# Smooth Parameterization of Rigid-Body Inertia

Caleb Rucker *Member, IEEE*, Patrick M. Wensing *Member, IEEE*

***Abstract*—In this letter, we propose a parameterization of the rigid-body inertia tensor that is singularity free, guarantees full physical consistency, and has a straightforward physical interpretation. Based on a version of log-Cholesky decomposition of the pseudo-inertia matrix, we construct a smooth isomorphic mapping from $\mathbb{R}^{10}$ to the set of fully physically consistent inertia tensors. This facilitates inertial estimation via unconstrained optimization on a vector space and avoids the non-uniqueness and singularities which we show are inherent to parameterizations of the inertia tensor based on eigenvalue decomposition and principal moments. The elements of our parameterization have straightforward physical meanings in terms of geometric transformations applied to a reference body. While adopting this new parameterization breaks the linear least squares structure of the system identification, theoretical results guarantee that all local optima of the resulting problems remain global optima. We compare the performance of three different parameterizations of inertia by performing inertial estimation on test data from a set of 1000 simulations of randomly sampled rigid bodies subject to external time-varying wrench and measurement noise. We also investigate the performance of the log-Cholesky parameterization on a dataset from MIT Cheetah 3 to empirically demonstrate the theoretical results. Overall, the results indicate that the log-Cholesky parameterization achieves fast convergence while providing a simple and intuitive parametric description of inertia.**

***Index Terms*—Dynamics, Inertial Estimation**

## I. Introduction

Estimation of the inertial parameters of rigid bodies was an early topic in robotics research [1]. For a single rigid body in 3D, the inertial parameters are the six unique elements of the symmetric rotational inertia tensor, the mass, and the three first mass moments along each coordinate axis. It is well known that the dynamics model of any manipulator or other multi-body system is linear in these parameters, which conveniently leads to a linear least squares estimation problem that serves as the basis for identification algorithms, either offline or adaptive.

However, as many researchers have recognized, not all combinations of parameter values actually correspond to some possible physical body. This gives rise to the notion of *physical consistency* of a set of parameters. Physical consistency was first defined by the notion that the rotational inertia tensor should be positive definite [2]–[5]. However, as shown by Traversaro, et al. [6] this is a necessary but not sufficient condition for physical consistency. In addition, a set of triangle inequalities on the principal moments of inertia must be satisfied, which was termed *full physical consistency* in [6]. It was later shown that the conditions for full physical consistency were equivalent to the requirement that the pseudo-inertia matrix be positive definite [7], such that the estimation problem is a constrained optimization problem subject to linear matrix inequalities (LMIs) [7]. This approach was applied to flying robots in [8]. Using the fact that physically-consistent parameters are isomorphic to $\mathcal{P}^4$ (the set of $4 \times 4$ positive definite matrices), Lee and Park illuminated the Riemannian geometry of consistent parameters, and proposed non-convex optimization methods for system identification based upon it [9]. Coordinate-invariant non-Euclidean regularization strategies were later cast as a convex optimization problem in [10].

In some applications, it would be desirable to formulate inertial estimation as an unconstrained optimization problem on a set of parameters that inherently embeds the requirements of physical consistency. For example, recent work in differentiable physics simulators motivated exploration of several parameterizations of rigid-body inertia that embedded physical consistency constraints [11]. Similarly, [12] learned an object's unknown inertial properties online in order to facilitate robotic manipulation, and needed to enforce physical consistency due to noisy observations. However, the parameterizations so far explored in these works have not been isomorphic mappings between $\mathbb{R}^{10}$ and the space of physically consistent parameters. A prominent mapping in both [11], [12], and others is based on Traversaro et al. [6], which uses the orientation of the principal axes, the second moments about them, the center of mass, and the mass, as base parameters instead of the elements of the inertia tensor. This intuitive approach is very appealing, but it requires the estimation problem to be performed on the $\mathrm{SO}(3)$ manifold (or to parameterize $\mathrm{SO}(3)$), and it further suffers from non-uniqueness and symmetry singularities, as we later show.

In this letter we review the requirements of physical consistency and highlight the fact that the space of physically consistent inertial parameters is actually isomorphic to $\mathbb{R}^{10}$. This is made evident first through a parameterization using the exponential of a symmetric matrix. We then propose a new parameterization of inertia based on a form of the log-Cholesky decomposition of the pseudo-inertia matrix. This parameterization is inherently physically consistent in the full sense and thus amenable to unconstrained optimization, yet it still resides on the vector space $\mathbb{R}^{10}$ and contains no singularities, constituting a diffeomorphic mapping. It further has an intuitive physical interpretation as it constitutes an affine transformation applied to a reference body. With this

Manuscript received: September 9, 2021; Revised: December 12, 2021; Accepted: January 10, 2021. This paper was recommended for publication by Editor Lucia Pallottino upon evaluation of the Associate Editor and Reviewers' comments. This work was supported by the National Science Foundation under CAREER Award IIS-1652588. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Caleb Rucker is with the Dept. of Mechanical, Aerospace, and Biomedical Engr., University of Tennessee, Knoxville, TN caleb.rucker@utk.edu

Patrick M. Wensing is with the Dept. of Aerospace & Mechanical Engr., University of Notre Dame, IN pwensing@nd.edu

Digital Object Identifier (DOI): see top of this page.

new parameterization, we demonstrate inertial estimation on simulated noisy data sets, and compare the results to other methods. Tests with experimental data from MIT Cheetah 3 confirm that while this new parameterization leads to a non-convex optimization problem, the diffeomorphism ensures that all local optima of the problem remain globally optimal.

## II. Rigid-Body Inertia and Consistency

The inertial parameters of a general rigid-body object are defined as functionals of the body's mass-density field $\rho : \mathbb{R}^3 \mapsto \mathbb{R}_{\geq 0}$ as

$$\begin{aligned} m &= \int_{\mathbb{R}^3} \rho(\mathbf{x}) \mathrm{d}V \\ \mathbf{h} &= \int_{\mathbb{R}^3} \mathbf{x}\rho(\mathbf{x}) \mathrm{d}V = [h_x \; h_y \; h_z]^\top \\ \bar{\mathbf{I}} &= \iiint_{\mathbb{R}^3} \begin{bmatrix} y^2+z^2 & -xy & -xz \\ -xy & x^2+z^2 & -yz \\ -xz & -yz & x^2+y^2 \end{bmatrix} \rho(\mathbf{x}) \mathrm{d}V \\ &= \begin{bmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{xy} & I_{yy} & I_{yz} \\ I_{xz} & I_{yz} & I_{zz} \end{bmatrix} \end{aligned} \tag{1}$$

where $m$ is the total mass, $\mathbf{h} = m\mathbf{c}$ is the first mass moment with $\mathbf{c} \in \mathbb{R}^3$ the center of mass, $\bar{\mathbf{I}}$ is the rotational inertia about the origin of the body-fixed reference frame, $\mathbf{x} = [x \; y \; z]^\top \in \mathbb{R}^3$ are Cartesian coordinates in the body-fixed reference frame, and $\mathrm{d}V = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$ is the volume differential.

The 6D spatial inertia tensor is then

$$\mathbf{I} = \begin{bmatrix} \bar{\mathbf{I}} & \mathbf{S}(\mathbf{h}) \\ \mathbf{S}(\mathbf{h})^\top & m\mathbf{1}_3 \end{bmatrix} \tag{2}$$

where $\mathbf{S} : \mathbb{R}^3 \mapsto \mathfrak{so}(3)$ such that $\mathbf{S}(\mathbf{x})\mathbf{y} = \mathbf{x} \times \mathbf{y}$ [13].

### A. Physical Consistency

**Definition II.1.** A vector

$$\boldsymbol{\pi} = [m \; h_x \; h_y \; h_z \; I_{xx} \; I_{yy} \; I_{zz} \; I_{xy} \; I_{yz} \; I_{xz}]^\top \in \mathbb{R}^{10} \tag{3}$$

composed of candidate inertial parameters is *fully physically consistent* if and only if there exists some finite-valued density field $\rho(\cdot) : \mathbb{R}^3 \mapsto \mathbb{R}_{\geq 0}$ such that $\boldsymbol{\pi}$ satisfies (1) and $m > 0$. Let us define the set of physically consistent parameter vectors:

$$\mathcal{C} = \{\boldsymbol{\pi} \in \mathbb{R}^{10} \mid \boldsymbol{\pi} \text{ is fully physically consistent}\}$$

Note that since $\rho$ is required to be finite valued above, this definition of physical consistency excludes point masses, line masses, and plane masses, all of which would require an infinite density in order to have nonzero mass. This is equivalent to requiring physical bodies to have nonzero thickness in all dimensions.

While defining physical consistency is straightforward, characterizing the set $\mathcal{C}$ is more challenging. A helpful starting place is provided by the theorem below, as stated in [6], which we provide here for the reader's convenience.

**Theorem II.1.** *$\boldsymbol{\pi}$ is fully physically consistent if and only if $\bar{\mathbf{I}} \succ 0$, $m > 0$, and the following triangle inequalities are satisfied*

$$D_x < D_y + D_z, \quad D_y < D_x + D_z, \quad D_z < D_x + D_y \tag{4}$$

*where $D_x$, $D_y$, and $D_z$ are principal moments of inertia such that $\bar{\mathbf{I}} = \mathbf{R}\mathbf{D}\mathbf{R}^\top$ with $\mathbf{D} = diag(D_x, D_y, D_z)$ and $\mathbf{R} \in \mathrm{SO}(3)$.*

*Proof.* See [6, Lemmas 1 and 2, and Theorem 1] □

The triangle inequalities above exist due to the relationship of $D_x$, $D_y$ and $D_z$ to the second moments $L_x$, $L_y$ and $L_z$ along the principal axes.

$$D_x = L_y + L_z, \quad D_y = L_x + L_z, \quad D_z = L_x + L_y \tag{5}$$

where

$$\begin{aligned} L_x &= \int_{\mathbb{R}^3} \tilde{x}^2 \rho(\mathbf{x}) \mathrm{d}V, \\ L_y &= \int_{\mathbb{R}^3} \tilde{y}^2 \rho(\mathbf{x}) \mathrm{d}V, \\ L_z &= \int_{\mathbb{R}^3} \tilde{z}^2 \rho(\mathbf{x}) \mathrm{d}V \end{aligned}$$

and $\tilde{\mathbf{x}} = [\tilde{x} \; \tilde{y} \; \tilde{z}]^\top$ are coordinates expressed in $\mathbf{R}$, i.e., $\mathbf{R}\tilde{\mathbf{x}} = \mathbf{x}$. Positivity of $L_x$, $L_y$, and $L_z$ then implies the triangle inequalities.

## III. Parameterization via the Eigenvalue Decomposition

As detailed in [6], the theorem above immediately suggests a natural way to parameterize $\mathcal{C}$ in terms of $\mathbb{R}_{>0} \times \mathbb{R}^3 \times \mathrm{SO}(3) \times \mathbb{R}^3_{>0}$, i.e., use as base parameters

$$m \in \mathbb{R}_{>0}, \; \mathbf{c} \in \mathbb{R}^3, \; \mathbf{R} \in \mathrm{SO}(3), \; \mathbf{L} = [L_x \; L_y \; L_z]^\top \in \mathbb{R}^3_{>0}$$

and then calculate $\mathbf{D}$ from $\mathbf{L}$ using (5), and the rotational inertia as $\bar{\mathbf{I}} = \mathbf{R}\mathbf{D}\mathbf{R}^\top$. Note that the domain of the mapping is not a vector space due to the involvement of $\mathrm{SO}(3)$, $\mathbb{R}_{>0}$, and $\mathbb{R}^3_{>0}$. To perform physically-consistent inertial parameter estimation, we would need to enforce the restrictions imposed by these structures. $\mathrm{SO}(3)$ could be handled by (1) further parameterizing $\mathbf{R}$ in terms of Euler angles, quaternions, exponential coordinates, or any other conventional coordinate expression for rotation, or (2) enforcing the structure of $\mathrm{SO}(3)$ as constraints during optimization, or (3) employing optimization on manifolds as proposed in [6].

### A. Symmetry-Induced Singularities

However, regardless of how the inequalities and $\mathrm{SO}(3)$ are handled, this parameterization of $\mathcal{C}$ contains singularities that were not revealed in prior research. To demonstrate this, suppose $\mathbf{R}$ is parameterized using exponential coordinates $\boldsymbol{\beta} \in \mathbb{R}^3$ as $\mathbf{R} = e^{\mathbf{S}(\boldsymbol{\beta})}$. Then, establishing a vector of base parameters

$$\boldsymbol{\phi} = [m \in \mathbb{R}_{>0}, \quad \mathbf{c} \in \mathbb{R}^3, \quad \boldsymbol{\beta} \in \mathbb{R}^3, \quad \mathbf{L} \in \mathbb{R}^3_{>0}]$$

we can construct a physically consistent $\boldsymbol{\pi}$ as outlined above. After calculating the Jacobian

$$\boldsymbol{\Gamma} = \frac{\partial \boldsymbol{\pi}}{\partial \boldsymbol{\phi}}$$

upon examination, we find that $L_x - L_y$, $L_y - L_z$, and $L_x - L_z$, are all factors of the determinant of $\boldsymbol{\Gamma}$, indicating that the mapping from $\boldsymbol{\phi}$ to $\boldsymbol{\pi}$ is singular whenever any two of the second moments are equal.

Physically interpreted, the singularity above occurs whenever the parameters describe a body with rotational symmetry of order three or greater about any axis (symmetry of order $n$ means that rotations by $360°/n$ about some axis do not change the object). Common examples of this include any linear extrusion of a regular polygon (e.g., square, equilateral triangle, regular hexagon, etc.) and any solid of revolution about an axis (e.g., a vase or baseball bat). As a concrete example, consider a cylinder with axis along $z$. In this case, $L_x = L_y$. The singularity can be understood intuitively, since rotating the body about the $z-$axis produces no change in the inertia matrix. Whenever these singular combinations of base parameters occur, there exist certain infinitesimal changes in $\boldsymbol{\pi}$ that cannot be realized by any corresponding infinitesimal change in $\boldsymbol{\phi}$ because of the singular Jacobian.

It is important to note that these symmetry singularities are independent from any other singularities that might arise from a parameterization of $\mathbf{R}$. Indeed, since the relevant factors of the determinant are purely functions of the second moment parameters $L_x$, $L_y$, and $L_z$, the symmetry singularities are present even if $\mathbf{R}$ is handled in a singularity-free way with quaternions [14] or optimization on manifolds [6]. The effect of the symmetry-induced singularities could potentially be significant in differentiable physics applications and online adaptive estimation. In practice, many robotic links will exhibit such symmetry at least approximately, and it would be desirable to avoid parameterizations that introduce symmetry singularities near the solution.

In addition to the symmetry singularities, there is a uniqueness issue: for any rigid-body inertia matrix, there are at least six different parameters sets $\boldsymbol{\phi}$ which produce it, since $L_1$, $L_2$, and $L_3$ can be assigned in any order and subsequently combined with a different rotation $\mathbf{R}$ to arrive at the same $\bar{\mathbf{I}}$. Thus, while physically intuitive, any parameterization based on the eigenvalue decomposition is not one-to-one, requires a strategy for handling $\mathrm{SO}(3)$, and contains additional singularities.

## IV. Vector-Space Parameterizations based on the Pseudo Inertia

In light of the previous section, one might assume that $\mathcal{C}$ is not isomorphic to any vector space (because $\mathrm{SO}(3)$ is involved), and that any attempt at parameterization in terms of $\mathbb{R}^{10}$ will suffer from the same problems. But this turns out not to be true. In this section, we explore two parameterizations that provide a bijective mapping from $\mathbb{R}^{10}$ to $\mathcal{C}$ without singularities. Both mappings are based on the 4x4 pseudo-inertia matrix, defined as [7]:

$$\mathbf{J} = \begin{bmatrix} \boldsymbol{\Sigma} & \mathbf{h} \\ \mathbf{h}^\top & m \end{bmatrix}$$

where

$$\boldsymbol{\Sigma} = \frac{1}{2}\mathrm{Tr}(\bar{\mathbf{I}})\mathbf{1}_3 - \bar{\mathbf{I}} \quad \text{and} \quad \bar{\mathbf{I}} = \mathrm{Tr}(\boldsymbol{\Sigma})\mathbf{1}_3 - \boldsymbol{\Sigma}$$

Defining $\mathbf{q}$ as the homogeneous coordinate representation of the point $\mathbf{x}$, i.e.,

$$\mathbf{q} = \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix}$$

The pseudo-inertia can be simply written as

$$\mathbf{J} = \int_V \mathbf{q}\mathbf{q}^\top \rho(\mathbf{x}) \, \mathrm{d}V. \tag{6}$$

It was proved in [7] that $\boldsymbol{\pi}$ is fully physically consistent if and only if $\mathbf{J}(\boldsymbol{\pi}) \succ 0$. This result has been useful in expressing physical consistency in terms of linear matrix inequalities [7], and for recognizing that $\mathcal{C}$ is equivalent to the Riemannian manifold $\mathcal{P}^4$ ($4 \times 4$ symmetric positive-definite matrices). This observation, in turn, provides a natural coordinate-free metric on $\mathcal{C}$ and is useful in improving regularization during inertial estimation [9].

### A. Exponential Map Parameterization

It has also been observed [9], [15] that a bijection can be constructed from $\mathbb{R}^{10}$ to $\mathcal{P}^4$ via the exponential of a symmetric matrix $\mathbf{Q} \in \mathcal{S}^4$:

$$\exp(\mathbf{Q}) = \sum_{k=0}^{\infty} \frac{\mathbf{Q}^k}{k!} = \mathbf{K} \in \mathcal{P}^4$$

and the matrix logarithm of $\mathbf{K} \in \mathcal{P}^4$ given by

$$\log(\mathbf{K}) = -\sum_{k=1}^{\infty} \frac{(\mathbf{1} - \mathbf{K})^k}{k} = \mathbf{Q} \in \mathcal{S}^4$$

When combined with the results above, this bijection establishes the isomorphism $\mathbb{R}^{10} \simeq \mathcal{C}$, so it is possible to use the exponential map to define a vector-space parameterization of $\mathcal{C}$. However, the physical meaning of this parameterization is not immediately clear, and its computation is not in a closed form, so it is difficult to gain intuition and insight or to form the Jacobian.

### B. Log-Cholesky Parameterization

Building on this prior work, we here further exploit the pseudo-inertia to identify a smooth isomorphic parameterization from $\mathbb{R}^{10}$ to $\mathcal{C}$, which is both physically intuitive and computationally closed form.

The real symmetric matrix $\mathbf{J}$ is positive definite if and only if there exists a real nonsingular matrix $\mathbf{M}$ such that

$$\mathbf{J} = \mathbf{M}\mathbf{M}^\top$$

The Cholesky decomposition is the unique factorization of this form where $\mathbf{M}$ is lower triangular with a positive diagonal. Similarly, there is also a unique decomposition

$$\mathbf{J} = \mathbf{U}\mathbf{U}^\top \tag{7}$$

where $\mathbf{U}$ is upper triangular with positive diagonal. While we could employ either upper or lower Cholesky decompositions,

the upper decomposition corresponds to a natural physical interpretation, as we will discuss below.

An upper triangular matrix $\mathbf{U}$ is nonsingular if and only if the diagonal elements are nonzero. Let us then construct a nonsingular upper triangular matrix $\mathbf{U}$ in the form

$$\mathbf{U} = e^{\alpha} \begin{bmatrix} e^{d_1} & s_{12} & s_{13} & t_1 \\ 0 & e^{d_2} & s_{23} & t_2 \\ 0 & 0 & e^{d_3} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{8}$$

Viewed in light of the decomposition (7) this can be seen as a version of the log-Cholesky decomposition with the bottom-right scalar term factored out.[1] Thus, given a vector of base parameters

$$\boldsymbol{\theta} = [\alpha, d_1, d_2, d_3, s_{12}, s_{23}, s_{13}, t_1, t_2, t_3]^\top \in \mathbb{R}^{10}$$

we can construct a physically consistent inertia matrix from the relationships above, resulting in the mapping $\boldsymbol{\pi} : \mathbb{R}^{10} \mapsto \mathcal{C}$, shown below:

$$\boldsymbol{\pi} = \mathbf{f}(\boldsymbol{\theta}) = e^{2\alpha} \begin{bmatrix} t_1^2 + t_2^2 + t_3^2 + 1 \\ t_1 e^{d_1} \\ t_1 s_{12} + t_2 e^{d_2} \\ t_1 s_{13} + t_2 s_{23} + t_3 e^{d_3} \\ s_{12}^2 + s_{13}^2 + s_{23}^2 + e^{2d_2} + e^{2d_3} \\ s_{13}^2 + s_{23}^2 + e^{2d_1} + e^{2d_3} \\ s_{12}^2 + e^{2d_1} + e^{2d_2} \\ -s_{12} e^{d_1} \\ -s_{12} s_{13} - s_{23} e^{d_2} \\ -s_{13} e^{d_1} \end{bmatrix} \tag{9}$$

The mapping is *onto*, and the uniqueness of the log-Cholesky decomposition implies that it is also one-to-one and invertible. Moreover, we can calculate the Jacobian

$$\boldsymbol{\Gamma} = \frac{\partial \boldsymbol{\pi}}{\partial \boldsymbol{\theta}},$$

and upon inspection we find that the only non-constant factors of the determinant of $\boldsymbol{\Gamma}$ are $e^{\alpha}$, $e^{d_1}$, $e^{d_2}$, and $e^{d_3}$, none of which can be zero $\forall \boldsymbol{\theta} \in \mathbb{R}^{10}$. The mapping $\mathbf{f}(\boldsymbol{\theta})$ and its inverse are thus smooth, constituting a diffeomorphism.

## V. Properties of the Log-Cholesky Parameterization

### A. Preservation of the Unique Optimum

The Log-Cholesky parameterization of inertia above converts the task of estimation from a linear least squares problem to a nonlinear optimization problem that is non-convex. This appears to be the price of embedding physical consistency in the parameter set. However, the resulting nonlinear optimization problem is still well posed, and all local optima are globally optimal because the mapping is diffeomorphic. To prove this, consider a linear least squares cost function $F(\boldsymbol{\pi})$. Suppose there is a minimum of the cost function at a physically consistent vector of parameters $\boldsymbol{\pi}^*$. This implies that

$$\frac{\partial F}{\partial \boldsymbol{\pi}}\Big|_{\boldsymbol{\pi}^*} = \mathbf{0}$$

and since the cost function is linear least squares, it is convex with all critical points in the domain being global optima. The optimum $\boldsymbol{\pi}^*$ corresponds to a unique set of parameters $\boldsymbol{\theta}^* = \mathbf{f}^{-1}(\boldsymbol{\pi}^*)$. In terms of the new parameterization, the gradient of the cost function is also zero at $\boldsymbol{\theta}^*$:

$$\frac{\partial F}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}^*} = \left(\frac{\partial F}{\partial \boldsymbol{\pi}} \boldsymbol{\Gamma}\right)\Big|_{\boldsymbol{\pi}^*} = \mathbf{0}$$

and since $\boldsymbol{\Gamma}$ is full rank everywhere, it follows that while the parameterization converts the linear least squares problem to a more general nonlinear one, this does not introduce any additional local minima[2].

### B. Physical Interpretation and Generalization

In order to physically interpret the log-Cholesky parameterization and gain intuition that may inform the estimation problem, recall that the pseudo inertia can be written in terms of the homogeneous coordinates of points in the body as in (6). Now, any affine transformation of a reference body's geometry can be written in the form

$$\mathbf{q}' = \mathbf{A}\mathbf{q} = \begin{bmatrix} \mathbf{B} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \mathbf{q}$$

where $\mathbf{B} \in \mathbb{R}^{3\times 3}$ is a linear transformation (encompassing, e.g., various compositions of rotation, shearing, and scaling) and $\mathbf{t} \in \mathbb{R}^3$ uniformly translates the object in space. While this transformation affects the geometry of the reference body without influencing its density, one could additionally transform the density of the body by scaling by some positive factor $\beta^2$. The combination of geometric transformation and density scaling results in the the pseudo-inertia of the transformed body:

$$\begin{aligned} \mathbf{J} &= \int_V \mathbf{q}'\mathbf{q}'^T \beta^2 \rho(\mathbf{x}) \mathrm{d}V \\ &= \int_V \mathbf{A}\mathbf{q}\mathbf{q}^T\mathbf{A}^\top \ \beta^2 \rho(\mathbf{x}) \mathrm{d}V \\ &= \mathbf{C}\mathbf{J}_0\mathbf{C}^\top \end{aligned} \tag{10}$$

where $\mathbf{C} = \beta\mathbf{A}$. We now recognize that our generalized log-Cholesky matrix $\mathbf{U}$ has the same basic form as $\mathbf{C}$ (where $\beta = e^{\alpha}$). Thus, our parameterization can be interpreted as applying an affine transformation along with a density scaling to some reference body with $\mathbf{J}_0 = \mathbf{1_{4\times 4}}$. Thus, we can generalize our parameterization (7) by applying it to an arbitrary reference object with pseudo-inertia $\mathbf{J}_0$,

$$\mathbf{J} = \mathbf{U}\mathbf{J}_0\mathbf{U}^\top \tag{11}$$

This change does not affect the uniqueness and invertibility of (7), as long as the chosen reference body itself is not degenerate (i.e., not a point mass, line mass, or plane mass). To see this, let $\mathbf{J} = \mathbf{U}_1\mathbf{U}_1^\top$ and $\mathbf{J}_0 = \mathbf{U}_2\mathbf{U}_2^\top$ be the unique Cholesky decompositions of $\mathbf{J}$ and $\mathbf{J}_0$. Then $\mathbf{U} = \mathbf{U}_1\mathbf{U}_2^{-1}$ exists as the unique upper triangular matrix with positive

[1]The log-Cholesky decomposition was used in [16] to parameterize variance-covariance matrices, which are closely related to pseudo-inertia matrices [7].

[2]This result can alternatively be understood as follows. We note that a convex function is not guaranteed to be convex under a nonlinear change of variables. However, under a diffeomorphic change of variables, any convex function is guaranteed to be goedesically convex [17], [18]. Similar to regular convexity, this property ensures that all critical points are globally optimal.

diagonal satisfying (11), due to the fact that upper triangular matrices with positive diagonal form a group.

Based on the interpretation of our parameterization as an affine transformation and density scaling of a reference body, we can now identify the following physical meanings of the elements of our parameter vector $\boldsymbol{\theta}$:

$$d_1,\ d_2,\ d_3$$

are dimensionless and stretch (scale) the reference body along $x, y, z$ by their respective exponentials. A positive $d_i$ indicates stretch, and negative values indicate compression. The transformation of the reference object will never cause it to degenerate to zero thickness in any dimension because the exponential can never be zero.

$$s_{12}, s_{13}, s_{23}$$

are dimensionless and shear the reference body in the $xy, xz, yz$ planes without changing the object's volume.

$$t_1, t_2, t_3$$

have length units and translate the reference body uniformly in $x, y, z$. The quantity $\alpha$ is dimensionless and effectively scales the density $\rho$ by the factor $e^{2\alpha}$. Again, the transformation can never cause the body's mass to degenerate to zero. In addition to being physically intuitive and eliminating all symmetry-induced singularities, the log-Cholesky parameterization of interia is computationally attractive since the Jacobian is algebraic and easy to calculate due to the closed-form mapping.

The generalization (11) provides a physically meaningful and intuitive way to incorporate prior information within inertial parameter estimation. An approximate inertial model (e.g., from a CAD model and/or physical measurements) can be used to determine the nominal pseudo-inertia $\mathbf{J}_0$ for a given robot link. Parametrizing the intertia as a transformation applied to the nominal geometry is attractive because the parameter magnitudes themselves indicate how far away from nominal the estimate is.

## VI. Comparison of Vector-Space Parameterizations

In this section we compare the three parameterizations of rigid-body inertia from Sections III, IV-A, and IV-B by testing their performance on an example inertial estimation problem.

### A. Integration of Randomly Sampled Rigid-Bodies

Our example problem assumes an unconstrained rigid body subject to known forces and moments over time. The Newton and Euler equations can be written compactly as

$$\mathbf{I}\dot{\boldsymbol{\xi}} + \boldsymbol{\xi} \times^* \mathbf{I}\boldsymbol{\xi} = \mathbf{w} \tag{12}$$

where $\boldsymbol{\xi} = [\boldsymbol{\omega}^\top \quad \mathbf{v}^\top]^\top$ is the 6-DOF body-frame velocity vector, $\mathbf{w} = [\mathbf{m}^\top \quad \mathbf{f}^\top]^\top$ is the known body-frame wrench consisting of applied force $\mathbf{f}$ and moment $\mathbf{m}$ vectors, and

$$\boldsymbol{\xi}\times^* = \begin{bmatrix} \mathbf{S}(\boldsymbol{\omega}) & \mathbf{S}(\mathbf{v}) \\ \mathbf{0}_{3\times3} & \mathbf{S}(\boldsymbol{\omega}) \end{bmatrix}$$

| Parameter | Sampling Interval |
|---|---|
| m | 0.01 kg to 1.0 kg |
| $c_x$ | -0.5 m to 0.5 m |
| $c_y$ | -0.5 m to 0.5 m |
| $c_z$ | -0.5 m to 0.5 m |
| $L_x$ | 0.01 kg m$^2$ to 1.0 kg m$^2$ |
| $L_y$ | 0.01 kg m$^2$ to 1.0 kg m$^2$ |
| $L_z$ | 0.01 kg m$^2$ to 1.0 kg m$^2$ |
| $\boldsymbol{\beta} \in \mathbb{R}^3$ where $R = e^{\mathbf{S}(\boldsymbol{\beta})}$ | Sphere of radius $2\pi$ |

Table I
Sampling Intervals for Ground-Truth Inertial Parameters used in Simulations

We performed simulations with 1000 randomly generated physically consistent inertia matrices by sampling uniform probability density functions for the mass, center of mass, second mass moments, and exponential rotation coordinates, using the eigenvalue construction. The ranges of these uniform distributions are outlined in Table I. The components of the external applied wrench were as follows:

$$\mathbf{w} = \begin{bmatrix} \sin(t) \\ \sin(t+\pi/3) \\ \sin(t+2\pi/3) \\ \sin(t+3\pi/3) \\ \sin(t+4\pi/3) \\ \sin(t+5\pi/3) \end{bmatrix}$$

and the simulation time interval was 0 to $2\pi$. By construction, this ensures that the time series of wrenches spans $\mathbb{R}^6$. Each simulation was calculated using Matlab's `ode23t` with data taken at equally spaced 0.1 second intervals. Simulated measurement noise and bias were then added to the resulting solution at the velocity level,

$$\tilde{\boldsymbol{\xi}}_i = \boldsymbol{\xi}_i + \boldsymbol{\epsilon_i}$$

where the error $\boldsymbol{\epsilon}_i \in \mathbb{R}^6$ was sampled from a uniform distribution of width $0.05$ (noise) and mean of $0.1$ (bias) in the heterogeneous units of $\boldsymbol{\xi}$, forming datasets on which inertial estimation can be performed by minimizing the sum of squared wrench residuals.

At each pair of discrete time points ($t_i$, and $t_{i+1}$) in each rigid body dataset, a residual wrench vector $\mathbf{r}_i$ was computed by a second-order discretization of (12) using the implicit midpoint method

$$\begin{aligned} \bar{\dot{\boldsymbol{\xi}}}_i &= \frac{\tilde{\boldsymbol{\xi}}_{i+1} - \tilde{\boldsymbol{\xi}}_i}{t_{i+1} - t_i} \\ \bar{\boldsymbol{\xi}}_i &= \frac{\tilde{\boldsymbol{\xi}}_{i+1} + \tilde{\boldsymbol{\xi}}_i}{2} \\ \mathbf{r}_i &= \mathbf{I}\bar{\dot{\boldsymbol{\xi}}}_i + \bar{\boldsymbol{\xi}}_i \times^* \mathbf{I}\bar{\boldsymbol{\xi}}_i - \mathbf{w}_i \end{aligned} \tag{13}$$

and all $\mathbf{r}_i$ vectors were concatenated to form the least squares residual vector.

### B. Comparison of Inertial Estimation Results

We performed inertial estimation on these 1000 datasets by solving the least squares problem using four different parameterizations of the inertia matrix: (1) the ten inertial

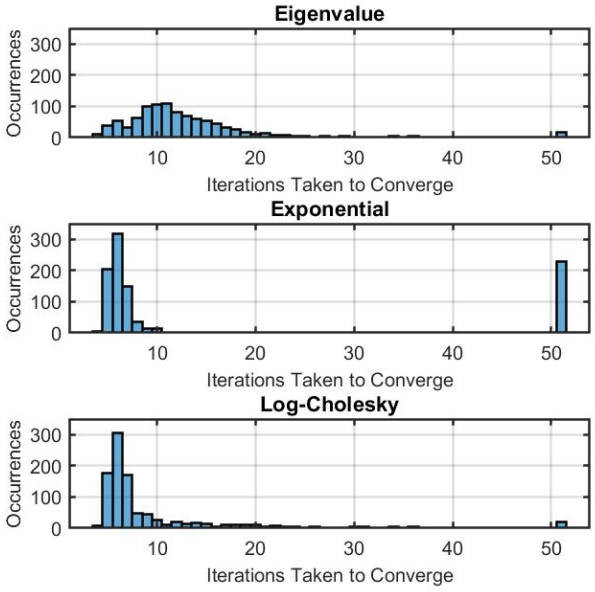


Figure 1. Results from 1000 estimation problems are shown for the different parameterizations discussed in this paper. The plots show a histogram for the number of iterations the algorithm took to satisfy the convergence criteria. The algorithms were terminated after a maximum of 50 iterations. The average time taken per iteration was 0.02 seconds, equal for all methods.

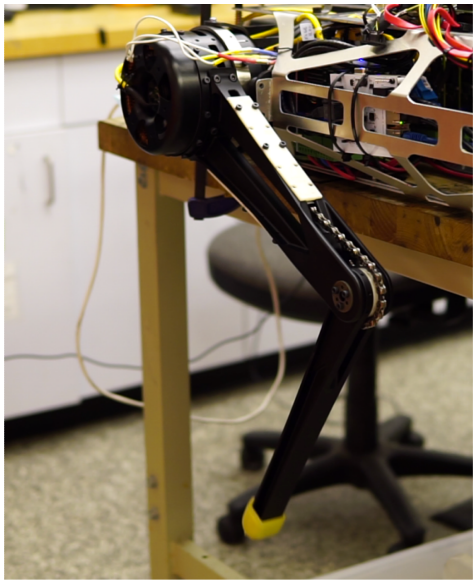

Figure 2. Experimental Setup for MIT Cheetah 3 System Identification dataset.

parameters themselves, (2) the eigenvalue construction in Section III (with exponential coordinates for rotation), (3) the exponential construction in IV-A, and (4) the log-Cholesky construction in section IV-B.

When the inertial parameters themselves (as given in (3)) are used as the variables, estimation is a linear least squares problem with a unique closed-form solution. However, out of the 1000 simulated rigid bodies, this resulting least squares solution produced a non-physical set of inertial parameters 15.1% of the time (151 occurrences). Of these, none resulted in negative mass. 42 nonphysical cases resulted in $\bar{\mathbf{I}}$ not positive definite, and the mean of the most negative eigenvalue from each of these cases was $-0.04$ kg m$^2$. All of the nonphysical cases violated at least one of the triangle inequalities (4), and the mean of the triangle violations was $-0.09$ kg m$^2$. In these cases, measurement noise, bias, and discretization error led to a nonphysical estimation result, as also observed in [6], [9].

The other three parameterizations create a nonlinear least squares estimation problem, but they ensure that the result is physically consistent. In each nonlinear case, we solved the nonlinear estimation problem using Matlab's `lsqnonlin()` function with the trust-region-reflective algorithm, no parameter bounds, and a maximum of 50 iterations. All problems were initialized with parameters such that the initial inertia matrix was the identity matrix.

Differences in computational performance can be seen in the number of iterations taken for the nonlinear estimation to converge. These data are encapsulated in Figure 1, which catalogs the occurances of the number of iterations taken for the algorithm to terminate. Number of iterations also serves as a proxy for time, because the average time taken per iteration (0.02 seconds) was equal for all methods. The eigenvalue parameterization displays a wider range of times to convergence with few failures. The exponential method converged quickly but had far more failures, and the log-Cholesky converged quickly with few failures, in addition to being closed-form and physically intuitive.

## VII. Application to MIT Cheetah Estimation

A final set of results assessed the performance of the new paramterization as applied to a dataset from the MIT Cheetah 3 [7], [19]. The robot was attached to the table as shown in Fig. 2, and one leg was moved through a maximally exciting trajectory. Joint angles and their derivatives were estimated based on a non-causal Savitzky–Golay filter [20] (fourth order, with a window size of 100 ms) applied to encoder data sampled at 1 kHz. Joint torques were estimated based on current readings reported by motor amplifiers, also sampled at 1 kHz. The joint torques were left unfiltered since they are the dependent variable within regression, and least squares methods are suited to handle noise in the dependent variable. Further details on the setup can be found in [7]. A regularized nonlinear least squares problem was then formulated as

$$\min_{\boldsymbol{\theta},\boldsymbol{\theta}_b,\boldsymbol{\theta}_c} \sum_{j=1}^{N_s} \| \mathbf{Y}_j \boldsymbol{\pi}(\boldsymbol{\theta}) + \mathrm{diag}(\dot{\boldsymbol{q}}_j)\mathbf{b}(\boldsymbol{\theta}_b) + \mathrm{diag}(\mathrm{sign}(\dot{\boldsymbol{q}}_j))\mathbf{b}_c(\boldsymbol{\theta}_c) - \boldsymbol{\tau}_j \|^2 + \gamma\, d(\boldsymbol{\pi}(\boldsymbol{\theta}) || \hat{\boldsymbol{\pi}}) \tag{14}$$

where $\mathbf{Y}_j$ gives the classical regressor at the $j$-th sample point, $\dot{\boldsymbol{q}}_j$ the joint rates at the $j$-th sample point, and $\mathbf{b} \in \mathbb{R}^3$ and $\mathbf{b}_c \in \mathbb{R}^3$ are the viscous and Coulomb friction coefficients, respectively. The final term $\gamma\, d(\boldsymbol{\pi}(\boldsymbol{\theta})||\hat{\boldsymbol{\pi}})$ is a coordinate-invariant regularizer (see [10]) that penalizes the deviation from a prior estimate $\hat{\boldsymbol{\pi}}$, with $\gamma = 10^{-2}$ the regularization weight chosen via cross-validation. Note that since not all

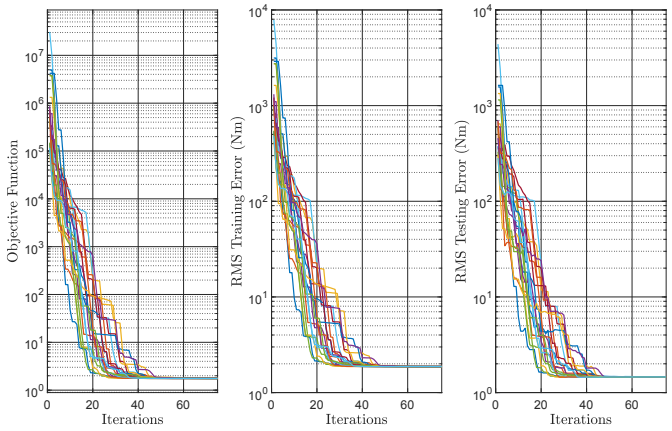


Figure 3. Example convergence from 20 random runs of system identification for the MIT Cheetah 3 dataset. The results are in empirical agreement with the theoretical analysis of Section V-A: while the new parameterization makes the system identification problem nonconvex, all initial conditions converge to a global minimizer.

parameters are identifiable [21], without this regularization, a subspace of the inertial parameters can be freely chosen, with their evolution during optimization determined due to numerical effects. To ensure that the friction terms were physically plausible, additional parameters $\boldsymbol{\theta}_b$ and $\boldsymbol{\theta}_c$ were optimized, with relation to the friction parameters as:

$$\mathbf{b}(\boldsymbol{\theta}_b) = \begin{bmatrix} e^{\boldsymbol{\theta}_b,1} \\ e^{\boldsymbol{\theta}_b,2} \\ e^{\boldsymbol{\theta}_b,3} \end{bmatrix} \qquad \mathbf{b}_c(\boldsymbol{\theta}_c) = \begin{bmatrix} e^{\boldsymbol{\theta}_c,1} \\ e^{\boldsymbol{\theta}_c,2} \\ e^{\boldsymbol{\theta}_c,3} \end{bmatrix}$$

The inertial parameters for one leg included 10 parameters for each link (ab/ad, thigh, and shank) and 10 parameters for each motor rotor (ab/ad, hip, knee) such that $\boldsymbol{\theta} \in \mathbb{R}^{60}$. The problem was solved using `fminunc()` in MATLAB with the trust-region algorithm. RMS residuals of 1.47 Nm, 1.49 Nm, and 1.43 Nm were observed on the ab/ad, hip, and knee torque predictions on a batch of testing data that was distinct from the training data. The prior and output parameters are summarized in the supplemental material.

A nearly identical overall RMS prediction error was observed across a set of trials wherein the initial guess to the optimization was randomized. Figure 3 shows the convergence over iterations for a representative set of 20 random initial guesses. For these tests, all parameters $(\boldsymbol{\theta}, \boldsymbol{\theta}_b, \boldsymbol{\theta}_c)$ were randomized on $[-1, 1]$, resulting in initial residuals up to $10^5$ Nm RMS. The objective shown on the left includes both the regularization and squared loss from (14). The RMS error on the training data is shown in the middle vs. iterations for the optimizer. The same parameter estimates at each iteration were then used to assess validation error on the separate testing data set, as shown on the right. The monotonic decrease in the testing error across training iterations suggests that overfitting was not present. Despite the significant initial suboptimality, all solutions converged to the the same RMS error. These tests empirically verify the theoretical property that all local minimizers of the non-convex (but unconstrained) problem (14) are global optimizers.

## VIII. Conclusions

We have proposed a method for parameterizing rigid-body inertia based on the Log-Cholesky decomposition of the symmetric positive definite pseudo-inertia matrix. We showed that this parameterization has the intuitive physical meaning of an affine transformation and density scaling applied to a reference body. Furthermore, it does not suffer from singularities that are inherent parameterizations based on an eigenvalue decomposition. Over a large set of simulated estimation problems with measurement noise, the method converged faster than an eigenvalue parameterization, and with less failures than an exponential parameterization. By construction, the method only produces inertia estimations that are physically consistent, in contrast to estimation based on the inertial parameters alone, which produced many physically inconsistent results. These results were also shown to scale well to practical identification problems such as those encountered when identifying legged robots. The results with the Cheetah 3 dataset further support the theoretical claim of the work in that unconstrained optimization can, in fact, lead to globally optimal, physically consistent solutions with our new parameterization.

Overall, the work paves the way for easily accessible unconstrained optimization tools to be employed for enforcing the physical plausibility of learned dynamics models. There are many possible extensions to be considered. One immediate extension is the application to closed-chain mechanisms. This generalization could be accomplished by combining the new parameterization with the identification methods in [22] to address kinematic loop-closure constraints. The method also would present applicability within physically consistent adaptive control schemes (e.g., [23]). In this case, continuous time update laws would directly inherit consistency guarantees when converted to discrete time if using our proposed parameterization.

## References

[1] C. G. Atkeson, C. H. An, and J. M. Hollerbach, "Estimation of Inertial Parameters of Manipulator Loads and Links," *The International Journal of Robotics Research*, vol. 5, no. 3, pp. 101–119, 1986.

[2] K. Yoshida, K. Osuka, H. Mayeda, and T. Ono, "When is the set of base parameter values physically impossible?" in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'94)*, vol. 1, 1994, pp. 335–342 vol.1.

[3] K. Yoshida and W. Khalil, "Verification of the positive definiteness of the inertial matrix of manipulators using base inertial parameters," *The International Journal of Robotics Research*, vol. 19, no. 5, pp. 498–510, 2000.

[4] K. Ayusawa, G. Venture, and Y. Nakamura, "Real-time implementation of physically consistent identification of human body segments," *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 6282–6287, 2011.

[5] M. Gautier, S. Briot, and G. Venture, "Identification of consistent standard dynamic parameters of industrial robots," in *2013 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, 2013, pp. 1429–1435.

[6] S. Traversaro, S. Brossette, A. Escande, and F. Nori, "Identification of fully physical consistent inertial parameters using optimization on manifolds," *IEEE International Conference on Intelligent Robots and Systems*, vol. 2016-Novem, pp. 5446–5451, 2016.

[7] P. M. Wensing, S. Kim, and J.-J. E. Slotine, "Linear Matrix Inequalities for Physically Consistent Inertial Parameter Identification: A Statistical Perspective on the Mass Distribution," *IEEE Robotics and Automation Letters*, vol. 3, no. 1, pp. 60–67, 2018.

[8] M. Ekal and R. Ventura, "A dual quaternion-based discrete variational approach for accurate and online inertial parameter estimation in free-flying obots," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 6021–6027.
[9] T. Lee and F. C. Park, "A Geometric Algorithm for Robust Multibody Inertial Parameter Identification," *IEEE Robotics and Automation Letters*, pp. 1–1, 2018.
[10] T. Lee, P. M. Wensing, and F. C. Park, "Geometric robot dynamic identification: A convex programming approach," *IEEE Transactions on Robotics*, vol. 36, no. 2, pp. 348–365, 2020.
[11] G. Sutanto, A. Wang, Y. Lin, M. Mukadam, G. Sukhatme, A. Rai, and F. Meier, "Encoding physical constraints in differentiable newton-euler algorithm," in *Proceedings of the 2nd Conference on Learning for Dynamics and Control*, ser. Proceedings of Machine Learning Research, A. M. Bayen, A. Jadbabaie, G. Pappas, P. A. Parrilo, B. Recht, C. Tomlin, and M. Zeilinger, Eds., vol. 120. PMLR, 10–11 Jun 2020, pp. 804–813.
[12] B. Sundaralingam and T. Hermans, "In-hand object-dynamics inference using tactile fingertips," *IEEE Transactions on Robotics*, vol. 37, no. 4, pp. 1115–1126, 2021.
[13] R. Featherstone, *Rigid Body Dynamics Algorithms*. Boston, MA: Springer US, 2008.
[14] C. Rucker, "Integrating rotations using nonunit quaternions," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 2979–2986, Oct 2018.
[15] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache, "Geometric means in a novel vector space structure on symmetric positive-definite matrices," *SIAM journal on matrix analysis and applications*, vol. 29, no. 1, pp. 328–347, 2007.
[16] J. C. Pinheiro and D. M. Bates, "Unconstrained parametrizations for variance-covariance matrices," *Statistics and Computing*, vol. 6, no. 3, pp. 289–296, 1996.
[17] T. Rapcsák, "Geodesic convexity in nonlinear optimization," *Journal of Optimization Theory and Applications*, vol. 69, no. 1, pp. 169–183, 1991.
[18] P. M. Wensing and J.-J. Slotine, "Beyond convexity—contraction and global convergence of gradient descent," *Plos one*, vol. 15, no. 8, p. e0236661, 2020.
[19] G. Bledt, M. J. Powell, B. Katz, J. Di Carlo, P. M. Wensing, and S. Kim, "MIT Cheetah 3: Design and control of a robust, dynamic quadruped robot," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 2245–2252.
[20] A. Savitzky and M. J. Golay, "Smoothing and differentiation of data by simplified least squares procedures." *Analytical chemistry*, vol. 36, no. 8, pp. 1627–1639, 1964.
[21] P. M. Wensing, G. Niemeyer, and J.-J. E. Slotine, "Observability in inertial parameter identification," *arXiv preprint arXiv:1711.03896*, 2017.
[22] M. Mistry, S. Schaal, and K. Yamane, "Inertial parameter estimation of floating base humanoid systems using partial force sensing," in *2009 9th IEEE-RAS International Conference on Humanoid Robots*. IEEE, 2009, pp. 492–497.
[23] T. Lee, J. Kwon, and F. C. Park, "A natural adaptive control law for robot manipulators," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 1–9.