0.27, -0.28, -0.92)$.

5.1. Rigid body motion and rotation experiments

We conducted a rigid body rotation experiment to demonstrate an application of the $\sqrt{\text{AO}}$-algorithm, and, consequently, that of Theorem 3.

5.1.1. Rigid body motion

A sketch of the rigid body and the apparatus we used to impose a motion on it in our rotation experiment are shown in Fig. 1(a). Photographs of our experimental setup are provided in Fig. B.6. A more abstract representation of our rigid body is shown in Fig. 1(b). As shown in Fig. 1(b), we take our rigid body to execute its motion in the physical Euclidean point space $\mathcal{E}$, which is a three dimensional real affine space. We term the inner product space associated with it the physical Euclidean vector space, and denote it as $\mathbb{E}$. An orthonormal basis for $\mathbb{E}$ is the set of vectors $(\boldsymbol{e}_i)_{i \in I}$, where $I = (1, 2, 3)$, shown in Fig. 1(b). In our formulation the vectors $\boldsymbol{e}_i$ have units of meters[7]. Let the position vector of a rigid body material particle $\mathcal{P}$ w.r.t. the origin o, shown in Fig. 1(b), at the time instance $\boldsymbol{\tau}$ be $^{\mathcal{P}}\boldsymbol{x}\left[\boldsymbol{\tau}\right]$. We then refer to

$$^{\mathcal{P}}\mathbf{x}\left[\tau\right] := \left(^{\mathcal{P}}\boldsymbol{x}\left[\boldsymbol{\tau}\right] \cdot \boldsymbol{e}_i\right)_{i \in I} \in \mathbb{R}^3, \tag{25}$$

as particle $\mathcal{P}'s$ non-dimensional position vector at the non-dimensional time instance $\tau \in \mathbb{R}$. In our formulation τ is defined such that τ seconds $= \boldsymbol{\tau}$. From here on we will drop the qualifiers "dimensional" and "non-dimensional," since these should be clear from context[8]. In a general rigid body motion $^{\mathcal{P}}\mathbf{x}\left[\tau\right]$ evolves as

$$^{\mathcal{P}}\mathbf{x}\left[\tau\right] = \mathbf{R}\left[\tau\right] {}^{\mathcal{P}}\mathbf{x}_0 + \mathbf{c}\left[\tau\right], \tag{26}$$

where $\mathbf{R}\left[\tau\right]$ is a rotation matrix, belonging to the special orthogonal group $SO(3, \mathbb{R})$ (a subset of $\mathcal{O}_3$), and $^{\mathcal{P}}\mathbf{x}_0, \mathbf{c}\left[\tau\right] \in \mathbb{R}^3$. The matrix $\mathbf{R}\left[\tau\right]$ and the vector $\mathbf{c}\left[\tau\right]$ are called the rotation matrix and the translation vector at the time instance τ, respectively.

The angular velocity vector of a rigid body from classical physics, $\boldsymbol{\omega}[\boldsymbol{\tau}]$, is related to $\mathbf{R}[\tau]$ as follows. The *angular velocity matrix in the body frame* $\overline{\mathbf{W}}\left[\tau\right]$ is defined as

$$\overline{\mathbf{W}}\left[\tau\right] = \mathbf{R}^\mathsf{T}\left[\tau\right]\mathbf{R}'\left[\tau\right], \tag{27}$$

where $\mathbf{R}'\left[\tau\right]$ is the derivative of $\mathbf{R}[\tau]$. It can be shown that $\overline{\mathbf{W}}\left[\tau\right] \in \mathcal{K}_3 \cong \mathfrak{so}(3, \mathbb{R})$, the Lie algebra of $SO(3, \mathbb{R})$. The *body angular velocity vector* is defined as

$$\overline{\mathbf{w}}\left[\tau\right] = \star\overline{\mathbf{W}}\left[\tau\right], \tag{28}$$

where "$\star$" denotes the map $\mathfrak{so}\left(3, \mathbb{R}\right) \ni \mathbf{W} \mapsto \star\mathbf{W} \in \mathbb{R}^3$,

$$\star\mathbf{W} = \left(\mathbf{W}_{\cdot 3 \cdot 2}, \mathbf{W}_{\cdot 1 \cdot 3}, \mathbf{W}_{\cdot 2 \cdot 1}\right).$$

The *angular velocity vector in the laboratory frame* is

$$\mathbf{w}[\tau] := \sum_{i \in I} \overline{\mathbf{w}}_{\cdot i}\left[\tau\right]\mathbf{e}_i[\tau], \tag{29}$$

where

$$\mathbf{e}_i\left[\tau\right] := \mathbf{R}\left[\tau\right]\mathbf{e}_i, \tag{30}$$

and $\mathbf{e}_1 := (1, 0, 0)$, $\mathbf{e}_2 := (0, 1, 0)$, and $\mathbf{e}_3 := (0, 0, 1)$[9]. The vector $\mathbf{w}[\tau]$ is, in fact, a non-dimensional form of $\boldsymbol{\omega}\left[\boldsymbol{\tau}\right]$.

---

[7]This formalism where the elements in a vector space have units, while their components w.r.t. any basis of that vector space are non-dimensional, was introduced in [22, 14].

[8]To aid the reader, we denote non-dimensional quantity using sans serif fonts. For example, we denote the non-dimensional counterpart to the dimensional/physical scalar $\boldsymbol{\tau}$ as τ, and the non-dimensional counterpart to the dimensional vector $\boldsymbol{x}$ as $\mathbf{x}$.

[9]The vectors $\mathbf{e}_i\left[\tau\right] := \mathbf{R}\left[\tau\right]\mathbf{e}_i$ are called the body basis vectors. Their dimensional counterparts appear in Fig. 1(b) as $\boldsymbol{e}_i\left[\boldsymbol{\tau}\right]$. These vectors can be thought of as being attached to the body and rotating with it (cf. Figs. 1(b.i) and 1(b.ii)).

5.1.2. Experimental configurations

The $\sqrt{AO}$-algorithm provides an approach to estimate the body angular velocity matrix, $\overline{\mathbf{W}}[\tau]$, of a rigid body when provided with acceleration measurements at four of its points (see §1.1 for details). It can handle data from the general rigid body motion given in (26). However, in order to make it easy for us to quantify the accuracy of the $\sqrt{AO}$-algorithm's estimated angular velocities, in our experiments we only impose a relatively simple, special case of the motion given in (26). Specifically, in our experiments we impose

$$\mathbf{c}[\tau] \equiv \mathbf{0}, \tag{31a}$$

i.e., no translation, and

$$\mathbf{R}[\tau] = \mathbf{n} \otimes \mathbf{n} + (\mathbf{I} - \mathbf{n} \otimes \mathbf{n}) \cos[\theta[\tau]] + (*\mathbf{n}) \sin[\theta[\tau]], \tag{31b}$$

where $\mathbf{I}$ is the identity element in $\mathcal{M}_3$, $\mathbf{n} \in \mathbb{R}^3$, $\theta[\cdot]$ is a continuous function, and "$\otimes$" is the dyadic product, such that if $\mathbf{u}$, $\mathbf{v} \in \mathbb{R}^n$ then $\mathbf{u} \otimes \mathbf{v} \in \mathcal{M}_n$, and $(\mathbf{u} \otimes \mathbf{v})_{.i \cdot j} = \mathbf{u}_{.i} \mathbf{v}_{.j}$. The "$*$" appearing in (31b) is the inverse of $\star$, mapping $\mathbf{n} \in \mathbb{R}^3$ to

$$*\mathbf{n} = \left(\left(\sum_{k \in I} \epsilon_{ikj} \mathbf{n}_{.k} \right)_{j \in I} \right)_{i \in I} \in \mathfrak{so}(3, \mathbb{R}),$$

where ϵ_{ijk} is the Levi-Cevita symbol. The vector $\mathbf{n}$ is called the rotation axis, and $\theta[\tau] \in \mathbb{R}$ the rotation angle at the time instance τ.

As per (31b) our rigid body rotates about the fixed axis $\mathbf{n}$. (The dimensional counterpart of $\mathbf{n}$, namely n, can be seen in, e.g., Fig. 1(b).) We use a brushless direct current (BLDC) motor (CPM-SCHP-3426D-ELNB, ClearPath, Teknic, USA; see Fig. 1(a)) to vary the rotation angle over time in a precise and systematic fashion. Configurations of our rigid body at several θ values are shown in Fig. 2(b).

We carried out three experiments, each with a different imposed $\theta[\cdot]$. The results from all three experiments are quite similar. Therefore, here we only present results from the first experiment, which we term *punctuated rotation*; and discuss the other two (which we refer to as *oscillatory rotation*, and *constant rotation*) in §Appendix B. In the punctuated rotation experiment

$$\theta[\tau] = \frac{\omega_\mathrm{m}}{2}\tau - \frac{2\omega_\mathrm{m}\tau_1}{\pi^3} \sum_{n=1,3,5} \frac{1}{n^3} \sin\left[\frac{2n\pi\tau}{\tau_1}\right], \tag{31c}$$

where ω_m and τ_1 are positive real numbers. The angular velocity in the punctuated rotation motion (31) has the form of a triangle wave with period τ_1 sec, varying between zero and ω_m Hertz[10], see Fig. 2(a.ii).

5.2. Acceleration measurements

Results from a trial of the punctuated rotation experiment are shown in Figs. 3–5. The rotation $\theta[\cdot]$ was that in (31c) with $\omega_\mathrm{m} = 31.41$, and $\tau_1 = 5.81$. (Graphs of $\theta[\cdot]$ and its derivative for these parameter values are shown in Figs. 2(a.i) and (a.ii), respectively.) The axis vector $\mathbf{n}$ in this trial was $(-0.27, -0.28, -0.92)$. (The dimensional form n is shown in Fig. 1(b).)

It follows from (27), (28), and (31b) that the body angular velocity vector $\overline{\mathbf{w}}$ is

$$\overline{\mathbf{w}}[\tau] = \theta'[\tau]\,\mathbf{n}. \tag{32}$$

The time-dependency of the components of $\overline{\mathbf{w}}$'s for the punctuated rotation experiment are shown in Fig. 5. The $\sqrt{AO}$-algorithm can estimate these components when provided with acceleration measurements at four of the rigid body's material particles. We discuss these measurements in the reminder of this section, and the application of the $\sqrt{AO}$-algorithm to these measurements in the next section. (For prediction of mild traumatic brain injury, the angular velocity vector $\mathbf{w}$'s magnitude is more critical than its direction. The

---

[10]It is more common to use rad/sec as the units of angular velocity. We, however, chose to use the equivalent units of Hertz.

body angular velocity vector $\overline{\mathbf{w}}$'s magnitude is always equal to that of $\mathbf{w}$'s. In the current trial, due to the simple nature of the imposed motion, their directions are same as well.)

Acceleration measurements at four of the rigid body's particles in the trial are shown in Fig. 3. We denote the four particles as $A_1, \ldots, A_4$ (marked, e.g., in Fig. 1(a)). The acceleration vector at each A_ℓ, $\ell = 1, 2, 3, 4$, was measured by a tri-axial accelerometer (Blue Trident, Vicon, UK; 1.6 kHz sampling rate) that was rigidly affixed to it. (These accelerometers are the blue objects in Fig. 1(a). In abstract copies of our rigid body in Figs. 1(b) and 2(b), they appear as blue cuboids.) Subfigures (a)–(d) in Fig. 3 correspond, respectively, to the measurements from A_1–A_4. Each subfigure shows graphs of three (discrete) functions. For example, subfigure (d) shows graphs of ${}^4\alpha_i\,[\cdot]$, $i = 1, 2, 3$. These functions are related to A_4's acceleration vector in the following manner. Say ${}^4\mathbf{a}[\tau]$ is the non-dimensional acceleration of A_4. Then ${}^4\alpha_i[\tau] := {}^4\mathbf{a}\,[\tau] \cdot {}^4\mathbf{e}_i\,[\tau]$, $i = 1, 2, 3$. The vectors $\left({}^4\mathbf{e}_i\,[\tau]\right)_{i \in I}$ give the directions of accelerometer A_4's measurement axes at the time instance τ^{11}. The functions graphed in the other subfigures (resp., $\left({}^1\alpha_i\,[\cdot]\right)_{i \in I}$, $\left({}^2\alpha_i\,[\cdot]\right)_{i \in I}$, and $\left({}^3\alpha_i\,[\cdot]\right)_{i \in I}$) are similarly related to the acceleration vector evolutions at the other points (resp., ${}^1\mathbf{a}\,[\cdot]$, ${}^2\mathbf{a}\,[\cdot]$, and ${}^3\mathbf{a}\,[\cdot]$).

5.3. Applying the $\sqrt{\mathrm{AO}}$-algorithm

To recall, the $\sqrt{\mathrm{AO}}$-algorithm allows us to determine the angular velocity of a rigid body from acceleration measurements. The $\sqrt{\mathrm{AO}}$-algorithm takes as input discrete forms of the accelerometer measurements ${}^\ell\boldsymbol{\alpha}\,[\cdot] := \left({}^\ell\alpha_i\,[\cdot]\right)_{i \in I}$, $\ell = 1, 2, 3, 4$, which we discussed in the previous section. Additionally, it requires the accelerometers' relative positions and orientations. These can be computed from the initial values of the accelerometers' position vectors^{12} ${}^\ell\mathbf{x}[0]$, $\ell = 1, 2, 3, 4$, and their measurement axes $\left({}^\ell\mathbf{e}_i\,[0]\right)_{i \in I}$. Our rigid body was 3D printed using fused deposition molding (FDM, material: polylactic acid (PLA), 3D printer: Original Prusa i3 MK3, Prusa Research, the Czech Republic). The parameters in its design were used to compute the initial values ${}^\ell\mathbf{x}[0]$, and $\left({}^\ell\mathbf{e}_i\,[0]\right)_{i \in I}$. These values for our punctuated rotation trial are given in the caption of Fig. 1.

On feeding in the accelerometer measurements ${}^\ell\boldsymbol{\alpha}\,[\tau_i]$, $i = 1, 2, \ldots$, and other experimental parameters, the $\sqrt{\mathrm{AO}}$-algorithm yields a discrete version of the function $\mathbb{R} \ni \tau \mapsto \mathbf{P}\,[\tau] \in \mathcal{M}_3$ (for details see [16, Eq. (4.1)]). The body angular velocity matrix of the rigid body at the time instance τ, $\overline{\mathbf{W}}\,[\tau]$, is related to this output as

$$r\,[\mathbf{B}\,[\tau]]\,\big[\overline{\mathbf{W}}\,[\tau]\big] = \mathbf{0}, \tag{33}$$

where $\mathbf{B}\,[\tau] \in \mathcal{K}_3^{\mathrm{sq}}$ is the symmetric part of $\mathbf{P}\,[\tau]$, and $r\,[\cdot]\,[\cdot]$ is defined in (2). As discussed in §1.1, due to experimental noise and errors, we do not obtain $\mathbf{B}\,[\tau]$ directly, but only an approximation for it, $\widetilde{\mathbf{B}}\,[\tau]$, which usually does not lie in $\mathcal{K}_3^{\mathrm{sq}}$. (More specifically, we do not obtain $\mathbf{P}\,[\tau]$ but only an approximation for it, $\widetilde{\mathbf{P}}\,[\tau]$, and take the symmetric part of $\widetilde{\mathbf{P}}\,[\tau] =: \widetilde{\mathbf{B}}\,[\tau]$ as an approximation for $\mathbf{B}\,[\tau]$. Graphs of some of the component functions of $\widetilde{\mathbf{B}}\,[\cdot]$ from the trial of our punctuated rotation experiment are shown in Fig. 4.) The next step of the $\sqrt{\mathrm{AO}}$-algorithm involves finding $\widehat{\mathbf{B}}\,[\tau_i]$, a best approximant for $\widetilde{\mathbf{B}}\,[\tau_i]$ in $\mathcal{K}_3^{\mathrm{sq}}$. We calculate $\widehat{\mathbf{B}}\,[\tau_i]$ by applying Algorithm 9. Specifically, we compute $\widehat{\mathbf{B}}\,[\tau_i]$ as $\mathbf{U}^\star\big[\widetilde{\mathbf{B}}\,[\tau_i]\big]$, which in the current case can be written as

$$\mathbf{U}^\star\big[\widetilde{\mathbf{B}}\,[\tau_i]\big] = \mathbf{N}\,[\tau_i]\,\mathrm{diag}_3\,[0, \mu_1^\star\,[\tau_i], \mu_1^\star\,[\tau_i]]\,\mathbf{N}\,[\tau_i]^\mathsf{T}, \tag{34a}$$

where

$$\mu_1^\star\,[\tau_i] = \begin{cases} (\lambda_2\,[\tau_i] + \lambda_3\,[\tau_i])\,/2, & \lambda_2\,[\tau_i] + \lambda_3\,[\tau_i] \leqslant 0, \\ 0, & \text{otherwise}, \end{cases} \tag{34b}$$

---

^{11}The dimensional counterparts of $\left({}^4\mathbf{e}_i\,[\cdot]\right)_{i \in I}$ at the initial and at a later time instance are, respectively, shown in Figs. 1(b.i) and 1(b.ii). In each figure they appear as differently colored arrows emanating from A_4. Recall that the accelerometers are rigidly fixed to the body and move with it. Therefore, the evolution of $\left({}^4\mathbf{e}_i\,[\tau]\right)_{i \in I}$ is the same as that in (30) with the body basis vector $\mathbf{e}_i\,[\tau]$ replaced with ${}^4\mathbf{e}_i\,[\tau]$, and the laboratory basis vector $\mathbf{e}_i$ replaced with the initial value of ${}^4\mathbf{e}_i\,[\cdot]$. In fact, the measurement axes of all four accelerometers, $\left({}^\ell\mathbf{e}_i\,[\tau]\right)_{i \in I}$, $\ell = 1, 2, 3, 4$, evolve similarly. Their evolution is illustrated in Fig. 1(b), and Fig. 2(b).

^{12}The dimensional form of ${}^4\mathbf{x}\,[\cdot]$ at the initial and a later time instance are shown in Fig. 1(b.i) and (b.ii), respectively.

a Rotation angle and speed

(i)

(ii)

Derivative of θ shown in (i)

Eq. (31c) for $\omega_m = 31.41$ and $\tau_1 = 5.81$

b Configurations at different time instances

(i) $\theta = 0$

(ii) $\theta = \frac{1}{5}\pi$

(iii) $\theta = \frac{2}{5}\pi$

(iv) $\theta = \frac{3}{5}\pi$

(v) $\theta = \frac{4}{5}\pi$

(vi) $\theta = \pi$

(vii) $\theta = \frac{6}{5}\pi$

(viii) $\theta = \frac{7}{5}\pi$

(ix) $\theta = \frac{8}{5}\pi$

(x) $\theta = \frac{9}{5}\pi$

Figure 2: Punctuated rotation experiment. Subfigures (a.i) and (a.ii), respectively, show the variation of the rotation angle and its derivative over time in the punctuated rotation experiment (see §5.1.2 for details). Subfigure (b) shows the configurations of the rigid body (gray cuboid) and the accelerometers (blue cuboids within the gray cuboid) attached to it for different rotation angles. The three arrows attached to the top corner of the gray cuboid denote the body basis vectors. The three arrows attached to a blue cuboid denote the accelerometer's measurement directions.

and $\lambda_2 [\tau_i]$, and $\lambda_3 [\tau_i]$ are the two smallest eigenvalues of $\widetilde{\mathbf{B}} [\tau_i]$. The matrix $\mathbf{N} [\tau_i] \in \mathcal{M}_3$ is constructed by taking the eigenvectors of $\widetilde{\mathbf{B}} [\tau_i]$ to be its columns, with the second and third columns corresponding to $\lambda_2 [\tau_i]$ and $\lambda_3 [\tau_i]$, respectively. The raw matrices $\widetilde{\mathbf{B}} [\tau_i]$ and their respective approximation $\widehat{\mathbf{B}} [\tau_i]$ are compared in Fig. 4.

The final step of the $\sqrt{\mathrm{AO}}$-algorithm involves computing the approximate body angular velocity matrix

Figure 3: Acceleration measurements in the punctuated rotation experiment (see §5.2 for details). Subfigures (a)–(d) correspond to the measurements from accelerometers A_1–A_4, respectively. The positions and orientations of the four accelerometers w.r.t. the rigid body is illustrated in Figs. 1(b) and 2(b). Each subfigure shows three graphs, labeled $i = 1$, $i = 2$, and $i = 3$. These, respectively, correspond to the acceleration components measured by the accelerometer in its 1, 2, and 3, directions (see, e.g., the three arrows attached to the cuboid in the inset in (a)).

$\widetilde{\mathbf{W}}[\tau_i]$ from the roots of $r\left[\widehat{\mathbf{B}}[\tau_i]\right][\cdot]$. For the current case the roots of $r\left[\widehat{\mathbf{B}}[\tau_i]\right][\cdot]$ come out to be

$$\widetilde{\mathbf{W}}[\tau_i] = \pm\mathbf{N}[\tau_i]\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\sqrt{-\mu_1^\star[\tau_i]} \\ 0 & \sqrt{-\mu_1^\star[\tau_i]} & 0 \end{pmatrix}\mathbf{N}[\tau_i]^\mathsf{T} \tag{35}$$

(see [16, §4.2] for details). The positive and negative signs in (35) correspond to the two roots of $r\left[\widehat{\mathbf{B}}[\tau_i]\right][\cdot]$. We choose the root that is closer to $\widetilde{\mathbf{W}}[\cdot]$'s value from the previous time step (see [16, §4.2] for details). Following (28), we compute an approximate body angular velocity vector at the time instance τ_i— $\widetilde{\mathbf{w}}[\tau_i]$—as

$$\widetilde{\mathbf{w}}[\tau_i] = \star\widetilde{\mathbf{W}}[\tau_i].$$

Graphs of the component functions of $\widetilde{\mathbf{w}}[\cdot]$ in the punctuated rotation trial are shown in Fig. 5.

5.4. Comparison

In addition to the graphs of $\widetilde{\mathbf{w}}[\cdot]$'s components, Fig. 5 also shows the graphs of the ideal body angular velocity's components, i.e. the components of $\overline{\mathbf{w}}[\cdot]$ given in (32), where $\theta[\cdot]$ is given by (31c) with $\omega_m = 31.41$ and $\tau_1 = 5.81$. The ideal angular velocity is what we expect to measure under ideal conditions, i.e., perfect rigid body motion, knowledge of experimental parameters, and acceleration measurements. As shown in Fig. 5, the angular velocity components measured using the $\sqrt{\text{AO}}$-algorithm closely align with their ideal counterparts. However, they exhibit some noise, particularly at time instances where the ideal components are non-smooth ($\tau = 2.91, 5.81, 8.72$, etc. in Fig. 5). Potential sources for this noise include accelerometer

14

measurement noise, as well as spurious dynamics introduced by artifacts in the motor's control system, chassis vibrations, and dynamic elastic deformations between the accelerometer particles.

We quantify the discrepancy between the measured and ideal angular velocities using the error metric

$$\frac{\left\|\widetilde{\overline{\mathbf{w}}}\left[\cdot\right] - \overline{\mathbf{w}}\left[\cdot\right]\right\|_{L^2(0,T)}}{\left\|\overline{\mathbf{w}}\left[\cdot\right]\right\|_{L^2(0,T)}}. \tag{36}$$

The norm $\|\cdot\|_{L^2(0,T)}$, where $T > 0$, is the standard L^2 function norm:

$$\left\|\overline{\mathbf{w}}\left[\cdot\right]\right\|_{L^2(0,T)} = \left(\int_0^T \left\|\overline{\mathbf{w}}\left[\tau\right]\right\|^2\right)^{1/2}, \tag{37}$$

and $[0,T]$ is $\overline{\mathbf{w}}\left[\cdot\right]$'s domain. Inside the integral in (37), $\|\cdot\|$ denotes the standard Euclidean norm on $\mathbb{R}^3$. For the punctuated rotation experiment, the error is 0.062 (or 6.2%). The error metrics for our oscillatory and constant rotation experiments are discussed in the §Appendix B.

For reference, we also estimated the body angular velocity using an alternative technique. In this technique the body angular velocity is estimated by time integrating the experimental body angular acceleration. It can be shown that

$$\overline{\mathbf{w}}'\left[\tau\right] = \star\left(\text{skew part of the } \mathbf{P}\left[\tau\right]\right), \tag{38}$$

where $\overline{\mathbf{w}}'\left[\cdot\right]$ is the body angular acceleration and is the derivative of $\overline{\mathbf{w}}\left[\cdot\right]$, which, of course, is the body angular velocity. The matrix valued function $\mathbf{P}\left[\cdot\right]$ in (38) is the same one that we discussed in the context of the $\sqrt{\text{AO}}$-algorithm's first step (see paragraph containing (33)). It depends on the acceleration measurements, and the relative positions and orientations of the accelerometers. For a derivation of (38) see [23, §3.2]. Following (38), we also estimated the body angular velocity in the experiment by integrating the skew part of $\widetilde{\mathbf{P}}\left[\cdot\right]$, which was obtained by applying the first step of the $\sqrt{\text{AO}}$-algorithm to that experiment. We call this alternative estimate of the body angular velocity the $\widetilde{\overline{\mathbf{w}}}\left[\cdot\right]$ *predicted from* AO-*algorithm*, or AO-$\widetilde{\overline{\mathbf{w}}}\left[\cdot\right]$ for short. The components of AO-$\widetilde{\overline{\mathbf{w}}}\left[\cdot\right]$ in the punctuated rotation experiment are graphed in Fig. 5. As observed, the discrepancies between the components of AO-$\widetilde{\overline{\mathbf{w}}}\left[\cdot\right]$ and their respective ideal components are significantly larger than those of the body angular velocity components estimated using the $\sqrt{\text{AO}}$-algorithm. Notably, the error metric for AO-$\widetilde{\overline{\mathbf{w}}}\left[\cdot\right]$ is 1.64 (or 164%), which is about 25 times greater than that of the body angular velocity estimated using the $\sqrt{\text{AO}}$-algorithm.

6. Concluding remarks

In this work, we have proved a result concerning the best approximation of square matrices within the set of matrices that are the square of a skew-symmetric matrix. An immediate application of our result is in numerical methods that aim to reproduce rigid body motion from acceleration measurements.

1. The present work uses the Frobenius norm to define a topology on the space of matrices. In other works such as [3], best approximation in other norms are considered, such as the induced ℓ^2-norm: for $\mathbf{A} \in \mathcal{M}_n$,

$$\|\mathbf{A}\|_2 := \sup_{\mathbf{x} \in \mathbb{R}^n} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|}.$$

For rigid body motion, studying the best approximation problem in other norms may yield different error properties, though we emphasize that all norms are equivalent on finite-dimensional vector spaces. Additionally, for other applications of the present problem, other norms may be more natural. Nevertheless, additional work is required to extend the results presented here to other norms, and it is not apparent whether closed-form solutions may exist for these other norms.

2. In the context of the $\sqrt{\text{AO}}$-algorithm, a natural question is how close is the estimated angular velocity matrix $\widetilde{\overline{\mathbf{W}}}$ to the actual angular velocity matrix $\overline{\mathbf{W}}$ at each time instance τ_i. In dimensions $n = 2$ and $n = 3$, we can show that when $\langle \overline{\mathbf{W}}, \widetilde{\overline{\mathbf{W}}} \rangle_F > 0$

$$\left\|\overline{\mathbf{W}} - \widetilde{\overline{\mathbf{W}}}\right\|_F^4 \leqslant C_n \left\|\mathbf{B} - \widehat{\mathbf{B}}\right\|_F^2,$$

Figure 4: Graphs of certain components of $\mathbf{B}[\cdot]$ in the punctuated rotation experiment. To avoid clutter we arbitrarily chose to only show the $(1,2)$, $(1,3)$, and $(2,2)$ components. Each component is graphed twice: once using the raw measurement for $\mathbf{B}[\cdot]$, namely $\widetilde{\mathbf{B}}[\cdot]$, which is computed directly from the acceleration measurements shown in Fig. 3, using a thick, solid, partially transparent line; and again using the best approximant of $\widetilde{\mathbf{B}}[\cdot]$, namely $\widehat{\mathbf{B}}[\cdot]$, which is obtained by applying Theorem 3 to $\widetilde{\mathbf{B}}[\cdot]$, using a thin, solid, opaque line. See §5.3 for details.

where $C_2 = 2$ and $C_3 = 8$. Additionally, because $\widehat{\mathbf{B}}$ is a best approximation of $\widetilde{\mathbf{B}}$, we have $\|\mathbf{B} - \widehat{\mathbf{B}}\|_F \leqslant 2\|\mathbf{B} - \widetilde{\mathbf{B}}\|_F$ and so

$$\left\|\overline{\mathbf{W}} - \overline{\widetilde{\mathbf{W}}}\right\|_F^4 \leqslant 4C_n \left\|\mathbf{B} - \widetilde{\mathbf{B}}\right\|_F^2.$$

Meanwhile, it may be shown via counterexample that no C_n exists for $n > 3$. Proof of these results will be presented in a forthcoming publication.

3. Since Theorem 3 makes it possible to use the $\sqrt{\text{AO}}$-algorithm on real data, we demonstrated this application by estimating the angular velocity of a rigid body using only acceleration measurements from four particles. Angular velocities measured via the $\sqrt{\text{AO}}$-algorithm and through naive time integration of angular accelerations are compared in Fig. 5. What is especially notable in Fig. 5 is that for the case in which the angular velocity is measured via time integration the errors grow with time; while in the case in which they are measured via the $\sqrt{\text{AO}}$-algorithm the errors remain bounded in time[13]. Based on these observations it is reasonable to conclude that the $\sqrt{\text{AO}}$-algorithm is a significant improvement over the naive way of measuring angular velocity using only material particle accelerations. In fact, considering the advantages of measuring angular velocity from particle accelerations, i.e., using accelerometers, rather than by using gyroscopes[14], the $\sqrt{\text{AO}}$-algorithm is potentially a significant advancement in measuring angular velocity in general. Since measurement of angular velocity is important in various fields—ranging from robotics to biomechanics—and, again, since Theorem 3 enables the $\sqrt{\text{AO}}$-algorithm's critical step, it follows that Theorem 3 and the mathematical theory that we developed in support of it are likely to a have a significant technological impact.

---

[13]There are mathematical justifications for both these observations. We will present these justifications in our forthcoming publication.

[14]We mentioned these advantages briefly in §1.1, but a thorough discussion is found in §Appendix A.

Acknowledgments

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author contributions

Y.W. carried out the research and experiments; B.E.G.P. performed the mathematical analysis; H.K. designed the research; all authors contributed to the writing of the manuscript.

Appendix A. Commercial off-the-shelf MEMS gyroscopes and accelerometers

A survey of commercially available off-the-shelf MEMS gyroscopes and accelerometers, along with their respective specifications, is presented in Table A.1.

Table A.1: Commercial off-the-shelf MEMS gyroscopes and accelerometers

Company	Product	Bandwidth (Hz)	Size (length×width×height mm^3)	Weight (g)
		Gyroscopes		
STMicroelectronics [24]	A3G4250D	110	$4 \times 4 \times 1.1$	—
	I3G4250D	110	$4 \times 4 \times 1.1$	—
	L3G3250A	140	$3.5 \times 3 \times 1$	—
	L3GD20	100	$4 \times 4 \times 1$	—
	L3G462A	110	$4 \times 4 \times 1.1$	—
	LPR410AL	110	$4 \times 5 \times 1$	—
Analog Devices [25]	ADXRS910	201	$10.30 \times 10.42 \times 3.58$	—
	ADXRS290	480	$4.5 \times 5.8 \times 1.2$	—
	ADXRS645	2000	$15 \times 8 \times 2.85$	—
	ADIS16137	400	$35.6 \times 44 \times 13.8$	—
	ADXRS642	2000	$6.85 \times 6.85 \times 3.8$	—
	ADXRS646	1000	$6.85 \times 6.85 \times 3.8$	—
	ADXRS453	77.5	$9 \times 9 \times 4$	—
	ADXRS620	2500	$6.85 \times 6.85 \times 3.8$	—
	ADXRS623	3000	$6.85 \times 6.85 \times 3.8$	—
	ADIS16060	1000	$8.35 \times 8.2 \times 5.2$	—
Tronics's Microsystems SA [26]	G4300	200	$12 \times 12 \times 5.5$	—
	G4050	30	$12 \times 12 \times 5.5$	—
	G3300	200	$19.6 \times 11.5 \times 2.9$	—
Bosch Sensortec[a][27]	BMI323	563	$2.5 \times 3.0 \times 0.83$	—

Table A.1: Commercial off-the-shelf MEMS gyroscopes and accelerometers (Continued)

	BMI270	751	$2.5 \times 3.0 \times 0.8$	—
	BMI088	523	$3 \times 4.5 \times 0.95$	—
	BMI085	684	$3 \times 4.5 \times 0.95$	—
Silicon Sensing [28]	CRM100	160	$5.7 \times 4.8 \times 1.2$	—
	CRS43	24	$29 \times 29 \times 18.4$	—
	CRH03	100	$47 \times 33.5 \times 25.4$	—
	CMS390	190	$10.4 \times 6.7 \times 2.7$	—
Gladiator Technologies [29]	G300D	600	$25.4 \times 25.4 \times 15.2$	—
	G200	200	8190 (volume)	—
	G150Z	200	18000 (volume)	—
Safran [30]	STIM210	262	$44.8 \times 38.6 \times 21.5$	52
	STIM277H	262	$44.8 \times 38.6 \times 21.5$	52
	STIM202	262	$44.75 \times 38.6 \times 20$	55
Accelerometers				
STMicroelectronics [24]	AIS328DQ	500	$4 \times 4 \times 1.8$	—
	H3LIS331DL	500	$3 \times 3 \times 1$	—
	IIS2DLPC	2500	$2 \times 2 \times 0.7$	—
	IIS3DWB	6300	$2.5 \times 3 \times 0.86$	—
	LIS2DS12	6400	$2 \times 2 \times 0.86$	—
	LIS2DUXS12	2500	$2 \times 2 \times 0.74$	—
Analog Devices [25]	ADXL1004	24000	$5 \times 5 \times 1.8$	—
	ADXL1001	11000	$5 \times 5 \times 1.8$	—
	ADXL382	8000	$2.9 \times 2.8 \times 0.87$	—
	ADXL203	2500	$5 \times 5 \times 2$	—
	ADXL354	1900	$6 \times 6 \times 2.25$	—
	ADXL316	1600	$4 \times 4 \times 1.45$	—
	ADXL355	1000	$6 \times 6 \times 2.25$	—
Tronics's Microsystems SA [26]	A3050	120	$12 \times 12 \times 5.5$	—
	A3140	300	$12 \times 12 \times 5.5$	—
	A3150	300	$12 \times 12 \times 5.5$	—
Bosch Sensortec[a][27]	BMI323	1677	$2.5 \times 3.0 \times 0.83$	—
	BMI270	684	$2.5 \times 3.0 \times 0.8$	—
	BMI088	523	$3 \times 4.5 \times 0.95$	—
	BMI085	684	$3 \times 4.5 \times 0.95$	—
Silicon Sensing [28]	CAS211	170	$10.4 \times 6 \times 2.2$	—
	CAS215	170	$10.4 \times 6 \times 2.2$	—
	CAS291	170	$10.4 \times 6.7 \times 2.7$	—
	CAS295	170	$10.4 \times 6.7 \times 2.7$	—

Table A.1: Commercial off-the-shelf MEMS gyroscopes and accelerometers (Continued)

Gladiator Technologies [29]	GA50	300	$8.9 \times 8.9 \times 3.2$	0.7
	A300D	800	$25.4 \times 25.4 \times 16.5$	19.25
	A40	140	9832 (volume)	—
Safran [30]	MS1000	200	$8.9 \times 8.9 \times 3.23$	1.5
	MS1000T	200	$8.9 \times 8.9 \times 3.23$	1.5
	SI1000	550	$8.9 \times 8.9 \times 3.23$	1.5
	VS1000	1150	$8.9 \times 8.9 \times 3.23$	1.5

[a] BMI serial product is an inertial measurement unit (IMU) which includes one triaxial accelerometer and one triaxial gyroscope.

Appendix B. Rotation test

The setup of the rotation test platform is shown in Fig. B.6.

Appendix B.1. Additional rotation tests

A sketch of the experimental setup for both constant and oscillatory rotation is shown in Fig. B.7(a.i). A drawing of a more abstract version of the rotation experiment is shown in Fig. B.7(a.ii). The setups for the constant and oscillatory rotation experiments are the same, except for the different rotation profiles used. In these two experiments, the rotation axis vector **n** was $(0, 0, 1)$. The initial values ${}^{\ell}\mathbf{x}[0]$, and $\left({}^{\ell}\mathbf{e}_i[0]\right)_{i \in I}$, $\ell = 1, 2, 3, 4$ are given in the caption of Fig. B.7.

Constant rotation. In the constant rotation experiment, the rotation angle $\theta[\cdot]$ is taken as

$$\theta[\tau] = \omega_m \tau, \tag{B.1}$$

where $\omega_m = 31.41$. Graphs of $\theta[\cdot]$ in (B.1) and its derivative are shown in Figs. B.7(b.i) and (b.ii), respectively. The comparison of the body angular velocity components calculated using $\sqrt{\text{AO}}$-algorithm and AO-algorithm in the constant rotation experiment is shown in Fig. B.8. Fig. B.8 also shows the graphs of the ideal body angular velocity's components, i.e. the components of $\overline{\mathbf{w}}[\cdot]$ given in Eq. (32) in the main manuscript, where $\theta[\cdot]$ is given by (B.1) with $\omega_m = 31.41$.

The error metric for the body angular velocity measured using $\sqrt{\text{AO}}$-algorithm in the trial of the constant rotation experiments is 0.0068 (or 0.68%), while the error metric for that using AO-algorithm is 0.166 (or 16.60%).

Oscillatory rotation. In the constant rotation experiment, the rotation angle $\theta[\cdot]$ is taken as

$$\theta[\tau] = \frac{\omega_m \tau_1}{8} + \frac{4\omega_m \tau_1}{\pi^3} \sum_{n=1,3,5} \frac{(-1)^{(n+1)/2}}{n^3} \cos\left[\frac{2n\pi\tau}{\tau_1}\right], \tag{B.2}$$

where $\omega_m = 31.41$ and $\tau_1 = 11.62$. Graphs of $\theta[\cdot]$ in (B.1) and its derivative are shown in Figs. B.7(c.i) and (c.ii), respectively. The comparison of the component functions of body angular velocity calculated using $\sqrt{\text{AO}}$-algorithm and AO-algorithm are shown in Fig. B.9. Fig. B.9 also shows the graphs of the ideal body angular velocity's components, i.e. the components of $\overline{\mathbf{w}}[\cdot]$ given in Eq. (32) in the main manuscript, where $\theta[\cdot]$ is given by (B.2) with $\omega_m = 31.41$ and $\tau_1 = 11.62$.

The error metric for the body angular velocity measured via the $\sqrt{\text{AO}}$-algorithm in the trial of the oscillatory rotation experiments is 0.038 (or 3.80%), while via AO-algorithm is 1.0956 (or 109.56%).

Appendix C. Proofs of preliminary results

For completeness, we present proofs of the preliminary results below.

Appendix C.1. Proof of Lemma 4

Lemma 4. For $\mathbf{S} \in \mathcal{S}_n$ and $\mathbf{K} \in \mathcal{K}_n$,

$$\|\mathbf{S} + \mathbf{K}\|_F^2 = \|\mathbf{S}\|_F^2 + \|\mathbf{K}\|_F^2.$$

Proof.

$$
\begin{aligned}
\|\mathbf{S} + \mathbf{K}\|_F^2 &= \langle \mathbf{S} + \mathbf{K}, \mathbf{S} + \mathbf{K} \rangle_F, \\
&= \|\mathbf{S}\|_F^2 + \|\mathbf{K}\|_F^2 + 2\langle \mathbf{S}, \mathbf{K} \rangle_F.
\end{aligned}
\tag{C.1}
$$

From the definition of the Frobenius inner product one can see that $\langle \mathbf{A}, \mathbf{B} \rangle_F = \langle \mathbf{A}^\mathsf{T}, \mathbf{B}^\mathsf{T} \rangle_F$. Hence,

$$
\begin{aligned}
\langle \mathbf{S}, \mathbf{K} \rangle_F &= \langle \mathbf{S}^\mathsf{T}, \mathbf{K}^\mathsf{T} \rangle_F, \\
&= \langle \mathbf{S}, -\mathbf{K} \rangle_F, \\
&= -\langle \mathbf{S}, \mathbf{K} \rangle_F,
\end{aligned}
$$

where the second equality is due to $\mathbf{K}$ being a skew-symmetric matrix. Therefore, $\langle \mathbf{S}, \mathbf{K} \rangle_F = 0$, and the result follows from (C.1).

$\square$

Appendix C.2. Proof of Lemma 5

Lemma 5. Let $\mathbf{A} \in \mathcal{M}_n$ and $\mathbf{Q} \in \mathcal{O}_n$. Then

$$\|\mathbf{AQ}\|_F = \|\mathbf{A}\|_F = \|\mathbf{QA}\|_F.$$

Proof. The Frobenius norm can be written in terms of the trace map $\mathsf{Tr}\,[\cdot] : \mathcal{M}_n \to \mathbb{R}$,

$$\mathsf{Tr}\,[\mathbf{A}] = \sum_{i=1}^{n} \mathbf{A}_{\cdot i \cdot i},$$

as

$$\|\mathbf{A}\|_F = \sqrt{\mathsf{Tr}\left[\mathbf{A}\mathbf{A}^\mathsf{T}\right]}.
\tag{C.2}$$

To prove the first equality in the lemma, we express $\|\mathbf{AQ}\|_F^2$ using the trace map as $\mathsf{Tr}\left[\mathbf{AQ}(\mathbf{AQ})^\mathsf{T}\right]$. Which, owing to the property of the transpose map that $(\mathbf{AQ})^\mathsf{T} = \mathbf{Q}^\mathsf{T}\mathbf{A}^\mathsf{T}$, simplifies to $\mathsf{Tr}\left[\mathbf{AQQ}^\mathsf{T}\mathbf{A}^\mathsf{T}\right]$. Since we know from Eq. (4) of the main manuscript that $\mathbf{QQ}^\mathsf{T} = \mathbf{I}_n$, this expression further simplifies to $\mathsf{Tr}\left[\mathbf{AA}^\mathsf{T}\right] = \|\mathbf{A}\|_F^2$. The result follows on taking the positive square root of this equation.

For showing the second equality in the lemma we follow a similar strategy and write $\|\mathbf{QA}\|_F^2$ in terms of the trace map as $\mathsf{Tr}\left[\mathbf{QA}(\mathbf{QA})^\mathsf{T}\right]$, or equivalently as $\mathsf{Tr}\left[\mathbf{QAA}^\mathsf{T}\mathbf{Q}^\mathsf{T}\right]$. The trace map has the property that $\mathsf{Tr}\,[\mathbf{AB}] = \mathsf{Tr}\,[\mathbf{BA}]$, where $\mathbf{A}, \mathbf{B} \in \mathcal{M}_n$. Using this property and commuting the composition of the first three matrices, namely $\mathbf{QAA}^\mathsf{T}$, and the last matrix, namely $\mathbf{Q}^\mathsf{T}$, in $\mathsf{Tr}\left[\mathbf{QAA}^\mathsf{T}\mathbf{Q}^\mathsf{T}\right]$ we get that $\|\mathbf{QA}\|_F^2 = \mathsf{Tr}\left[\mathbf{Q}^\mathsf{T}\mathbf{QAA}^\mathsf{T}\right]$. This equation simplifies to $\|\mathbf{QA}\|_F^2 = \|\mathbf{A}\|_F^2$, since $\mathbf{Q}^\mathsf{T}\mathbf{Q} = \mathbf{I}_n$, and $\mathsf{Tr}\left[\mathbf{AA}^\mathsf{T}\right] = \|\mathbf{A}\|_F^2$. The result follows on taking the positive square root of this equation.

$\square$

Appendix C.3. Proof of Lemma 7

Lemma 7. The sets $\mathcal{V}_n$ and $\mathcal{K}_n^{\mathrm{sq}}$ are the same.

Proof. Proof that $\mathcal{K}_n^{\mathrm{sq}} \subseteq \mathcal{V}_n$ may be found in [16, Appendix C]. The proof relies on the fact that the eigendecomposition of $\mathbf{K} \in \mathcal{K}_n$ has pure imaginary eigenvalues (which come in conjugate pairs) and orthogonal (complex) eigenvectors. Below we show that $\mathcal{V}_n \subseteq \mathcal{K}_n^{\mathrm{sq}}$. (This result can also be found in [31].)

1. Let $\mathbf{A} \in \mathcal{V}_n$, i.e., $\mathbf{A} = \mathbf{N}\mathbf{D}\mathbf{N}^{\mathsf{T}}$, where $\mathbf{N} \in \mathcal{O}_n$, when n is even (resp. odd) $\mathbf{D} = \mathrm{diag}_n\left[\mu_1, \mu_1, \ldots, \mu_k, \mu_k\right]$ (resp. $\mathbf{D} = \mathrm{diag}_n\left[0, \mu_1, \mu_1, \ldots, \mu_k, \mu_k\right]$), and $k = n/2$ (resp. $k = (n-1)/2$).

2. When n is even (resp. odd) we define

$$\mathbf{K} = \mathrm{diag}\left[\begin{pmatrix} 0 & -\sqrt{-\mu_1} \\ \sqrt{-\mu_1} & 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 & -\sqrt{-\mu_k} \\ \sqrt{-\mu_k} & 0 \end{pmatrix}\right] \text{[15]} \tag{C.3a}$$

$$\left(\text{resp. } \mathbf{K} = \mathrm{diag}\left[(0), \begin{pmatrix} 0 & -\sqrt{-\mu_1} \\ \sqrt{-\mu_1} & 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 & -\sqrt{-\mu_k} \\ \sqrt{-\mu_k} & 0 \end{pmatrix}\right]\right). \tag{C.3b}$$

It is straightforward to show that $\mathbf{K}^{\mathsf{T}} = -\mathbf{K}$, and hence $\mathbf{K} \in \mathcal{K}_n$. Moreover, a direct calculation[16] gives $\mathbf{K}^2 = \mathbf{D}$.

3. Let $\mathbf{B} := \mathbf{N}\mathbf{K}\mathbf{N}^{\mathsf{T}}$, then

$$\mathbf{B}^{\mathsf{T}} = (\mathbf{N}\mathbf{K}\mathbf{N}^{\mathsf{T}})^{\mathsf{T}} = \mathbf{N}\mathbf{K}^{\mathsf{T}}\mathbf{N}^{\mathsf{T}} = -\mathbf{N}\mathbf{K}\mathbf{N}^{\mathsf{T}} = -\mathbf{B}. \tag{C.6}$$

The third equality in (C.6) follows from the fact that $\mathbf{K}$ is skew-symmetric. Furthermore,

$$\mathbf{B}^2 = (\mathbf{N}\mathbf{K}\mathbf{N}^{\mathsf{T}})(\mathbf{N}\mathbf{K}\mathbf{N}^{\mathsf{T}}) = \mathbf{N}\mathbf{K}^2\mathbf{N}^{\mathsf{T}} = \mathbf{N}\mathbf{D}\mathbf{N}^{\mathsf{T}} = \mathbf{A}. \tag{C.7}$$

The second equality in (C.7) follows from the fact that $\mathbf{N}$ is orthogonal; the third from the fact that $\mathbf{K}^2 = \mathbf{D}$ (see Step 2); and the last equality from $\mathbf{A}$'s definition. Since we know that $\mathbf{B} \in \mathcal{K}_n$, from (C.6), and $\mathbf{B}^2 = \mathbf{A}$, from (C.7), it follows from $\mathcal{K}_n^{\mathrm{sq}}$'s definition that $\mathbf{A} \in \mathcal{K}_n^{\mathrm{sq}}$.

□

We briefly comment on the uniqueness of the matrix $\mathbf{B}$ in the proof of Lemma 7. In particular, this matrix may not be unique. Taking the $\mathbf{K}$ given in (C.3), multiplying one or more of its diagonal blocks with -1, and performing an orthogonal similarity transformation on it using $\mathbf{N}$ will yield a matrix which too lies in $\mathcal{K}_n$ and squares to $\mathbf{A}$.

Appendix C.4. Proof of Lemma 8

Lemma 8. Let $\mathbf{\Lambda} = \mathrm{diag}_n\left[\lambda_1, \lambda_2, \ldots, \lambda_n\right]$ where $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$, and let $\mathbf{D} = \mathrm{diag}_n\left[\mu_1, \mu_2, \ldots, \mu_n\right]$ where $\mu_1 \geqslant \mu_2 \geqslant \ldots \geqslant \mu_n$. Then,

$$\min_{\mathbf{Q} \in \mathcal{O}_n} \|\mathbf{\Lambda} - \mathbf{Q}\mathbf{D}\mathbf{Q}^{\mathsf{T}}\|_F = \|\mathbf{\Lambda} - \mathbf{D}\|_F. \tag{11}$$

---

[15]For square matrices $\mathbf{A}_1 \in \mathcal{M}_{n_1}, \ldots, \mathbf{A}_m \in \mathcal{M}_{n_m}$ (which may have different sizes $n_1, \ldots, n_m$, respectively), we let $\mathrm{diag}\left[\mathbf{A}_1, \ldots, \mathbf{A}_m\right]$ denote the block diagonal matrix

$$\begin{pmatrix} \mathbf{A}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_m \end{pmatrix} \in \mathcal{M}_{\sum_i n_i}, \tag{C.4}$$

where the $\mathbf{0}$ entries denote matrices whose entries are all zero (and whose dimensions we infer by context).

[16]A useful identity for carrying out this calculation is

$$\begin{pmatrix} 0 & -\mu \\ \mu & 0 \end{pmatrix}^2 = \begin{pmatrix} -\mu^2 & 0 \\ 0 & -\mu^2 \end{pmatrix}, \tag{C.5}$$

where $\mu \in \mathbb{R}$.

Proof. Let $\mathbf{Q} \in \mathcal{O}_n$. The matrices $\boldsymbol{\Lambda}$ and $\mathbf{Q}\mathbf{D}\mathbf{Q}^\mathsf{T}$ are both symmetric and have real eigenvalues $(\lambda_1, \ldots, \lambda_n)$ and $(\mu_1, \ldots, \mu_n)$ arranged in descending order. Applying [32, Corollary 6.3.8] (where, using the notation from the reference, we take matrices $\mathbf{A}$ and $\mathbf{A} + \mathbf{E}$ to be $\boldsymbol{\Lambda}$ and $\mathbf{Q}\mathbf{D}\mathbf{Q}^\mathsf{T}$, respectively) gives

$$\sum_{i=1}^{n} (\lambda_i - \mu_i)^2 \leqslant \|\boldsymbol{\Lambda} - \mathbf{Q}\mathbf{D}\mathbf{Q}^\mathsf{T}\|_F^2.$$

The left-hand-side is identically $\|\boldsymbol{\Lambda} - \mathbf{D}\|_F^2$. Since this inequality holds for any $\mathbf{Q} \in \mathcal{O}_n$ (and is an equality in the particular case when $\mathbf{Q} = \mathbf{I}_n$), the proof follows. $\qquad\square$

Appendix C.5. Alternate Proof of Lemma 8

We also present a second, longer proof which does not make use of [32, Corollary 6.3.8]. Before stating the proof, we first discuss an isomorphism between the space of matrices and the space of n-tuples of vectors.

The Euclidean space $\widetilde{\mathcal{M}}_n$. Consider the set $\widetilde{\mathcal{M}}_n := \overbrace{\mathbb{R}^n \times \mathbb{R}^n \times \ldots \mathbb{R}^n}^{n\text{-times}}$. Let $\widetilde{\mathbf{X}} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ and $\widetilde{\mathbf{Y}} = (\mathbf{y}_1, \ldots, \mathbf{y}_n)$ belong to $\widetilde{\mathcal{M}}_n$. Define the function $\langle \cdot, \cdot \rangle_\sim : \widetilde{\mathcal{M}}_n \times \widetilde{\mathcal{M}}_n \to \mathbb{R}$,

$$\langle \widetilde{\mathbf{X}}, \widetilde{\mathbf{Y}} \rangle_\sim = \sum_{i=1}^{n} \langle \mathbf{x}_i, \mathbf{y}_i \rangle. \tag{C.8}$$

It can be shown that $\widetilde{\mathcal{M}}_n$ together with $\langle \cdot, \cdot \rangle_\sim$ is a Euclidean vector space. In fact, the inner product spaces $(\mathcal{M}_n, \langle \cdot, \cdot \rangle_F)$ and $\left(\widetilde{\mathcal{M}}_n, \langle \cdot, \cdot \rangle_\sim \right)$ are isomorphic. For example, there exists the bijective, linear, inner product preserving map $\mathtt{iso}\,[\cdot] : \mathcal{M}_n \to \widetilde{\mathcal{M}}_n$,

$$\mathtt{iso}\,[\mathbf{X}] = \left((X_{j1})_{j\in\mathcal{I}}, (X_{j2})_{j\in\mathcal{I}}, \ldots, (X_{jn})_{j\in\mathcal{I}} \right), \tag{C.9}$$

where X_{ij} is the i-j^{th} component of $\mathbf{X}$. Note that the map $\mathtt{iso}\,[\cdot]$ is also a homeomorphism. The inverse of $\mathtt{iso}\,[\cdot]$ can be explicitly written as

$$\mathtt{iso}^{-1}\,[(\mathbf{y}_1, \ldots, \mathbf{y}_n)] = (\mathbf{y}_i)_{i\in\mathcal{I}}^\mathsf{T}. \tag{C.10}$$

Proof.

1. We begin by showing the minimization problem in (11) is equivalent to maximizing $g\,[\cdot] : \mathcal{O}_n \to \mathbb{R}$, where

$$g\,[\mathbf{Y}] = \langle \boldsymbol{\Lambda}, \mathbf{Y}\mathbf{D}\mathbf{Y}^\mathsf{T} \rangle_F.$$

   Note that the function $g\,[\cdot]$ is continuous, and that its domain, namely $\mathcal{O}_n$, is a compact subset[17] of the Euclidean vector space $\mathcal{M}_n$. Therefore, there exists a $\mathbf{Y}^\star \in \mathcal{O}_n$ such that

$$\max_{\mathbf{Y}\in\mathcal{O}_n} g\,[\mathbf{Y}] = g\,[\mathbf{Y}^\star].$$

   (1.*a*) Via matrix algebra and the definition of the Frobenius inner product and norm,

$$\begin{aligned}
\|\boldsymbol{\Lambda} - \mathbf{Q}\mathbf{D}\mathbf{Q}^\mathsf{T}\|_F^2 &= \langle \boldsymbol{\Lambda} - \mathbf{Q}\mathbf{D}\mathbf{Q}^\mathsf{T}, \boldsymbol{\Lambda} - \mathbf{Q}\mathbf{D}\mathbf{Q}^\mathsf{T} \rangle_F, \\
&= \|\boldsymbol{\Lambda}\|_F^2 + \|\mathbf{Q}\mathbf{D}\mathbf{Q}^\mathsf{T}\|_F^2 - 2\langle \boldsymbol{\Lambda}, \mathbf{Q}\mathbf{D}\mathbf{Q}^\mathsf{T} \rangle_F, \\
&= \|\boldsymbol{\Lambda}\|_F^2 + \|\mathbf{D}\|_F^2 - 2\langle \boldsymbol{\Lambda}, \mathbf{Q}\mathbf{D}\mathbf{Q}^\mathsf{T} \rangle_F, \\
&= \|\boldsymbol{\Lambda}\|_F^2 + \|\mathbf{D}\|_F^2 - 2g\,[\mathbf{Q}].
\end{aligned} \tag{C.11}$$

   For the third equality, we used Lemma 5 to equate $\|\mathbf{Q}\mathbf{D}\mathbf{Q}^\mathsf{T}\|_F$ with $\|\mathbf{D}\|_F$.

---

[17]For any $\mathbf{Q} \in \mathcal{O}_n$, we have $\|\mathbf{Q}\|_F = \sqrt{n}$, which follows from Eq. (1.2) of the main manuscript. Therefore $\mathcal{O}_n$ is bounded. The set $\mathcal{O}_n$ is also closed since it is the pre-image of the closed set $\{\mathbf{0}\} \subset \mathcal{M}_n$ under the continuous map $\mathcal{M}_n \ni \mathbf{Y} \mapsto (\mathbf{Y}^\mathsf{T}\mathbf{Y} - \mathbf{I}_n) \in \mathcal{M}_n$, where $\mathbf{0}$ is the matrix whose entries are all zero.

(1.*b*) It follows from (C.11) and the definition of $\mathbf{Y}^\star$ that

$$\min_{\mathbf{Q}\in\mathcal{O}_n} \|\mathbf{\Lambda} - \mathbf{Q}\mathbf{D}\mathbf{Q}^\mathsf{T}\|_F = \|\mathbf{\Lambda} - \mathbf{Y}^\star\mathbf{D}\,\mathbf{Y}^{\star\,\mathsf{T}}\|_F. \tag{C.12}$$

2. Completing the proof now requires us to show that $g\,[\mathbf{I}_n] = g\,[\mathbf{Y}^\star]$, i.e. $\mathbf{I}_n$ is a maximizer of $g\,[\cdot]$. For ease of computations, we will instead show that $\tilde{g}\,[\cdot] = g \circ \mathtt{iso}^{-1}\,[\cdot]$, which is defined over the set

$$\widetilde{\mathcal{O}}_n = \Big\{(\mathbf{y}_1,\ldots,\mathbf{y}_n) \in \widetilde{\mathcal{M}}_n \ : \ \langle\mathbf{y}_i,\mathbf{y}_j\rangle = \delta_{ij},\ i,j\in\mathcal{I}\Big\},$$

achieves its maximum value at $\widetilde{\mathbf{I}}_n = \mathtt{iso}[\mathbf{I}_n]$. By the use of isomorphism, $\mathbf{I}_n$ is also a maximizer of $g\,[\cdot]$.

(2.*a*) A more explicit form of $\tilde{g}\,[\cdot]$ is

$$\tilde{g}\,[(\mathbf{y}_1,\ldots,\mathbf{y}_n)] = \big\langle\mathbf{\Lambda}, \mathtt{iso}^{-1}\,[(\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_n)]\,\mathbf{D}\,\mathtt{iso}^{-1}\,[(\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_n)]^\mathsf{T}\big\rangle_F,$$
$$= \sum_{i=1}^{n}\mu_i\langle\mathbf{y}_i,\mathbf{\Lambda}\mathbf{y}_i\rangle, \tag{C.13}$$

where we recall that each $\mathbf{y}_i \in \mathbb{R}^n$.

3. We first show Step 2 for the case where $\lambda_1 > \lambda_2 > \ldots > \lambda_n$ and $\mu_1 > \mu_2 > \ldots > \mu_n$ and $\lambda_i, \mu_i \neq 0$ for all $i \in \mathcal{I}$.

(3.*a*) Let $\widetilde{\mathbf{Y}}^\star$ be a global maximum of $\tilde{g}\,[\cdot]$. The set $\widetilde{\mathcal{O}}_n$ is a smooth manifold and $\tilde{g}\,[\cdot]$ is smooth everywhere on $\widetilde{\mathcal{O}}_n$. Hence, $\widetilde{\mathbf{Y}}^\star$ is a stationary point of $\tilde{g}\,[\cdot]$.

(3.*b*) A necessary condition for $(\mathbf{y}_1^\star,\mathbf{y}_2^\star,\ldots,\mathbf{y}_n^\star) := \widetilde{\mathbf{Y}}^\star$ to be a stationary point of $\tilde{g}\,[\cdot]$ is that for each $i \in \mathcal{I}$

$$\partial_i\tilde{g}\left[\widetilde{\mathbf{Y}}^\star\right] + \sum_{m=1}^{n}\sum_{k=1}^{n}L_{mk}\partial_i H_{mk}\left[\widetilde{\mathbf{Y}}^\star\right] = \mathbf{0} \in \mathbb{R}^n, \tag{C.14a}$$

where

$$H_{ij}\,[\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_n] := \langle\mathbf{y}_i,\mathbf{y}_j\rangle - \delta_{ij}, \tag{C.14b}$$

$i,j \in \mathcal{I}$. In (C.14a), the quantities $\partial_i\tilde{g}\left[\widetilde{\mathbf{Y}}^\star\right]$, $\partial_i H_{mk}\left[\widetilde{\mathbf{Y}}^\star\right]$, respectively, denote

$$\frac{\partial\tilde{g}\left[\widetilde{\mathbf{Y}}^\star\right]}{\partial\mathbf{y}_i} := \left(\frac{\partial\tilde{g}\left[\widetilde{\mathbf{Y}}^\star\right]}{\partial\widetilde{\mathbf{y}}_{i\cdot1}},\ldots,\frac{\partial\tilde{g}\left[\widetilde{\mathbf{Y}}^\star\right]}{\partial\,\mathbf{y}_{i\cdot n}}\right), \tag{C.14c}$$

$$\frac{\partial H_{mk}\left[\widetilde{\mathbf{Y}}^\star\right]}{\partial\mathbf{y}_i} := \left(\frac{\partial H_{mk}\left[\widetilde{\mathbf{Y}}^\star\right]}{\partial\mathbf{y}_{i\cdot1}},\ldots,\frac{\partial H_{mk}\left[\widetilde{\mathbf{Y}}^\star\right]}{\partial\mathbf{y}_{i\cdot n}}\right). \tag{C.14d}$$

The quantities L_{mk}, $m,k \in \mathcal{I}$, in (C.14a) are components of a real symmetric matrix. That is, $L_{mk} = \mathbf{L}_{\cdot m\cdot k}$, where $\mathbf{L} \in \mathcal{S}_n$.

(3.*c*) It follows from (C.13) and (C.14) that

$$2\mu_i\mathbf{\Lambda}\mathbf{y}_i^\star - 2\sum_{k=1}^{n}L_{ik}\mathbf{y}_k^\star = \mathbf{0}, \tag{C.15}$$

holds for each $i \in \mathcal{I}$. Taking the inner product of (C.15) with $\mathbf{y}_j^\star$, noting that $\langle\mathbf{y}_k^\star,\mathbf{y}_j^\star\rangle = \delta_{kj}$, and simplifying we get that

$$L_{ij} = \mu_i\langle\mathbf{y}_j^\star,\mathbf{\Lambda}\mathbf{y}_i^\star\rangle, \tag{C.16}$$

where $i,j \in \mathcal{I}$. Computing L_{ij} and L_{ji} from (C.16) and setting them equal to each other, since they are the components of a symmetric matrix, we get that

$$\mu_i\langle\mathbf{y}_j^\star,\mathbf{\Lambda}\mathbf{y}_i^\star\rangle = \mu_j\langle\mathbf{y}_i^\star,\mathbf{\Lambda}\mathbf{y}_j^\star\rangle,$$
$$= \mu_j\langle\mathbf{\Lambda}^\mathsf{T}\mathbf{y}_i^\star,\mathbf{y}_j^\star\rangle,$$
$$= \mu_j\langle\mathbf{\Lambda}\mathbf{y}_i^\star,\mathbf{y}_j^\star\rangle. \tag{C.17}$$

24

The second equality in (C.17) follows from the definition of the matrix transpose. The third equality in (C.17) follows because $\mathbf{\Lambda}$ is a diagonal matrix, and hence a symmetric matrix. Thus, by the symmetry of the inner products, $(\mu_i - \mu_j)\langle \mathbf{y}_j^\star, \mathbf{\Lambda}\mathbf{y}_i^\star \rangle = 0$ for any $i, j \in \mathcal{I}$. By assumption, $\mu_i \neq \mu_j$ when $i \neq j$, and so

$$\langle \mathbf{y}_j^\star, \mathbf{\Lambda}\mathbf{y}_i^\star \rangle = 0, \tag{C.18}$$

for $i \neq j$. Along with (C.16), we must have that $L_{ij} = 0$ for $i \neq j$.

(3.d) Given $L_{ij} = 0$ for $i \neq j$ and $L_{ii} = \mu_i \langle \mathbf{y}_i^\star, \mathbf{\Lambda}\mathbf{y}_i^\star \rangle$, (C.15) may be simplified to

$$2\mu_i \left(\mathbf{I}_n - \mathbf{y}_i^\star \otimes \mathbf{y}_i^\star \right) \mathbf{\Lambda}\mathbf{y}_i^\star = \mathbf{0}, \tag{C.19}$$

where $\mathbf{y}_i^\star \otimes \mathbf{y}_i^\star \in \mathcal{M}_n$ such that $(\mathbf{y}_i^\star \otimes \mathbf{y}_i^\star)_{\cdot m \cdot n} = \mathbf{y}_{i \cdot m}^\star \mathbf{y}_{i \cdot n}^\star$. Because $\mu_i \neq 0$, it must either be the case that $\mathbf{\Lambda}\mathbf{y}_i^\star = \mathbf{0}$, or $\mathbf{\Lambda}\mathbf{y}_i^\star$ and $\mathbf{y}_i^\star$ are parallel. We can rule out the former: because $\lambda_j \neq 0$ for all $j \in \mathcal{I}$, then $\mathbf{\Lambda}\mathbf{y}_i^\star = \mathbf{0}$ implies that $\langle \mathbf{y}_i^\star, \mathbf{y}_i^\star \rangle = 0$, which contradicts the fact that $(\mathbf{y}_1^\star, \ldots, \mathbf{y}_n^\star) \in \tilde{\mathcal{O}}_n$ and therefore $\langle \mathbf{y}_i^\star, \mathbf{y}_i^\star \rangle = 1$. Thus,

$$\mathbf{\Lambda}\mathbf{y}_i^\star = \alpha \mathbf{y}_i^\star, \tag{C.20}$$

for some $\alpha \in \mathbb{R}$.

(3.e) Let $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ be the canonical set of basis vectors of $\mathbb{R}^n$. That is, for each $i \in \mathcal{I}$, $\mathbf{e}_i = (\delta_{ik})_{k \in \mathcal{I}}$. Since $\lambda_i \neq \lambda_j$ for $i \neq j$ and $\langle \mathbf{y}_i^\star, \mathbf{y}_i^\star \rangle = 1$, then (C.20) implies that for each $i \in \mathcal{I}$ the vector $\mathbf{y}_i^\star$ has to be equal to $\pm \mathbf{e}_j$ where $j \in \mathcal{I}$. The constraint that $\langle \mathbf{y}_i^\star, \mathbf{y}_j^\star \rangle = 0$ for $i \neq j$ implies that two different $\mathbf{y}_i^\star$ cannot correspond to the same canonical basis vector. Hence, it follows that $(\mathbf{y}_1^\star, \ldots, \mathbf{y}_n^\star)$ has to be of the form $\left(\pm \mathbf{e}_{\sigma[1]}, \pm \mathbf{e}_{\sigma[2]}, \ldots, \pm \mathbf{e}_{\sigma[n]} \right)$, where $\sigma[\cdot]$ is a permutation on $\mathcal{I}$.

(3.f) Thus, the stationary points of $\tilde{g}[\cdot]$, which includes the global maximizer $(\mathbf{y}_1^\star, \ldots, \mathbf{y}_n^\star)$, belong to the finite set

$$\left\{ \left(\pm \mathbf{e}_{\sigma[1]}, \pm \mathbf{e}_{\sigma[2]}, \ldots, \pm \mathbf{e}_{\sigma[n]} \right) \ : \ \sigma[\cdot] \text{ is a permutation on } \mathcal{I} \right\}. \tag{C.21}$$

The image of this set under $\tilde{g}[\cdot]$ is

$$\left\{ \sum_{i=1}^n \mu_i \lambda_{\sigma[i]} \ : \ \sigma[\cdot] \text{ is a permutation on } \mathcal{I} \right\}, \tag{C.22}$$

which must contain the maximum value of $\tilde{g}[\cdot]$. The maximal element of this finite set is $\sum_i \lambda_i \mu_i$[18].

(3.g) The maximal element of the set (C.22) corresponds to the identity permutation. Hence, it follows from (C.21) that the set of all $\widetilde{\mathbf{Y}}^\star$, i.e., the set of all global maximizers of $\tilde{g}[\cdot]$, equals

$$\tilde{S}_n^\star = \{ (\pm \mathbf{e}_1, \pm \mathbf{e}_2, \ldots, \pm \mathbf{e}_n) \}, \tag{C.23}$$

among which is $\widetilde{\mathbf{I}}_n$.

4. In this step we relax the conditions we required on λ_i in Step 3, and assume only that λ_i are non-increasing. However, we still require μ_i to be strictly decreasing and non-zero.

(4.a) For any $\varepsilon > 0$, define $\lambda_i(\varepsilon) = \lambda_i + \varepsilon(n + 1 - i)$, and let $\mathbf{\Lambda}_\varepsilon = \mathsf{diag}_n \left[\lambda_1(\varepsilon), \lambda_2(\varepsilon), \ldots \right]$. When ε is sufficiently small, we guarantee that $\lambda_i(\varepsilon) > \lambda_{i+1}(\varepsilon)$ and $\lambda_i(\varepsilon) \neq 0$ hold for all i. That is, for ε sufficiently small $\lambda_i(\varepsilon)$ satisfy the same conditions that λ_i did in Step 3.

(4.b) Define $\tilde{g}_\varepsilon[\cdot] : \tilde{\mathcal{O}}_n \to \mathbb{R}$ as

$$\tilde{g}_\varepsilon[(\mathbf{y}_1, \ldots, \mathbf{y}_n)] = \sum_{i=1}^n \mu_i \langle \mathbf{y}_i, \mathbf{\Lambda}_\varepsilon \mathbf{y}_i \rangle. \tag{C.24}$$

It follows from (C.13) that

$$\tilde{g}_\varepsilon\left[\widetilde{\mathbf{Y}} \right] = \tilde{g}\left[\widetilde{\mathbf{Y}} \right] + \varepsilon \sum_{i=1}^n \mu_i \langle \mathbf{y}_i, \mathsf{diag}_n\left[n, n-1, \ldots, 1 \right] \mathbf{y}_i \rangle,$$

---

[18]This is because $(\lambda_i)_{i \in \mathcal{I}}$ and $(\mu_i)_{i \in \mathcal{I}}$ are descending lists of numbers. To elaborate further note that if $a_1 \geqslant a_2$ and $b_1 \geqslant b_2$, then $(a_1 - a_2)(b_1 - b_2) \geqslant 0$ implies that $a_1 b_1 + a_2 b_2 \geqslant a_1 b_2 + a_2 b_1$.

where we recall that $\widetilde{\mathbf{Y}} := (\mathbf{y}_1, \ldots, \mathbf{y}_n)$. For the second term, we have

$$\varepsilon \sum_{i=1}^{n} \mu_i \langle \mathbf{y}_i, \mathsf{diag}_n \left[n, n-1, \ldots, 1 \right] \mathbf{y}_i \rangle \leqslant \varepsilon \sum_{i=1}^{n} |\mu_i| n \langle \mathbf{y}_i, \mathbf{y}_i \rangle = \varepsilon n \sum_{i=1}^{n} |\mu_i|,$$

which is a constant that is independent of $\widetilde{\mathbf{Y}}$. Hence,

$$\left| \tilde{g}_\varepsilon \left[\widetilde{\mathbf{Y}} \right] - \tilde{g} \left[\widetilde{\mathbf{Y}} \right] \right| \leqslant \varepsilon n \sum_{i=1}^{n} |\mu_i|. \tag{C.25}$$

(4.c) Recall that $\widetilde{\mathbf{I}}_n$ belongs to the set $\tilde{S}_n^\star$ defined in (C.23). Then, for any $\widetilde{\mathbf{Y}}$

$$\tilde{g} \left[\widetilde{\mathbf{I}}_n \right] - \tilde{g} \left[\widetilde{\mathbf{Y}} \right] = \left(\tilde{g} \left[\widetilde{\mathbf{I}}_n \right] - \tilde{g}_\varepsilon \left[\widetilde{\mathbf{I}}_n \right] \right) + \left(\tilde{g}_\varepsilon \left[\widetilde{\mathbf{I}}_n \right] - \tilde{g}_\varepsilon \left[\widetilde{\mathbf{Y}} \right] \right) + \left(\tilde{g}_\varepsilon \left[\widetilde{\mathbf{Y}} \right] - \tilde{g} \left[\widetilde{\mathbf{Y}} \right] \right). \tag{C.26}$$

Using (C.25) for the first and third terms, we bound the previous equation from below:

$$\tilde{g} \left[\widetilde{\mathbf{I}}_n \right] - \tilde{g} \left[\widetilde{\mathbf{Y}} \right] \geqslant \left(\tilde{g}_\varepsilon \left[\widetilde{\mathbf{I}}_n \right] - \tilde{g}_\varepsilon \left[\widetilde{\mathbf{Y}} \right] \right) - 2\varepsilon n \sum_{i=1}^{n} |\mu_i|. \tag{C.27}$$

Note that $\tilde{g}_\varepsilon \left[\cdot \right]$ satisfies all the conditions imposed on $g \left[\cdot \right]$ in Step 3, and hence $\tilde{g}_\varepsilon \left[\cdot \right]$ achieves its global maxima at $\widetilde{\mathbf{I}}_n$. That is, $\tilde{g}_\varepsilon \left[\widetilde{\mathbf{I}}_n \right] \geqslant \tilde{g}_\varepsilon \left[\widetilde{\mathbf{Y}} \right]$ for all $\widetilde{\mathbf{Y}} \in \tilde{\mathcal{O}}_n$. Then it follows from (C.27) that

$$\tilde{g} \left[\widetilde{\mathbf{I}}_n \right] - \tilde{g} \left[\widetilde{\mathbf{Y}} \right] \geqslant -2\varepsilon n \sum_{i=1}^{n} |\mu_i|. \tag{C.28}$$

Since $\varepsilon > 0$ can be arbitrarily small, the above inequality implies that

$$\tilde{g} \left[\widetilde{\mathbf{I}}_n \right] - \tilde{g} \left[\widetilde{\mathbf{Y}} \right] \geqslant 0. \tag{C.29}$$

We could have chosen any $\widetilde{\mathbf{Y}}^\star \in \tilde{S}_n^\star$ and arrived at the same conclusion; thus we can also conclude that $\tilde{S}_n^\star$ is the set of global maximizers of $\tilde{g} \left[\cdot \right]$ even in this more general case.

5. Finally, we relax the conditions on μ_i as well. That is, we consider a $\tilde{g} \left[\cdot \right]$ in which we only require that λ_i and μ_i each be a non-increasing sequence. The procedure we follow is similar to the one in the previous step.

(5.a) As we did for λ_i in Step 4, we construct a perturbed sequence $\mu_i(\varepsilon)$, where $\mu_i(\varepsilon) := \mu_i + \varepsilon(n+1-i)$. Again, for ε sufficiently small $\mu_i(\varepsilon)$ will be a strictly decreasing sequence and $\mu_i(\varepsilon) \neq 0$ for each $i \in \mathcal{I}$. We define $\mathbf{D}(\varepsilon) = \mathsf{diag}_n \left[\mu_1(\varepsilon), \ldots, \mu_n(\varepsilon) \right]$.

(5.b) Let $\tilde{g}_\epsilon \left[\cdot \right] : \tilde{\mathcal{O}}_n \to \mathbb{R}$,

$$\tilde{g}_\varepsilon \left[(\mathbf{y}_1, \ldots, \mathbf{y}_n) \right] = \langle \mathbf{\Lambda}, \mathbf{Y} \mathbf{D}(\varepsilon) \mathbf{Y}^{\mathsf{T}} \rangle_F, \tag{C.30}$$

where $\mathbf{Y} = \mathsf{iso}^{-1} \left[\mathbf{y}_1, \ldots, \mathbf{y}_n \right]$. Similar to the previous step, it can be shown with (C.13) that

$$\tilde{g}_\epsilon \left[\widetilde{\mathbf{Y}} \right] = \tilde{g} \left[\widetilde{\mathbf{Y}} \right] + \sum_{i=1}^{n} \varepsilon(n+1-i) \langle \mathbf{y}_i, \mathbf{\Lambda} \mathbf{y}_i \rangle. \tag{C.31}$$

For the second term,

$$|\langle \mathbf{y}_i, \mathbf{\Lambda} \mathbf{y}_i \rangle| \leqslant \max_j |\lambda_j| \langle \mathbf{y}_i, \mathbf{y}_i \rangle = \max_j |\lambda_j|. \tag{C.32}$$

Thus,

$$\left| \tilde{g}_\epsilon \left[\widetilde{\mathbf{Y}} \right] - \tilde{g} \left[\widetilde{\mathbf{Y}} \right] \right| \leqslant \varepsilon \max_j |\lambda_j| \sum_{i=1}^{n} (n+1-i),$$

$$= \frac{1}{2} n(n+1) \varepsilon \max_j |\lambda_j|. \tag{C.33}$$

(5.c) Using calculations similar to those in Step (4.c) we get from (C.33)

$$\tilde{g}\left[\widetilde{\mathbf{I}}_n\right] - \tilde{g}\left[\widetilde{\mathbf{Y}}\right] \geqslant \left(\tilde{g}_\epsilon\left[\widetilde{\mathbf{I}}_n\right] - \tilde{g}_\epsilon\left[\widetilde{\mathbf{Y}}\right]\right) - n(n+1)\varepsilon \max_j |\lambda_j|, \qquad (C.34)$$

where $\widetilde{\mathbf{I}}_n$ is an element of $\tilde{S}_n^\star$. Note that $\tilde{g}_\epsilon[\cdot]$ satisfies all the conditions imposed on $\tilde{g}[\cdot]$ in Step 4, hence $\tilde{g}_\epsilon[\cdot]$ achieves its global maximum at $\widetilde{\mathbf{I}}_n$. That is, $\tilde{g}_\epsilon\left[\widetilde{\mathbf{I}}_n\right] \geqslant \tilde{g}_\epsilon\left[\widetilde{\mathbf{Y}}\right]$ for all $\widetilde{\mathbf{Y}} \in \tilde{\mathcal{O}}_n$. Thus,

$$\tilde{g}\left[\widetilde{\mathbf{I}}_n\right] - \tilde{g}\left[\widetilde{\mathbf{Y}}\right] \geqslant -2\epsilon n\sqrt{n}\sum_{i=1}^n |\lambda_i|. \qquad (C.35)$$

Using arguments similar to those we used towards the end of Step (4.c) it can be concluded from (C.35) that even this more general $\tilde{g}[\cdot]$ achieves its maximum value on the set $\tilde{S}_n^\star$.

□

References

[1] P. R. Halmos, Positive approximants of operators, Indiana University Mathematics Journal 21 (10) (1972) 951–960.

[2] T. Hayden, J. Wells, Approximation by matrices positive semidefinite on a subspace, Linear Algebra and its Applications 109 (1988) 115–130.

[3] N. J. Higham, Computing a nearest symmetric positive semidefinite matrix, Linear Algebra and its Applications 103 (1988) 103–118.

[4] N. J. Higham, Computing the polar decomposition—with applications, SIAM Journal on Scientific and Statistical Computing 7 (4) (1986) 1160–1174.

[5] P. Cardou, G. Fournier, P. Gagnon, A nonlinear program for angular-velocity estimation from centripetal-acceleration measurements, IEEE/ASME Transactions on Mechatronics 16 (5) (2011) 932–944.

[6] K. J. Orlik-Rückemann, Aerodynamic aspects of aircraft dynamics at high angles of attack, Journal of Aircraft 20 (9) (1983) 737–752.

[7] A. Colagrossi, M. Lavagna, A spacecraft attitude determination and control algorithm for solar arrays pointing leveraging sun angle and angular rates measurements, Algorithms 15 (2) (2022).

[8] H. Ma, S. Xu, Magnetometer-only attitude and angular velocity filtering estimation for attitude changing spacecraft, Acta Astronautica 102 (2014) 89–102.

[9] J. Liang, O. Ma, Angular velocity tracking for satellite rendezvous and docking, Acta Astronautica 69 (11) (2011) 1019–1028.

[10] G. Guglieri, F. Maroglio, P. Pellegrino, L. Torre, Design and development of guidance navigation and control algorithms for spacecraft rendezvous and docking experimentation, Acta Astronautica 94 (1) (2014) 395–408.

[11] J. Lobo, L. Marques, J. Dias, U. Nunes, A. T. de Almeida, Sensors for mobile robot navigation, in: Autonomous Robotic Systems, Springer London, London, 1998, pp. 50–81.

[12] B. Barshan, H. Durrant-Whyte, Inertial navigation systems for mobile robots, IEEE Transactions on Robotics and Automation 11 (3) (1995) 328–342.

[13] E. G. Takhounts, M. J. Craig, K. Moorhouse, J. McFadden, V. Hasija, Development of brain injury criteria (BrIC), Stapp Car Crash Journal 57 (2013) 243.

[14] Y. Wan, W. Fang, R. W. Carlsen, H. Kesari, A finite rotation, small strain 2D elastic head model, with applications in mild traumatic brain injury, Journal of the Mechanics and Physics of Solids 179 (2023) 105362.

[15] R. W. Carlsen, A. L. Fawzi, Y. Wan, H. Kesari, C. Franck, A quantitative relationship between rotational head kinematics and brain tissue strain from a 2-D parametric finite element analysis, Brain Multiphysics 2 (2021) 100024.

[16] Y. Wan, A. L. Fawzi, H. Kesari, Determining rigid body motion from accelerometer data through the square-root of a negative semi-definite tensor, with applications in mild traumatic brain injury, Computer Methods in Applied Mechanics and Engineering 390 (2022) 114271.

[17] K. Parsa, Dynamics, state estimation, and control of manipulators with rigid and flexible subsystems, Ph.D. thesis, McGill University (2003).

[18] P. Cardou, J. Angeles, Angular velocity estimation from the angular acceleration matrix, Journal of Applied Mechanics 75 (2) (2008) 021003.

[19] K. Parsa, T. A. Lasky, B. Ravani, Design and mechatronic implementation of an accelerometer-based, kinematically redundant inertial measurement unit, in: Proceedings, 2005 IEEE/ASME International Conference on Advanced Intelligent Mechatronics., IEEE, 2005, pp. 644–651.

[20] D. P. Bertsekas, Nonlinear programming, Athena Scientific, 1999.

[21] S. Axler, Linear Algebra Done Right, Springer, 2015.

[22] Y. Wan, R. D. González-Cruz, D. Hoffman-Kim, H. Kesari, A mechanics theory for the exploration of a high-throughput, sterile 3D *in vitro* traumatic brain injury model, Biomechanics and modeling in mechanobiology 23 (4) (2024) 1179–1196.

[23] M. M. Rahaman, W. Fang, A. L. Fawzi, Y. Wan, H. Kesari, An accelerometer-only algorithm for determining the acceleration field of a rigid body, with application in studying the mechanics of mild traumatic brain injury, Journal of the Mechanics and Physics of Solids 143 (2020) 104014.

[24] STMicroelectronic, https://www.st.com, accessed on 13 February 2025 (1987).

[25] Analog Devices, https://www.analog.com, accessed on 13 February 2025 (1965).

[26] Tronic's Microsystems SA, https://www.tronics.tdk.com, accessed on 13 February 2025 (1997).

[27] Bosch Sensortec, https://www.bosch-sensortec.com, accessed on 13 February 2025 (2005).

[28] Silicon Sensing, https://siliconsensing.com, accessed on 13 February 2025 (1999).

[29] Gladiator Technologies, https://gladiatortechnologies.com, accessed on 13 February 2025 (2005).

[30] Safran, https://www.safran-group.com, accessed on 13 February 2025 (2005).

[31] R. F. Rinehart, Skew matrices as square roots, The American Mathematical Monthly 67 (2) (1960) 157–161.

[32] R. A. Horn, C. R. Johnson, Matrix Analysis, 2nd Edition, Cambridge University Press, 2013.

Figure 5: Body angular velocity in the punctuated rotation experiment. Subfigures (a)–(c), respectively, show graphs of the body angular velocity's three components. (These are equivalently the components of the angular velocity w.r.t. the body basis vectors, which are attached to the rigid body and rotate with it. The body basis vectors are illustrated, e.g., by the three arrows attached to the top corner of the gray cuboid (the rigid body) in the inset in (a).) Each component is graphed thrice: once from the ideal body angular velocity, specifically the $\overline{\mathbf{w}}[\cdot]$ given by (32) and (31c) for $\omega_m = 31.41$ and $\tau_1 = 5.81$, using a thick, solid, slightly transparent line; once from the body angular velocity yielded by the $\sqrt{\text{AO}}$ algorithm, namely $\widetilde{\mathbf{w}}[\cdot]$, using a thin, solid, opaque line; and finally from the body angular velocity yielded by the AO-algorithm, namely AO-$\widetilde{\mathbf{w}}[\cdot]$, using hollow markers. See §5.4 for details. Subfigure (d) shows the magnitude of the angular velocity. It too is similarly graphed thrice.

29

Figure B.6: Rotation test platform. Four triaxial accelerometers are mounted to a sensor holder. The sensor holder is rotating, driven by a feedback control BLDC motor, which is controlled by the computer.

a Apparatus of the experiment of constant and oscillatory rotation

(i)

Rotation direction

Accelerometer#4, A_4

A_2

Rotation axis, N

Sensor holder

A_3

A_1

Motor

(ii)

$e_3 [0]$

n

θ

$e_1 [0]$

$e_2 [0]$

${}^4e_3 [0]$

${}^4e_2 [0]$

${}^4e_1 [0]$

A_4

$\mathcal{E}$

e_3 ${}^4x [0]$

O

e_2

e_1

b Rotation angle and speed in the constant rotation experiment

(i)

θ (rad)

Eq. (2.1) for $\omega_m = 31.41$

τ (s)

(ii)

θ' (rad/s)

Derivative of the θ shown in (i)

τ (s)

c Rotation angle and speed in the oscillatory rotation experiment

(i)

θ (rad)

Eq. (2.2) for $\omega_m = 31.41$ and $\tau_1 = 11.62$

τ (s)

(ii)

θ' (rad/s)

Derivative of the θ shown in (i)

τ (s)

Figure B.7: Set-up and geometry parameters of the constant and oscillatory rotation experiment. (a.i) shows a sketch of the rigid body rotation experiment. (a.ii) shows the initial configuration of the rigid body in both constant and oscillatory experiment. The components of four accelerometer's initial position vectors are ${}^1x[0] = (0.064, 0.001, -0.016), {}^2x[0] = (-0.031, -0.055, -0.016)$, ${}^3x[0] = (-0.032, 0.055, -0.016)$, and ${}^4x[0] = (-0.015, 0.001, 0.103)$. The components of $\left({}^1e_i[0]\right)_{i\in(1,2,3)}$ are $((-1,0,0), (0,-1,0), (0,0,1))$; of $\left({}^2e_i[0]\right)_{i\in(1,2,3)}$ are $((0.50, 0.86, 0), (-0.86, 0.50, 0), (0,0,1))$; of $\left({}^3e_i[0]\right)_{i\in(1,2,3)}$ are $((0.50, -0.86, 0), (0.86, 0.50, 0), (0,0,1))$; of $\left({}^4e_i[0]\right)_{i\in(1,2,3)}$ are $((-1,0,0), (0,-1,0), (0,0,1))$; and of $(e_i[0])_{i\in(1,2,3)}$ are $((1,0,0), (0,1,0), (0,0,1))$. Subfigures (b.i) and (b.ii), respectively, shows the variation of the rotation angle and its derivative over time in the constant rotation experiment. Subfigures (c.i) and (c.ii), respectively, shows the variation of the rotation angle and its derivative over time in the oscillatory rotation experiment.

Figure B.8: Body angular velocity in the constant rotation experiment. Subfigures (a)–(c), respectively, show graphs of the body angular velocity's three components. (These are equivalently the components of the angular velocity w.r.t. the body basis vectors, which are attached to the rigid body and rotate with it. The body basis vectors are illustrated, e.g., by the three arrows attached to the top corner of the gray cuboid (the rigid body) in the inset in (a).) Each component is graphed thrice: once from the ideal body angular velocity, specifically the $\overline{\mathbf{w}}\left[\cdot\right]$ given by Eq. (32) in the main manuscript and (B.1) for $\omega_m = 31.41$ using a thick, solid, slightly transparent line; once from the body angular velocity yielded by the $\sqrt{\text{AO}}$ algorithm, namely $\widetilde{\overline{\mathbf{w}}}\left[\cdot\right]$, using a thin, solid, opaque line; and finally from the body angular velocity yielded by the AO-algorithm, namely AO-$\widetilde{\overline{\mathbf{w}}}\left[\cdot\right]$, using hollow markers. Subfigure (d) shows the magnitude of the angular velocity. It too is similarly graphed thrice.

Figure B.9: Body angular velocity in the oscillatory rotation experiment. Subfigures (a)–(c), respectively, show graphs of the body angular velocity's three components. (These are equivalently the components of the angular velocity w.r.t. the body basis vectors, which are attached to the rigid body and rotate with it. The body basis vectors are illustrated, e.g., by the three arrows attached to the top corner of the gray cuboid (the rigid body) in the inset in (a).) Each component is graphed thrice: once from the ideal body angular velocity, specifically the $\overline{\mathbf{w}}[\cdot]$ given by Eq. (32) in the main manuscript and (B.2) for $\omega_m = 31.41$ and $\tau_1 = 11.62$ using a thick, solid, slightly transparent line; once from the body angular velocity yielded by the $\sqrt{AO}$ algorithm, namely $\widetilde{\mathbf{w}}[\cdot]$, using a thin, solid, opaque line; and finally from the body angular velocity yielded by the AO-algorithm, namely AO-$\widetilde{\mathbf{w}}[\cdot]$, using hollow markers. Subfigure (d) shows the magnitude of the angular velocity. It too is similarly graphed thrice.