



저작자표시-비영리 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.
- 이차적 저작물을 작성할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

공학박사학위논문

**Geometric Methods for Dynamic Model-Based
Identification and Control of Multibody Systems**

기하학적 방법론을 이용한
다물체 시스템 모델 추정 및 제어

2019 년 8 월

서울대학교 대학원

기계항공공학부

이 태 운

Geometric Methods for Dynamic Model-Based Identification and Control of Multibody Systems

TAEYOON LEE

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in the
DEPARTMENT OF MECHANICAL AND AEROSPACE ENGINEERING

at
SEOUL NATIONAL UNIVERSITY

July 2019

ABSTRACT

Geometric Methods for Dynamic Model-Based Identification and Control of Multibody Systems

by

Taeyoon Lee

Department of Mechanical and Aerospace Engineering
Seoul National University

Multibody mechanical systems constitute a large and important class of input-output systems that are the subject of model-based planning and control. Model-based control and planning methods in robotics have recently gained more attention, as complex, high-dimensional robots, e.g., humanoids and quadruped robots, are beginning to perform highly dynamic tasks. While the performance of these methods is in general affected by the accuracy of the model, common situations for complex high-dimensional systems raise a number of difficulties in estimating the model parameters in a robust and generalizable manner.

In this thesis, we demonstrate that many of the challenges can be mitigated by appealing to coordinate-invariant, differential geometric methods. The key lies in the finding that mass-inertial parameters reside in a curved space of positive definite matrices endowed with a natural Riemannian metric which captures the distance between the parameter pairs in a physically meaningful, coordinate-invariant way. Taking this geometric perspective as our point of departure, we present geometric, coordinate-invariant algorithms that allow robust estimation of the parameters in various stages and situations of estimation.

We first propose geometric formulations and algorithms for robust offline parameter identification of multibody mechanical systems. In particular, we provide a convex programming approach to the geometric dynamic parameter identification through the use of second-order approximations of the Riemannian distance. Not only does this allow for the use of fast convex optimization algorithms that are guaranteed to converge to a global solution, but also ensures coordinate-invariance while allowing for the inclusion of additional convex constraints as imposed by physical considerations and other practical requirements. Our geometric identification methods are validated through extensive experiments on a wide range of systems ranging from a robot manipulator to a legged robot and a human subject. The results show markedly improved robustness and generalizability *vis-à-vis* existing vector space methods.

Then we address the problem of generating optimal excitation trajectories for parameter identification. We suggest a new set of optimality criteria that encodes the information from the trajectory samples in a coordinate-invariant way. The resulting optimal excitation trajectories are coordinate invariant and can be obtained efficiently and robustly using recursive analytic gradients of the criteria. The proposed geometric framework is also used to devise a coordinate-invariant algorithm for characterizing the effectively identifiable set of parameters given a set of excitation trajectory samples. The suggested

method is particularly useful for robust identification of high-dimensional systems like humanoid robot that can execute only a limited range of feasible trajectories. The improved robustness and accuracy of our geometric approach in comparison to existing methods is demonstrated through numerical experiments involving industrial manipulators and a humanoid robot.

Finally, we propose a geometric parameter adaptation law for adaptive control of robot manipulators. Toward deriving our geometric adaptation law, we extend the way of defining Riemannian manifold structure on the space of feasible inertial parameters to more general types of mechanical parameters including, e.g., joint frictions and stiffness. Then we show that a coordinate-invariant choice of a Lyapunov function that can be naturally defined on the so-called Hessian manifold of mechanical parameters leads to an adaptation law that can be viewed as a natural gradient descent flow on the corresponding manifold. Perhaps most importantly, our geometric approach considerably reduces the degree to which engineering choices must be made in the adaptation gain matrix compared to the existing methods. Our geometric adaptive control framework is further extended to robust adaptive control where arbitrary convex constraints imposed on the parameters can be taken into account with geometric projection methods. The efficacy of our method is verified with adaptive trajectory tracking control task involving a seven-dof robot manipulator through both simulation and real experiment.

Keywords: System identification, Adaptive control, Optimal experimental design, Optimal trajectory generation, Convex optimization, Matrix manifolds, Riemannian geometry.

Student Number: 2015-20749

Contents

Abstract	iii
List of Tables	xiii
List of Figures	xv
1 Introduction	1
1.1 A Geometric Approach	4
1.2 Organization	7
2 Preliminaries	11
2.1 Introduction	11
2.2 Symmetric Positive Definite Matrix Manifold	12
2.2.1 Affine-invariant Riemannian Metric	13
2.2.2 Log-det Bregman Divergence	14
2.3 Matrix Lie Groups	16
2.3.1 The Rotation Group	17
2.3.2 The Euclidean Group	18

3	Geometric Dynamic Identification of Multibody Systems	21
3.1	Introduction	21
3.2	Preliminaries	24
3.2.1	Physically Consistent Rigid Body Inertial Parameters	24
3.2.2	Linear Least Squares based Identification	27
3.3	Geometry of Rigid Body Inertial Parameters	31
3.3.1	Riemannian Distance Metric	32
3.3.2	Entropic Divergence	37
3.3.3	Constant Pullback Metric	38
3.3.4	Distribution Awareness of Geometric Distances	39
3.4	Geometric Identification with Geodesic Least Squares	43
3.4.1	Intrinsic Riemannian Error Criterion	43
3.4.2	Cyclic Optimization Algorithm	46
3.5	Geometric Identification with Convex Relaxations	48
3.5.1	Provable Comparative Analysis Scheme	50
3.6	Experiments	56
3.6.1	AMBIDEX Robot Manipulator	56
3.6.2	MIT Cheetah 3 Robot	63
3.6.3	Human with Low-Cost Affordable Sensors	71
3.7	Discussion	75
3.8	Conclusion	77
4	Geometric Criteria for Excitation Trajectory Optimization	79
4.1	Introduction	79
4.2	Preliminaries	81
4.2.1	Optimal Design of Experiments	81

4.2.2	Excitation Criteria for Multibody Systems	84
4.2.3	Coordinate Invariance and Normalization	86
4.3	Geometric Excitation Criteria	89
4.3.1	Motivation	90
4.3.2	Pushforward Metric on Observable Parameters	90
4.3.3	Coordinate-invariant Criterion	94
4.4	Optimal Excitation Trajectory Generation	97
4.5	Determination of Effectively Identifiable Parameter Set	99
4.5.1	Reduced Identification	100
4.5.2	Reduced Optimal Excitation	102
4.6	Simulation Study	102
4.6.1	SCARA with Unknown Payload	103
4.6.2	KUKA iiwa R280 Manipulator	107
4.6.3	Atlas V5 Humanoid Robot	111
4.7	Conclusion	116
5	Geometric Robust Adaptive Control of Robot Manipulators	117
5.1	Introduction	117
5.2	Preliminaries	120
5.2.1	Adaptive Control of Robot Manipulators	121
5.3	Barrier-Hessian Manifolds	125
5.3.1	Rigid Body Inertial Parameters	127
5.3.2	Joint Friction/Stiffness Parameters	130
5.4	Geometric Parameter Update Laws	132
5.4.1	Geometric Adaptation Law	134
5.4.2	Geometric Projection Law	138

5.5	Simulation Study: Barret WAM7 Manipulator	140
5.5.1	Full Adaptation	142
5.5.2	Unknown Payload Adaptation	143
5.6	Experiment: AMBIDEX Robot Manipulator	147
5.7	Conclusion	150
6	Conclusion	153
6.1	Summary	153
6.2	Future Work	156
6.3	Concluding Remark	158
A	Proofs and Supplemental Derivations	159
A.1	Supplemental Derivations	159
A.2	Proof of Proposition 3.2	161
A.3	Supplemental Propositions for Section 3.5.1	164
A.4	Proof of Proposition 4.1	166
A.5	Proof of Proposition 5.1	168
A.6	Proof of Proposition 5.2	169
A.7	Proof of Proposition 5.5	170
A.8	Proof of Proposition 5.6	170
A.9	Proof of Proposition 5.7	171
A.10	Proof of Proposition 5.8	172
A.11	Proof of Proposition 5.9	172
B	Algorithms and Implementation Details	173
B.1	Mappings Associated with Inertial Parameters	173
B.2	Inertial Parameter Perturbation Strategy	175

B.3 Recursive Regressor Gradient for Multibody Systems	175
B.3.1 Joint Torque Regressor	178
B.3.2 Ground Reaction Wrench Regressor	181
B.4 Matrix Inverse-free Computation of Geometric Adaptation Laws	181
Bibliography	183
Abstract	196

List of Tables

3.1	Identification Results on a Human Subject : RMS Error of Predicted f_z [N], x_{ZMP} [cm], y_{ZMP} [cm]	74
4.1	2-dof SCARA robot with unknown payload: identification results with “condition number” optimization	106
4.2	KUKA iiwa R280: identification results with “E-optimality” optimization	110
4.3	Identification results on Atlas V5 humanoid with parameter reduction: Output validation	113
4.4	Identification results on Atlas V5 humanoid with parameter reduction: Validation on joint space dynamics	115

List of Figures

- 1.1 Three unit mass ellipsoids, each of uniform density and drawn to the same scale; 1kg ellipsoidal bodies A, B and C with principal radii of $[0.1, 0.1, 0.1]$ m, $[0.15, 0.06, 0.08]$ m and $[0.02, 0.25, 0.02]$ m respectively. . 6
- 3.1 The effect of imposing LMI while minimizing a least-squares objective function. The grey ellipsoidal contour indicates the least-squares objective function while the blue-shaded region is the feasible convex region of the search space in \mathbb{R}^{10n} defined by the LMI. 3.1(a) When the pure least-squares minimizer Φ^{LS} is physically consistent, it coincides with the least squares minimizer subject to the LMI, Φ^{LS+LMI} . 3.1(b) When Φ^{LS} is not physically consistent, Φ^{LS+LMI} lies on the boundary of \mathcal{M}^n . 30

3.2	Minimal geodesic path between two mass density functions (green ellipses) under the Riemannian metric 3.3.19. Each mass density function is visualized by an ellipse whose center and shape matches the center of mass and covariance matrix, respectively. 3.2(a) shows a plot of the distances from the fixed mass density function (black ellipse) to ones with the same mass and covariance, but constant deviation in positions of center of mass. 3.2(b) shows the geodesic path	35
3.3	Comparisons of squared distances between inertial parameters of ellipsoidal bodies, each of uniform density and drawn to the same scale. In each plot, the distances are normalized with the value at a particular configuration of the two bodies (marked on the plots with black vertical dashed lines). Two unit systems are used to evaluate the distances: kg-m (SI unit) and inch-pound. The inertial parameters of the target bodies (grey ellipsoids) are used to evaluate the pullback form of the Riemannian metric for the constant pullback metric. The distance from the fixed target density (grey ellipsoid) to the perturbed ones (red ellipsoids) are compared while varying the position and orientation, but with total mass fixed.	41

3.4	<p>Figures 3.4(a) compares the distance from the fixed target density (grey ellipsoid) to the perturbed ones (red ellipsoids) while varying the total volume, but with total mass fixed. Figure 3.4(b) compares the distance from the target body (grey sphere) with varied total mass m to the identical body perturbed by adding uniformly distributed point mass noise inside the body whose total mass Δ_m is also drawn from the uniform distribution on $[-0.3 \cdot m, 0.3 \cdot m]$. For each case of the target body with total mass m, the mean of the distances to 500 samples of the perturbed bodies is evaluated; the deviation between the “normalized” Euclidean (SI, inch-pound) distances is very small so that their plots in Figure 3.4(b) almost overlap each other. The “normalized” geometric (const. pullback, entropic and Riemannian) distances likewise have minimal deviation from one another.</p>	42
3.5	<p>Pictorial description of the geometric meaning of the physically consistent inertial parameter identification problem. The green cone-shaped shaded region indicates the manifold $\mathcal{M}^n \simeq \mathcal{P}(4)^n$ embedded in \mathbb{R}^{10n}. The blue hyperplanes each represent the linear constraints in $A\Phi = b$. Each of the black arrows indicates the projection of Φ to the hyperplanes whose paths depend on the metric. Red dashed circles denote the projections of Φ to physically inconsistent values outside the green shaded region.</p>	44

3.6	<p>Conceptual sketch of solving a constrained point(${}^0\Phi$)-to-set(\mathcal{X}_c) projection problem for least-squares identification of inertial parameters with different distance measures for regularization. The grey cone the represents the region of physically consistent inertial parameters, while the green set represents the boundary of the set \mathcal{X}_c with bounded least-squares error. Φ_{Euc}^* and Φ_{Geom}^* are the projected values of the prior value ${}^0\Phi$ to the set \mathcal{X}_c with Euclidean distance and geometric distance, e.g., entropic divergence or Riemannian distance, respectively. The ellipsoids depicted inside the grey boxes represent the visualization of the equivalent mass densities of the inertial parameters along the paths of projections.</p>	52
3.7	<p>Inertial parameter identification results for the AMBIDEX robot manipulator with different regularizers and varying the number of training samples. Friction parameters are fixed to the values identified from the full sample set. The red dashed horizontal line indicates the RMS test error with the prior value ${}^0\bar{\Phi}$ and the black bold horizontal line indicates the level of RMS test error using full set of 4,100 training samples for identification.</p>	60
3.8	<p>Visualization of the identified inertial parameters of the AMBIDEX manipulator with 30 training samples using different distance measures for regularization. The ellipsoids represent uniform mass densities uniquely realizable from the inertial parameters ϕ_i of each link. The ellipsoids for point mass parameters $\bar{\mu}$ are obtained by the transformation rule $\bar{\Phi} = R\bar{\mu}$. 61</p>	61
3.9	<p>Comparison of the ground ZMP prediction trajectories with the identified inertial parameters and prior value.</p>	63

3.10	Identification results on the MIT Cheetah 3 leg with different regularizers and varying the number of training samples. Friction parameters are fixed to the values identified from the full samples. The red dashed horizontal line indicates the RMS test error with the prior value ${}^0\Phi$ while the black bold horizontal line indicates the level of RMS test error using the full 7,379 training samples for identification.	66
3.11	Visualization of the identified inertial parameters of the MIT Cheetah 3 leg with 30 training samples using different distance measures for regularization. The ellipsoids represent uniform mass densities uniquely realizable from the inertial parameters ϕ_i of each link. The ellipsoids for the point mass parameters μ are obtained by the transformation rule $\Phi = R\mu$.	68
3.12	Identification results on the MIT Cheetah 3 leg with Euclidean, constant pullback, and entropic distance based regularizers validated on floating-base galloping motion trajectory. According to the rule (B.2.11), the noise condition $\delta=0.5$, $N_m=10$ is used for perturbing the ground truth inertial parameters from the fixed prior value ${}^0\Phi$. The mean and variance of RMS prediction errors on the joint torque, body force, and body moment are presented.	69
3.13	Human inertial parameter identification results while varying the regularization factor $\gamma = \alpha \cdot \text{tr}(A^T A)$. The ellipsoids represent uniform mass densities uniquely realizable from the inertial parameters ϕ_i of each link. The ellipsoids for point mass parameters μ are obtained by the transformation rule $\Phi = R\mu$	73
4.1	1-link, 1-dof robot	86

4.2	Contour maps of A-optimality excitation criterion for the one-dof robot, shown at logarithmic scale and evaluated with different coordinate choices	88
4.3	Geometric distortion view of the proposed excitation criterion. The metrics characterized with (green and red) bold ellipses are predefined respectively on the full parameter space and observation space. The (green and red) dashed ellipses represent respectively the pushforward and pull-back of the metrics to the observable(base) parameter space under the linear mappings B and A_B	91
4.4	The figures show 2-dof SCARA robot with unknown payload(red solid sphere) and reference frames considered, and joint specifications of 7-dof KUKA iiwa R280 industrial manipulator.	103
4.5	Atlas V5 joint specifications	111
4.6	Snapshots of 10-th (top) and 20-th (bottom) balancing motion trajectories used for (reduced) dynamic identification and validation	112
4.7	Plot of eigenvalues of the normalized covariance matrix $H_0^{1/2}CH_0^{1/2}$ for Atlas V5 balancing trajectories. The horizontal dashed line represents the threshold value $\epsilon_\lambda = 10^{-2}$ used for parameter selection.	113
5.1	Example of a 2-dimensional barrier-Hessian manifold (\mathcal{M}, h)	126
5.2	Geometric adaptation on (\mathcal{M}, h)	136
5.3	Geometric adaptation on (\mathcal{M}, h) with projection at the boundary of \mathcal{C} .	140
5.4	Plot of joint tracking RMS error when adapting the entire set of the inertial parameters with various levels of noise in the initial estimate . .	143

5.5	Plot of joint tracking RMS error when adapting the entire set of the inertial parameters on a varied sequence of desired trajectories, with a fixed noise level for the initial estimate; $\sigma = 0.4$. RMS errors for ‘No-adaptation’ all exceeded 7 deg	144
5.6	Joint tracking RMS error plots for adaptation of three kinds of unknown loads; 5.6(a) Sphere, 5.6(b) Ellipsoid1, and 5.6(c) Ellipsoid2. The plots above show mean RMS tracking error for repeated rounds of periodic desired trajectories, while the bottom plot shows a time plot of RMS tracking error. The color legends for all the plots are depicted in 5.6(c). Mean RMS tracking error for ‘No-adaptation’ in 5.6(c) all exceed 70 deg.	146
5.7	Snapshots of adaptation scenarios on the AMBIDEX robot manipulator. 500g (yellow colored) sandbags are placed on multiple body links.	148
5.8	Joint tracking RMS errors during adaptive control of AMBIDEX robot manipulator subject to link inertia and friction uncertainties	149
5.9	Joint tracking RMS error trajectory during adaptive control of AMBIDEX robot manipulator with 500g sandbag loaded at link 10.	150
B.1	Fixed-base tree-structured multibody system	176

1

Introduction

Accurate dynamic models are becoming increasingly necessary for robots performing highly dynamic tasks. Simplified dynamic models (e.g., the inverted pendulum, articulated point mass models, or centroidal dynamics models [1]) are easy to compute, and often quite effective for robots performing a narrow range of tasks like bipedal walking. However, for robots to perform a wider and more versatile set of tasks requiring fast and accurate motions, or physical interactions involving contact forces, complete multibody dynamic models are often essential [2, 3, 4, 5, 6]. Efficient robot dynamics algorithms and advances in current computing technology make real-time computation of the dynamics possible for even moderately complex, high-dof robots [7, 8, 9]. Accurate complete dynamic models are also essential for building a high-quality simulator; not only can simulators be used for test beds for hand-designed controllers, they can also be used to provide a virtual environment in which robots learn complex controllers on their own by repeated trial and error [10, 11, 12, 13]. There are many other applications in which dynamic modeling is important as well as for robot control. Particularly, human can also be modeled and represented as a rigid multibody system like robots,

and estimating subject-specific dynamic models is important for the dynamic analysis of human gait and sports motions [14].

Of course, it goes without saying that a dynamic model is only as good as the accuracy of the model parameters like mass-inertial parameters, friction parameters, etc. Not surprisingly, there is an extensive literature on estimating dynamic model parameters, commonly referred to as *system identification* or *dynamic calibration*, with most approaches exploiting the fact that the dynamic equations are linear in the parameters [15]; the equations of motion for mechanical systems are second-order differential equations of the form

$$\tau = M(q, p)\ddot{q} + b(q, \dot{q}, p) = Y(q, \dot{q}, \ddot{q})p, \quad (1.0.1)$$

where $(q, \dot{q}) \in \mathbb{R}^n \times \mathbb{R}^n$ is the state, $\tau \in \mathbb{R}^n$ is the control input, $M(q, p) \in \mathbb{R}^{n \times n}$ is the symmetric positive-definite mass matrix, $p \in \mathbb{R}^N$ are the model parameters, and $b(q, \dot{q}, p) \in \mathbb{R}^n$ captures potential, Coriolis, and other terms as governed by the laws of physics; these terms can be collected and represented as $Y(q, \dot{q}, \ddot{q})p$, where $Y(q, \dot{q}, \ddot{q}) \in \mathbb{R}^{n \times N}$. System identification typically proceeds by first taking measurements of τ , q , \dot{q} , \ddot{q} at multiple time instances t_1, \dots, t_m along some reference trajectory, and formulate the calibration problem as a least-squares estimation problem for p

One reason why this problem is less straightforward is that the p cannot assume any arbitrary vector value, but rather must satisfy a set of complex nonlinear constraints dictated by physics, e.g., masses cannot be negative, and inertia tensors must be symmetric positive-definite. Further complicating matter is that in practice the measurement conditions and the quality of reference trajectories significantly impact the accuracy of the identification. The measurements should be sufficiently complete and reliable, and the reference trajectory should be sufficiently rich so as to excite the full spectrum of the

dynamics, which can be high-dimensional and quite complex; the number of parameters N can range from 30-60 for a typical industrial manipulator to several hundred for a humanoid robot. Importantly, the set of admissible reference trajectories may be restricted by stability, safety, and other considerations (e.g., a humanoid robot should always maintain postural stability, and the actuators must be capable of generating needed torques). Moreover, for such high-dimensional systems, the available measurements may be limited by the number and location of the sensors (e.g., only a subset of a robot's joints may be equipped with torque sensors). Also, with the growing emphasis on making robots more affordable, measurements obtained from low-cost sensors typically will be incomplete and quite noisy.

The problem of estimating the dynamic parameters of a robot can become even more complicated, when the dynamic parameters change from the initial values during the task. A typical case involves many of the recent lightweight robot manipulators for which loading moderately heavy object at the end-effector can largely change the dynamic behaviors. Then the mass-inertial parameters need to be consistently re-identified in an online manner. Similarly, for mobile-base robots that can be used for cargo transportation, the ability to sense and immediately adapt to different payload conditions will be highly required for maintaining dynamic balance. This type of problem can be generally framed as *adaptive control* in which real-time closed-loop stability of the controlled system must be taken into account in the design of both the online parameter estimator and the model-based controller. While many robot adaptive control algorithms have been developed toward theoretically ensuring the stability of the system, there are still many practical problems and limitations to be solved.

Perhaps the most fundamental difficulty regarding parameter estimation for adaptive control lies in the fact that only limited trajectory data is available for online parameter estimation. This is essentially the same problem as discussed above on how the richness

and quality of reference trajectories can affect the accuracy of identification. Another critical limitation that existing robot adaptive control algorithms have in common is the requirement that users must choose a valid initial adaptation gain matrix. This can be a time-consuming process requiring repeated trial and error. Although in many cases stability of the system is theoretically guaranteed without regarding to the choice of adaptation gain matrix, in practice this choice significantly affects the performance, e.g., speed of convergence and transient behavior. Further, a naive choice of the gain matrix more often than not causes severe numerical instability issue mainly due to the common theoretic stability guarantees derived under somewhat ideal conditions, e.g., noise-free measurements, exact continuous-time implementation of the parameter adaptation law, etc.

For these and other reasons, estimating dynamic model parameters for multibody mechanical systems in both offline and online settings still remains a theoretical and practical challenge.

1.1 A Geometric Approach

In this thesis, we show that many of the aforementioned difficulties in dynamic parameter estimation of multibody mechanical systems can be mitigated by appealing to differential geometric methods. More specifically, our main contribution lies in reformulating the existing problem formulations by exploiting the fact that the collection of mass-inertial parameters constitutes a Riemannian manifold with a naturally defined Riemannian metric. This is in contrast to almost all of the existing methods adopting the standard Euclidean metric between the parameter values.

We first show through both examples and analysis that our Riemannian distance metric more accurately accounts for the underlying mass distributions of bodies in a

coordinate-invariant way. To illustrate, consider the three unit mass ellipsoids shown in Figure 1.1, each of uniform mass density and drawn to the same scale. Ideally we seek to compare the mass distributions of these three bodies. Since this space is infinite-dimensional, the mass distributions for each ellipsoid are instead reduced to a finite-dimensional set of ten mass-inertial parameters, with the parameter values depending on the choice of body-fixed reference frame and units. Evaluating these values in SI units and with respect to the body frames drawn in black, for the standard Euclidean distance metric $d_{Euc}(\cdot, \cdot)$ it can be shown that $d_{Euc}(A, B) = d_{Euc}(A, C)$, which mildly defies intuition. Worse, if the parameter values are now obtained with respect to the green reference frames, the same Euclidean distance metric results in the even more nonsensical $d_{Euc}(A, B) \gg d_{Euc}(A, C)$ by a factor of four; clearly a meaningful distance metric should not depend on the choice of body frame. Using a Riemannian distance $d_{Riem}(\cdot, \cdot)$ on the other hand leads to the eminently more reasonable $d_{Riem}(A, B) < d_{Riem}(A, C)$, with the calculated distances invariant to the choice of body frame.

Based on this Riemannian metric, we first formulate and minimize a coordinate-invariant error criterion for robust parameter identification. We show that accuracy and robustness of the identification can be vastly improved as compared to the Euclidean counterparts. Yet, a continued drawback of existing geometric methods is that the corresponding optimization problems are inherently non-convex, have numerous local minima, and are computationally highly intensive to solve. We address these twin deficiencies of computational complexity and non-convexity, and forge the Riemannian geometric framework into an efficient, robust, and practical tool for robot dynamic parameter identification. The key idea is a convex formulation of the regularization term based on a second-order approximation of the Riemannian distance. Not only does this allow for the use of fast convex optimization algorithms that are guaranteed to converge to

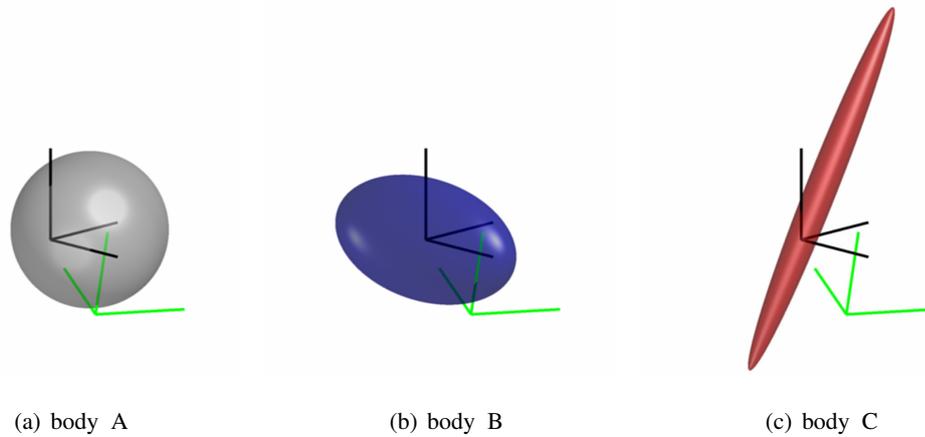


Figure 1.1: Three unit mass ellipsoids, each of uniform density and drawn to the same scale; 1kg ellipsoidal bodies A, B and C with principal radii of $[0.1, 0.1, 0.1]$ m, $[0.15, 0.06, 0.08]$ m and $[0.02, 0.25, 0.02]$ m respectively.

a global solution, but also ensures coordinate-invariance while allowing for the inclusion of additional convex constraints as imposed by physical considerations and other practical requirements.

Taking the same coordinate-invariant, geometric perspective, we also consider the problem of generating sufficiently rich trajectories for mechanical systems identification. More specifically, we address the problem of how to effectively generate reference trajectory samples that encode maximal information—the optimal excitation trajectory generation problem—taking into account physical constraints and measurement characteristics, thus leading to more accurate and robust parameter identification. Our main focus will be on minimizing a new coordinate-invariant set of excitation criteria admitting physically meaningful interpretations. Then we show how the same geometric framework can be applied to infer the optimal reduced set of effectively identifiable parameters from the given trajectory samples. Such reduction of the model parameters

can be useful not only for finding optimal excitation trajectories, but also parameter identification for complex high-dimensional systems like humanoid robots, in which the execution of reference trajectories that sufficiently excite the entire structurally identifiable set of parameters is difficult or impractical to achieve.

The final contribution of this thesis is a geometric adaptive control framework for robot manipulators. Toward deriving our geometric adaptation law, we extend the Riemannian manifold structure defined on the space of feasible inertial parameters to general mechanical parameters, e.g., including frictions and stiffness, defined on an arbitrary convex set. More specifically, Hessian manifold structure induced by a particular choice of convex barrier function is defined on the space of mechanical parameters; the proposed Riemannian manifold structure on the inertial parameters can be regarded as a special case. Then a Lyapunov-stable parameter update law can be naturally derived using the Bregman divergence associated with the corresponding convex barrier function. Whereas existing adaptation laws can be viewed as a gradient update law on a flat Euclidean space, our method can be viewed as a natural gradient-like update law on a curved space endowed with a Riemannian metric. Our geometric adaptation law is coordinate invariant and considerably reduces the degree to which engineering choices must be made in the constant adaptation gain. We also extend our method to incorporate robust adaptation strategies in which arbitrary convex constraints on the parameters are respected by non-smooth projection strategies.

1.2 Organization

In Chapter 2, we review the geometric properties of some important matrix manifolds, positive definite matrices and matrix Lie groups. Affine-invariant Riemannian geometry of a positive definite matrix manifold will serve as a basis for defining a Riemannian

manifold structure on the inertial parameter space. Notations and operations presented in characterizing two particular matrix Lie groups, Rotation group and Euclidean group, will be adopted in deriving recursive coordinate-invariant algorithms for optimal excitation trajectory generation of multibody systems.

Chapter 3 is devoted to proposing geometric formulations and algorithms for robust parameter identification of multibody mechanical systems. We first show how the choice of physical coordinate system explicitly enters into existing identification formulations. We then define a natural geometric structure on the inertial parameter space and show how to formulate the geometric identification problem in a coordinate-invariant, and also computationally efficient, way. Our geometric identification methods are validated through extensive experiments on high-dimensional systems ranging from a robot manipulator to a legged robot and a human subject both with restricted sensor measurements.

In Chapter 4 we propose a set of geometric optimal excitation criteria that can be optimized to generate high-quality reference trajectories. The resulting optimal trajectories are coordinate invariant, and can be obtained efficiently and robustly using recursive analytic gradients of the criteria. For high-dimensional systems that can execute only a limited range of feasible trajectories, we also show how our geometrical framework can be used to optimally identify a reduced set of parameters for the given set of trajectories. The improved robustness and accuracy of our geometric approach vis-à-vis existing methods is demonstrated through numerical experiments involving industrial manipulators and a high-dimensional humanoid robot.

Chapter 5 addresses the problem of adaptive control of robot manipulators in which the parameters are adapted online while also guaranteeing the stability of the plant interconnected with the model-based controller. We show that defining a coordinate-invariant choice of a Lyapunov function leads to a natural, stable adaptation law, which actually

turns out to have a form of natural gradient descent flow on the Hessian manifold. The method is further extended to robust adaptive control where arbitrary convex constraints imposed on the parameters can be taken into account with geometric projection methods. Our method is verified with adaptive trajectory tracking control task involving a seven-dof robot manipulator through both simulation and real experiment.

2

Preliminaries

2.1 Introduction

In this chapter, we summarize geometric properties of $n \times n$ symmetric positive definite matrix manifold, $\mathcal{P}(n)$, and two important instances of matrix Lie groups, special orthogonal group $SO(n)$ and special Euclidean group $SE(n)$.

Affine-invariant geometric structure of $\mathcal{P}(n)$ introduced in Section 2.2 will serve as a basis for defining a natural Riemannian manifold structure on the space of rigid body inertial parameters in Chapter 3. We also present a Log-det divergence measure on $\mathcal{P}(n)$ as a convex second-order approximation of the Riemannian distance. This divergence metric would be used in Chapter 3 for convex formulation of geometric dynamic identification, and in Chapter 5 as a valid Lyapunov function candidate for deriving geometric adaptation laws for robot adaptive control.

Section 2.3 briefly summarizes basic concepts of Lie groups and presents notations and operations associated with Rotation group and Euclidean group of rigid body motions, $SO(3)$ and $SE(3)$, and their corresponding Lie algebras, $so(3)$ and $se(3)$.

The presented notations and operations shall be mainly adopted for describing recursive coordinate-invariant algorithms involving rigid multibody dynamics and its derivatives in Chapter 4.

2.2 Symmetric Positive Definite Matrix Manifold

The space of real-symmetric $n \times n$ positive-definite matrices is defined as,

$$\mathcal{P}(n) = \{P \in \mathbb{R}^{n \times n} \mid P = P^T, P \succ 0\}. \quad (2.2.1)$$

$\mathcal{P}(n)$ is embedded in $\mathcal{S}(n)$, the space of real-symmetric $n \times n$ matrices, as an open and convex subset. Also, $\mathcal{P}(n)$ is a differentiable manifold whose tangent space at a point $P \in \mathcal{P}(n)$, denoted $T_P\mathcal{P}(n)$, can be identified with symmetric matrices $\mathcal{S}(n)$. A basis for $\mathcal{S}(n)$ can be constructed in the usual way, i.e., the basis element $E_{ij} \in \mathcal{S}(n)$, where $i \geq j$, is a symmetric matrix whose ij and ji elements are one, and the remaining elements zero.

For $P \in \mathcal{P}(n)$ and $V \in \mathcal{S}(n)$, the matrix exponential map $\exp : \mathcal{S}(n) \rightarrow \mathcal{P}(n)$ and matrix logarithm map $\log : \mathcal{P}(n) \rightarrow \mathcal{S}(n)$ defined by

$$\exp(V) = \sum_{k=0}^{\infty} \frac{V^k}{k!} = I + V + \frac{V^2}{2!} + \frac{V^3}{3!} + \dots \quad (2.2.2)$$

and

$$\log(P) = - \sum_{k=1}^{\infty} \frac{(I - P)^k}{k} = (I - P) - \frac{(I - P)^2}{2!} - \frac{(I - P)^3}{3!} + \dots \quad (2.2.3)$$

are both uniquely defined and inverses of each other.

2.2.1 Affine-invariant Riemannian Metric

For $P \in \mathcal{P}(n)$ and $X, Y \in T_P\mathcal{P}(n) = \mathcal{S}(n)$, the Riemannian metric invariant under the group action

$$G * P = GPG^T, \quad (2.2.4)$$

where $G \in GL(n)$ is any $n \times n$ nonsingular matrix, also called the affine-invariant metric [16], is given by

$$\langle X, Y \rangle_P = \frac{1}{2} \text{tr} (P^{-1}XP^{-1}Y). \quad (2.2.5)$$

Then the length of a curve $P(t) \in \mathcal{P}(n)$, $a \leq t \leq b$, can be given by

$$l(P) = \int_a^b \sqrt{\langle \dot{P}(t), \dot{P}(t) \rangle_{P(t)}} dt. \quad (2.2.6)$$

The minimal geodesic $\gamma : [0, 1] \rightarrow \mathcal{P}(n)$ connecting two arbitrary points $P_1, P_2 \in \mathcal{P}(n)$, is given by

$$\gamma(t) = P_1^{1/2} (P_1^{-1/2} P_2 P_1^{-1/2})^t P_1^{1/2} \quad (2.2.7)$$

The tangent vector of the geodesic at P_1 is defined by the Riemannian logarithm map

$$\text{Log}_{P_1}(P_2) = P_1^{1/2} \log(P_1^{-1/2} P_2 P_1^{-1/2}) P_1^{1/2}. \quad (2.2.8)$$

The inverse of the Riemannian log map, the Riemannian exponential map is also defined. Given an element $V \in \mathcal{S}(n)$, the minimal geodesic emanating from a point $P \in \mathcal{P}(n)$ in the direction of the tangent vector V can be computed as follows:

$$\text{Exp}_P(V) = P^{1/2} \exp(P^{-1/2} V P^{-1/2}) P^{1/2}. \quad (2.2.9)$$

The distance between two arbitrary points P_1, P_2 can be defined in the usual way by the length of the above minimal geodesic. This can be calculated as the norm of $\text{Log}_{P_1}(P_2)$

at P_1 given by

$$d_{\mathcal{P}(n)}(P_1, P_2) = \|\text{Log}_{P_1}(P_2)\|_{P_1} = \left(\sum_{i=1}^n (\log(\lambda_i))^2 \right)^{1/2} \quad (2.2.10)$$

where λ_i are the eigenvalues of $P_1^{-1/2}P_2P_1^{-1/2}$, or equivalently, those of $P_1^{-1}P_2$. It is straightforward to check that the geodesic distance $d_{\mathcal{P}(n)}(\cdot, \cdot)$ derived above from the affine-invariant Riemannian metric (2.2.5) is also invariant under the $GL(n)$ group action on P_1 and P_2 .

2.2.2 Log-det Bregman Divergence

The Bregman divergence is widely used as a pseudo-distance metric in the applied mathematics and engineering field (e.g., in information geometry, machine learning, optimization, etc.). It generally does not satisfy fundamental requirements for a distance metric such as symmetry or the triangle inequality condition. However, the Bregman divergence remains of great interest for many practical and efficient algorithms where the use of exact distance is intractable or limited.

The Bregman divergence associated with a strictly convex function $h : \Omega \rightarrow \mathbb{R}$ for points $p, q \in \Omega$, is defined by the difference between the value of h at p and the value of the first-order Taylor expansion of h around q evaluated at p :

$$d_h(p||q) = h(p) - h(q) - \nabla h(q)^T(p - q).$$

Below are some general properties of the Bregman divergence:

- (i) Non-negativity: $d_h(p||q) \geq 0$ for all $p, q \in \Omega$ and $d_h(p||q) = 0$ if and only if $p = q$.
- (ii) Convexity: $d_h(p||q)$ is strictly convex in its first argument p .

(iii) Linearity: For two convex and differentiable functions h_1 and h_2 , and a nonnegative scalar λ , $d_{h_1+\lambda h_2}(p||q) = d_{h_1}(p||q) + \lambda d_{h_2}(p||q)$ holds.

(iv) Pythagorean theorem: Let $\text{Proj}_{\mathcal{C}}(p_0)$ be the projection of p_0 onto a convex set $\mathcal{C} \subset \Omega$ defined by

$$\text{Proj}_{\mathcal{C}}(p_0) \triangleq \arg \min_{q \in \mathcal{C}} d_h(q||p_0). \quad (2.2.11)$$

Then

$$d_h(p||p_0) \geq d_h(p||\text{Proj}_{\mathcal{C}}(p_0)) + d_h(\text{Proj}_{\mathcal{C}}(p_0)||p_0) \quad (2.2.12)$$

holds for all $q \in \mathcal{C}$.

(v) Approximation: the following second-order approximation generally holds:

$$d_h(p||p + dp) = \frac{1}{2} dp^T \nabla^2 h(p) + o(\|dp\|^2), \quad (2.2.13)$$

where $\nabla^2 h$ is the hessian of h .

Now, when the domain Ω is restricted to $\mathcal{P}(n)$, the Bregman divergence associated with a minus log-det function, $h(P) = -\log |P|$, $P \in \mathcal{P}(n)$, is given by

$$d_h(P_1||P_2) = \log \frac{|P_2|}{|P_1|} + \text{tr}(P_2^{-1}P_1) - 4 \quad (2.2.14)$$

$$= \sum_{i=1}^4 (-\log(\lambda_i) + \lambda_i - 1), \quad (2.2.15)$$

where λ_i are the eigenvalues of $P_2^{-1}P_1$, or equivalently, $P_2^{-1/2}P_1P_2^{-1/2}$. The log-det divergence d_F above also satisfies affine invariance (i.e., it is invariant under the $GL(n)$ group action $*$ on P_1 and P_2). Further, as $h(P) = -\log |P|$ is convex, the log-det divergence (2.2.14) is convex with respect to its first argument P_1 .

Remark 1 The Bregman divergence of a log-det (2.2.14) is also referred to as Stein's loss [17], the Burg matrix divergence [18], or simply the log-det divergence [19].

Remark 2 The log-det divergence d_h approximates the affine-invariant Riemannian metric (2.2.5) up to second-order; that is, for two infinitesimally close positive-definite matrices $P, P + dP \in \mathcal{P}(n)$, the following holds:

$$d_h(P + dP || P) = \underbrace{\frac{1}{2} \text{tr}(P^{-1} dP P^{-1} dP)}_{\langle dP, dP \rangle_P} + o(\lambda_i(P^{-1} dP)^2) \quad (2.2.16)$$

2.3 Matrix Lie Groups

A Lie group \mathbf{G} is a group which is a differentiable manifold with the smooth product and inverse group operations, i.e., $(X, Y) \mapsto X * Y$ and $X \mapsto X^{-1}$, where $*$ and $(\cdot)^{-1}$ respectively denote the product and inverse operations, which are both smooth functions for $X, Y \in \mathbf{G}$. The Lie algebra \mathfrak{g} associated with a Lie group \mathbf{G} is defined as the tangent vector space at the identity element of \mathbf{G} with the bilinear operation called the Lie bracket $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ satisfying

- (i) Skew-symmetry: $[x, y] = -[y, x]$,
- (ii) Jacobi identity: $[x, [y, z]] + [y, [x, z]] + [z, [x, y]] = 0$,

for all $x, y, z \in \mathfrak{g}$.

Matrix Lie group is a Lie group whose elements can be identified as matrices. For general linear group $GL(n)$ consisting of real $n \times n$ nonsingular matrices and its better known subgroups, the associated product and inverse operations can be defined by ordinary matrix multiplication and matrix inverse respectively. Further, tangent space at the identity element together with the matrix commutator identified as the Lie bracket operator constitutes a Lie algebra.

2.3.1 The Rotation Group

The special orthogonal group $SO(n)$ consists of the real $n \times n$ orthogonal matrices with determinant one, i.e.,

$$SO(n) = \{R \in \mathbb{R}^{n \times n} \mid RR^T = R^T R = \mathbb{1}_n, \det(R) = +1\}, \quad (2.3.17)$$

where we denote $\mathbb{1}_n$ to be $n \times n$ identity matrix. $SO(3)$ possesses the structure of both a differentiable manifold and a continuous group (under matrix multiplication), and is a well-known example of a matrix Lie group. From the definition it immediately follows that $R^{-1} = R^T$. The Lie algebra $so(n)$, which recall can be identified with the tangent space to $SO(n)$ at the identity element, consists of the real $n \times n$ skew-symmetric matrices:

$$so(n) = \{\Omega \in \mathbb{R}^{n \times n} \mid \Omega^T = -\Omega\}. \quad (2.3.18)$$

The Lie bracket of two elements $U, V \in so(n)$ is given by their matrix commutators:

$$[U, V] = UV - VU. \quad (2.3.19)$$

When $n = 3$, $R \in SO(3)$ represents the orientation of a rigid body situated in \mathbb{R}^3 . The Lie algebra $so(3)$ consists of 3×3 real skew-symmetric matrices of the form

$$\begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}. \quad (2.3.20)$$

For convenience, given any $\omega \in \mathbb{R}^3$, we denote its 3×3 skew-symmetric matrix representation $[\cdot] : \mathbb{R}^3 \rightarrow so(3)$ as

$$[\omega] = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}. \quad (2.3.21)$$

The Lie bracket between two elements in $so(3)$ can be further identified with the vector product in \mathbb{R}^3 as follows. Setting $U = [u] \in so(3)$ and $V = [v] \in so(3)$ for some $u, v \in \mathbb{R}^3$, a straightforward calculation reveals that

$$\begin{aligned} [U, V] &= [u][v] - [v][u] \\ &= [u \times v]. \end{aligned}$$

The Lie bracket can also be expressed using the (small) adjoint map representation $ad_u : so(3) \rightarrow so(3)$ as

$$ad_u(v) = u \times v. \quad (2.3.22)$$

Finally given any $R \in SO(3)$ and $[\omega] \in so(3)$, one can associate with R the (large) adjoint map $Ad_R : so(3) \rightarrow so(3)$ as follows:

$$Ad_R([\omega]) = R[\omega]R^T = [R\omega]. \quad (2.3.23)$$

We can write the above more compactly in vector form as $Ad_R(\omega) = R\omega$.

2.3.2 The Euclidean Group

The special Euclidean group $SE(n)$ consists of the real $(n+1) \times (n+1)$ real matrices of the form

$$\begin{bmatrix} R & p \\ 0 & 1 \end{bmatrix}, \quad (2.3.24)$$

where $R \in SO(n)$ and $p \in \mathbb{R}^n$ (the 0 in the last row denotes an n -dimensional row vector of zeros). In the case of most interest to us, $n = 3$, $SE(3)$ describes the set of all possible relative displacements between two reference frames situated in physical space. The Lie algebra of $SE(3)$, denoted $se(3)$, consists of matrices of the form,

$$\begin{bmatrix} [\omega] & v \\ 0 & 0 \end{bmatrix}, \quad (2.3.25)$$

where $[\omega] \in so(3)$ and $v \in \mathbb{R}^3$. The above matrix can also be represented more compactly in six-dimensional vector form via the following notation. Given an element

$$\begin{bmatrix} [\omega] & v \\ 0 & 0 \end{bmatrix} = [S] \in se(3), \quad (2.3.26)$$

the equivalent six-dimensional vector representation of $[S]$ will be denoted

$$S = \begin{bmatrix} \omega \\ v \end{bmatrix}. \quad (2.3.27)$$

The Lie bracket of two elements $S_1 = (\omega_1, v_1) \in se(3)$, $S_2 = (\omega_2, v_2) \in se(3)$ can be evaluated as follows:

$$[[S_1], [S_2]] = [S_1][S_2] - [S_2][S_1] \quad (2.3.28)$$

$$= \begin{bmatrix} [\omega_1 \times \omega_2] & [\omega_1]v_2 - [\omega_2]v_1 \\ 0 & 0 \end{bmatrix}. \quad (2.3.29)$$

The above Lie bracket can be denoted more compactly in terms of the vector representation S_1, S_2 . Using the (small) adjoint representation $\text{ad}_{S_1}(S_2)$ to denote the Lie bracket between S_1 and S_2 , we write

$$\text{ad}_{S_1}(S_2) = \begin{bmatrix} [\omega_1] & 0 \\ [v_1] & [\omega_1] \end{bmatrix} \begin{bmatrix} \omega_2 \\ v_2 \end{bmatrix}. \quad (2.3.30)$$

Given $S = (\omega, v) \in se(3)$, the 6×6 matrix representation of the (small) adjoint mapping $\text{ad}_S : se(3) \rightarrow se(3)$ is denoted

$$[\text{ad}_S] = \begin{bmatrix} [\omega] & 0 \\ [v] & [\omega] \end{bmatrix}. \quad (2.3.31)$$

Given $X = \begin{bmatrix} R & p \\ 0 & 1 \end{bmatrix} \in SE(3)$ and $[S] = \begin{bmatrix} [\omega] & v \\ 0 & 0 \end{bmatrix} \in se(3)$, the (large) adjoint mapping $\text{Ad}_X : se(3) \rightarrow se(3)$ is defined as

$$\text{Ad}_X([S]) = X[S]X^{-1} \quad (2.3.32)$$

$$= \begin{bmatrix} [R\omega] & Rv + [p]R\omega \\ 0 & 0 \end{bmatrix}. \quad (2.3.33)$$

In vector form the above can also be written as

$$\text{Ad}_X(S) = \begin{bmatrix} R & 0 \\ [p]R & R \end{bmatrix} \begin{bmatrix} \omega \\ v \end{bmatrix}. \quad (2.3.34)$$

The 6×6 matrix representation of the linear map Ad_X is denoted

$$[\text{Ad}_X] = \begin{bmatrix} R & 0 \\ [p]R & R \end{bmatrix}. \quad (2.3.35)$$

3

Geometric Dynamic Identification of Multibody Systems

3.1 Introduction

The identification of the mass and inertial parameters of a robot, and more generally of any system that can be modeled and represented as a rigid multibody system, e.g., digital characters and humans, is a prerequisite to any dynamic model-based method for motion planning and control. Not surprisingly, there is an extensive literature on methods for inertial parameter identification [15, 20, 21, 22, 23], also referred to as dynamic calibration in analogy to the even more well-studied problem of kinematic calibration. Nearly all of the previous approaches exploit the fact that a robot's dynamics are linear in the mass and inertial parameters. Typical approaches take force/torque and kinematic measurements at multiple points along some reference trajectory, and formulate the identification problem as a least-squares problem [15]. However, the identification of inertial parameters using these standard least-squares methods turns out to be highly sensitive

to various aspects of the measurement conditions.

Trajectories used in identification should be sufficiently rich and result in a well-conditioned regressor matrix so that the identified parameters are not overly sensitive to measurement noise. For standard open-chain manipulators, achieving these conditions is often feasible [24]. For more complex robots like humanoids and legged robots, however, possible reference motions are usually highly restricted, e.g., the robot must always maintain balance, and the actuators must be capable of generating needed torques. Also, with the growing emphasis on making robots more affordable, measurements obtained from low-cost sensors typically will be incomplete and quite noisy. Several works [25, 26] have shown that, for high-dof systems, the accuracy of the parameters identified using standard least-squares estimation methods is often far from satisfactory.

An appealing approach to make identification robust to the quality of excitation trajectories and measurement noise is to exploit prior knowledge for the inertial parameters. Regularization toward some known nominal values for inertial parameters (e.g., obtained from CAD data) is helpful when reference motion trajectories are not maximally exciting or when excitation is ill conditioned [25, 26]. Also, one can explicitly constrain parameters to ones that are physically feasible. Inertial parameters are in fact subject to a set of positive-definiteness requirements dictated by physics, also referred to as the physical consistency condition. The physical consistency requirements can be strictly guaranteed via a set of linear matrix inequality constraints, leading to a semidefinite programming problem [27], [28]; orthogonal decomposition-based manifold optimization techniques offer an alternate solution to this problem [29]. However, unless it happens that the unconstrained least squares minimizer is already physically consistent, solutions to these inequality constrained formulations will necessarily lie on the boundary of physical consistency, e.g., some of the resulting links may end up being flat along one or more dimensions, or reduce to a point mass or even zero mass.

In this Chapter, we claim that such ill-posed behavior is traceable in large part to the use of standard Euclidean distance metric (more specifically, as a regularizer to an ill-posed least squares error criterion that does not account for prior information in a physically consistent and natural way). The key idea of our approach rests on the fact that the collection of inertial parameters constitutes a Riemannian manifold with a naturally defined Riemannian metric. Applying a one-to-one mapping between a link's inertial parameters and the space of 4×4 symmetric positive-definite matrices first pointed out in [28], we formulate and minimize a coordinate-invariant error criterion based on this natural Riemannian metric. We demonstrate that regularization based on this coordinate-invariant notion of distance can significantly improve the accuracy and robustness of mass-inertial parameter estimation, even for high-dimensional systems subject to noisy measurements and ill-conditioned reference trajectories.

Yet, a continued drawback of existing geometric methods is that the corresponding optimization problems are inherently non-convex, have numerous local minima, and are computationally highly intensive to solve. To this, we propose a convex formulation under the same coordinate-invariant Riemannian geometric framework that directly addresses these and other deficiencies of the geometric approach. The main idea behind our approach is a second-order approximation of the Riemannian distance that allows for the convex regularization of the inertial parameter identification problem. Not only does this allow for the use of fast convex optimization algorithms that are guaranteed to converge to a global solution, but also ensures coordinate-invariance while allowing for the inclusion of additional convex constraints as imposed by physical considerations and other practical requirements. We also show that different choices for the distance metric used in regularization can be viewed as a choice of a projection mapping to some fixed subset of the Riemannian manifold of mass-inertial parameters. With this insight, we propose a quantitative method for uniformly comparing the performance of

different mass-inertial parameter estimation methods. Both the practical and qualitative advantages of our approach are validated through extensive identification case studies involving the AMBIDEX manipulator, the MIT Cheetah 3, and a human subject.

3.2 Preliminaries

In this section, we first characterize a physically consistent set of rigid body inertial parameters. Then, we briefly review conventional linear least squares based approaches for inertial parameter identification, and point out geometric issues and problems induced by the ignored geometry of the underlying parameter space.

3.2.1 Physically Consistent Rigid Body Inertial Parameters

The full set of inertial parameters for a single rigid body is conventionally characterized by its mass, first mass moment, and rotational inertia tensor, all described in some body-fixed reference frame $\{b\}$. The first mass moment is denoted as $h_b = m \cdot p_b \in \mathbb{R}^3$ where m is the mass and $p_b \in \mathbb{R}^3$ is the position of center of mass (CoM). The rotational inertia is given by the 3×3 symmetric tensor $I_b \in \mathcal{S}(3)$. Hence, there are ten independent parameters for a single rigid body often represented in a vectorized form as

$$\phi_b = \phi_b(m, h_b, I_b) = [m, h_b^T, I_b^{xx}, I_b^{yy}, I_b^{zz}, I_b^{xy}, I_b^{yz}, I_b^{zx}]^T \in \mathbb{R}^{10}. \quad (3.2.1)$$

Also, recall the 6×6 spatial inertia tensor description of inertial parameters [7, 9],

$$G(\phi_b) = \int_{\Omega} \begin{bmatrix} [x_b] \\ \mathbb{I}_3 \end{bmatrix} \begin{bmatrix} [x_b] \\ \mathbb{I}_3 \end{bmatrix}^T \rho(x_b) dV_b = \begin{bmatrix} I_b & [h_b] \\ [h_b]^T & m\mathbb{I}_3 \end{bmatrix} \in \mathcal{S}(6). \quad (3.2.2)$$

where $\rho : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}_0^+$ is a non-negative mass density function for a given rigid body. The spatial inertia $G(\phi_b)$ is known to serve as a kinetic energy metric for a rigid

body, i.e.,

$$T = \frac{1}{2} V_b^T G(\phi_b) V_b, \quad (3.2.3)$$

where $V_b = (w_b, v_b) \in se(3)$ is the 6×1 vector representation of a spatial body velocity. The positive-definiteness condition on the spatial inertia tensor,

$$G(\phi_b) \succ 0, \quad (3.2.4)$$

is also equivalent to $m > 0$ and $I_b^C = I_b - [h_b][h_b]^T/m \succ 0$ by the Schur complement lemma.¹ Satisfaction of these conditions is sufficient for $G(\phi_b)$ to be a valid kinetic energy metric. However, these conditions are insufficient for inertial parameters to be physically realizable from some non-negative mass density function [29, 28].

Toward characterizing an exact condition for *density-realizability* of inertial parameters, authors in [28] employed a 4×4 symmetric matrix representation of the parameters called the pseudo-inertia matrix. Given a rigid body, its pseudo inertia is given by

$$P(\phi_b) = \int_{\Omega} \begin{bmatrix} x_b \\ 1 \end{bmatrix} \begin{bmatrix} x_b \\ 1 \end{bmatrix}^T \rho(x_b) dV_b = \begin{bmatrix} \Sigma_b & h_b \\ h_b^T & m \end{bmatrix} \in \mathcal{S}(4), \quad (3.2.5)$$

where

$$\Sigma_b = \int_{\Omega} x_b x_b^T \rho(x_b) dV_b \in \mathcal{S}(3) \quad (3.2.6)$$

is the *density-weighted second moment matrix*. This matrix has a one-to-one linear correspondence with I_b as $\Sigma_b = \frac{1}{2} \text{tr}(I_b) \mathbb{I}_3 - I_b$ and likewise $I_b = \text{tr}(\Sigma_b) \mathbb{I}_3 - \Sigma_b$. A physical consistency condition that is both necessary and sufficient is simply the positive definiteness of the pseudo inertia, i.e.,

$$P(\phi_b) \succ 0. \quad (3.2.7)$$

¹By the parallel axis theorem, I_b^C is the rotational inertia matrix about the center of mass

In addition to ensuring a positive mass and positive definite rotational inertia, the condition (3.2.7) also ensures that the principal moments of inertia satisfy a set of triangle inequalities [29] given by

$$\lambda_1(I_b^C) + \lambda_2(I_b^C) > \lambda_3(I_b^C), \quad (3.2.8)$$

$$\lambda_2(I_b^C) + \lambda_3(I_b^C) > \lambda_1(I_b^C), \quad (3.2.9)$$

$$\lambda_3(I_b^C) + \lambda_1(I_b^C) > \lambda_2(I_b^C). \quad (3.2.10)$$

Satisfaction of these inequalities is the critical ingredient for $P(\phi_b) \succ 0$ to be sufficient for physical plausibility of ϕ_b . To illustrate with a simple example, consider the rotational inertia tensor matrix at the center of mass given by $I_b^C = \text{diag}(1, 0.3, 0.3)$ which is positive definite but does not satisfy the triangle inequality condition, i.e., $0.3 + 0.3 < 1$. The corresponding second moment matrix is calculated as $\Sigma_b^C = \frac{1}{2} \text{tr}(I_b^C) \mathbb{I}_3 - I_b^C = \text{diag}(-0.2, 0.5, 0.5) \neq 0$. Since the density-weighted covariance matrix of any non-negative mass density function is positive definite, we can infer that the inertial parameters giving rise to I_b^C are not physically feasible.

Relying on the pseudo inertia representation, authors in [28] also introduced a constraint on ϕ_b that ensures the existence of an associated mass density function supported within a bounding ellipsoid. Consider an ellipsoid defined by

$$\mathcal{E} = \{x_b \in \mathbb{R}^3 : \begin{bmatrix} x_b \\ 1 \end{bmatrix}^T Q_b \begin{bmatrix} x_b \\ 1 \end{bmatrix} \leq 0\}, \quad (3.2.11)$$

The inertial parameters ϕ_b can be realized through mass purely inside \mathcal{E} if and only if ϕ_b satisfies (3.2.7) and

$$\text{tr}(P(\phi_b) Q_b) \leq 0. \quad (3.2.12)$$

Parameters ϕ_b satisfying these conditions are called \mathcal{E} -density realizable. Since $P : \mathbb{R}^{10} \rightarrow \mathcal{S}(4)$ is a one-to-one linear mapping, the positive definiteness requirement (3.2.7)

together with the linear inequality constraint (3.2.12) imposes a convex feasible set for ϕ_b in \mathbb{R}^{10} .

3.2.2 Linear Least Squares based Identification

Owing to the structure of a general coupled rigid-body system, any inertial parameter identification problem boils down to fitting a system of linear equations [15].

To illustrate, for a n -link open chain with joint variables $q \in \mathbb{R}^n$ and joint torque vector $\tau \in \mathbb{R}^n$, the dynamic equations are of the form

$$\tau = M(q, \Phi_b)\ddot{q} + b(q, \dot{q}, \Phi_b) = Y(q, \dot{q}, \ddot{q})\Phi_b, \quad (3.2.13)$$

where

$$\Phi_b = [\phi_{b_1}^T, \dots, \phi_{b_n}^T]^T \in \mathbb{R}^{10n} \quad (3.2.14)$$

denotes the complete set of mass and inertial parameters for the n links, $M(q, \Phi_b) \in \mathbb{R}^{n \times n}$ is the mass matrix, and $b(q, \dot{q}, \Phi_b) \in \mathbb{R}^n$ denotes the vector of Coriolis and gravitational forces; these terms can be collected and represented as $Y(q, \dot{q}, \ddot{q})\Phi_b$, where $Y(q, \dot{q}, \ddot{q}) \in \mathbb{R}^{n \times 10n}$. Then, measurements of τ , q , \dot{q} , \ddot{q} at multiple points along some reference trajectory each impose n linear equations $Y\Phi_b = \tau \in \mathbb{R}^n$. These quantities can be stacked to form a single large overdetermined system of linear equations of the form

$$A\Phi_b = b. \quad (3.2.15)$$

Remark 3.1. *For a floating-base multibody system, e.g. humanoid robots and humans, the evolution of the net system momentum [23], [25] is commonly adopted for identification purposes rather than focusing on the full-body dynamics. Perhaps remarkably, when the system is an open-chain kinematic tree, this approach provides an identical observable set of inertial parameters as in the full dynamics case. Since this reduction*

only considers effects of external forces on the total change of momentum, non-rigid components, e.g., nonlinear friction or elasticity in the joint-space dynamics, can be ignored. Therefore, this approach can be particularly useful for pure inertial parameter identification purposes when inertias involved in closed kinematic loops (e.g., motor rotors) are negligible. An identical formulation can also be used for open-chain manipulators with force/torque sensor embedded at the base link [30].

Now, the inertial parameter identification problem is commonly formulated as a regularized least-squares estimation problem of the form

$$\min_{\Phi_b} \underbrace{\|A\Phi_b - b\|_{\Sigma^{-1}}^2}_{\text{measurement error}} + \gamma \underbrace{\|\Phi_b - {}^0\Phi_b\|^2}_{\text{regularizer}}. \quad (3.2.16)$$

The first term represents the squared measurement error; elements of the diagonal matrix $\Sigma = \text{diag}(\sigma_1^2, \dots, \sigma_m^2) \in \mathbb{R}^{m \times m}$ weight each of the residuals according to the uncertainty or variance of the measurements. If we have perfect measurements of a full rank matrix A , and b is subject to zero-mean Gaussian noise of covariance Σ , then the least-squares minimizer without the regularization term is known to be the maximum-likelihood solution with the desirable statistical property of being asymptotically unbiased. Since in practice A is neither full rank nor perfectly measured, the regularization term with scale factor γ is added to make the problem well-posed, and to keep the estimated parameter values close to the given nominal values ${}^0\Phi_b$. Ignoring the physical consistency constraints, the optimal least-squares solution can be expressed in closed form as $\Phi_b^{LS} = (A^T \Sigma^{-1} A + \gamma I)^{-1} (A^T \Sigma^{-1} b + \gamma \cdot {}^0\Phi_b)$.

However, as has been observed also in the recent literature [27], [25], pure unconstrained least squares solutions are generally not reliable in terms of physical consistency, and attempts have been made toward obtaining physically consistent optimal solutions, i.e. from projection-based or recursive trial and error based methods (these

and related methods are surveyed in [27]), to more recent methods based on orthogonal decomposition-based manifold optimization techniques [29], or imposing a set of linear matrix inequality (LMI) constraints on the link inertial parameters ϕ_{b_i} as $P(\phi_{b_i}) \succeq 0$ [27], [28] in the optimization formulation. The latter approach of imposing LMI guarantees not only physically consistent and global minimum solution, but also leads to a semidefinite programming problem, for which reliable and efficient numerical methods are widely available. However, as depicted in Figure 3.1, the constrained solution Φ_b^{LS+LMI} may end up lying on the boundary of the manifold \mathcal{M}^n (i.e., a rank-deficient positive-semidefinite matrix P_{b_i} , leading to an i^{th} rigid link that is, e.g., flat along one dimension, a point mass, or even zero mass) whenever the unconstrained global minimum Φ_b^{LS} lies outside \mathcal{M}^n due to the convexity of the original objective function and the search space $\mathcal{M}^n \simeq P(4)^n$. The problem typically becomes even more pronounced for high-dof structures like humans and humanoids, and it is not uncommon to enforce stronger sufficiency conditions for physical consistency [26], [25], such as imposing user-specified convex bounds on, e.g., the link masses and the locations of mass centers.

However, most importantly, we claim that it is the use of the standard Euclidean distance metric, specifically as a regularizer to an ill-posed least squares error criterion that doesn't account for prior information in a physically consistent and invariant way. The inertial parameter identification problem is unlike other parameter fitting problems, e.g., arbitrary function approximation, in that the parameters to be identified are physical quantities. While a choice of physical units and reference frames is necessary for a numerical implementation, a well-posed problem for inertial parameter estimation should not depend on these choices. That is, the physical meaning of the identified parameters should be the same regardless of physical units and reference frames in the problem formulation. As will be further elucidated in the following sections, the use of standard

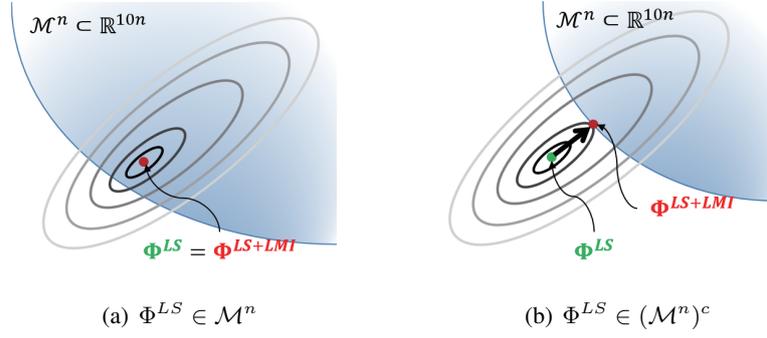


Figure 3.1: The effect of imposing LMI while minimizing a least-squares objective function. The grey ellipsoidal contour indicates the least-squares objective function while the blue-shaded region is the feasible convex region of the search space in \mathbb{R}^{10n} defined by the LMI. 3.1(a) When the pure least-squares minimizer Φ^{LS} is physically consistent, it coincides with the least squares minimizer subject to the LMI, Φ^{LS+LMI} . 3.1(b) When Φ^{LS} is not physically consistent, Φ^{LS+LMI} lies on the boundary of \mathcal{M}^n .

Euclidean metric as a regularizer not only makes the identification coordinate dependant, but also fail to capture the natural geometric structure of the underlying parameter space. This eventually makes the identification sensitive and leads to a physically inconsistent unconstrained estimate Φ_b^{LS} , especially when the measurement data is insufficient and noisy. In line with the preceding discussion on the nature of the constrained solution, this implies that under the conventional regularized least squares formulation as in (3.2.16), imposing the physical consistency constraints generally will not improve matters much, and the solutions are highly susceptible to lying on the boundary of the user-specified feasible region.

From the Maximum a Posteriori (MAP) estimation point of view of the formulation (3.2.16), the regularization term exactly represents a prior distribution on Φ_b . In fact, the regularization term can be replaced by an arbitrary quadratic function, $\|\Phi_b - {}^0\Phi_b\|_{Q^{-1}}$

for some positive-definite matrix Q . The formulation is equivalent to a MAP estimation problem where the likelihood is given by a Gaussian density function, $p(b|A, \Phi_b) = \mathcal{N}(A\Phi_b, \Sigma)$, and the prior is given by a Gaussian density function, $p(\Phi_b) = \mathcal{N}({}^0\Phi_b, Q/\gamma)$. In most practical situations, we are only provided with the prior nominal value ${}^0\Phi_b$ from, e.g., CAD. By contrast, the uncertainty information for the nominal value, corresponding to the covariance of the prior, Q/γ , is hardly obtainable from any source. For this reason, the identity matrix as in (3.2.16), or some scale-normalizing diagonal matrix is commonly chosen for Q , leaving the scalar γ as a user-tuned hyperparameter. These somewhat limiting choices of distance metric always result in coordinate-dependent estimation of inertial parameters.

3.3 Geometry of Rigid Body Inertial Parameters

This section is devoted to defining a natural Riemannian geometric structure of the feasible inertial parameter space \mathcal{M} , which is to be exploited for geometric, coordinate-invariant formulations of inertial parameter identification problem. The content provided here will also generally serve as a core basis for developing geometric algorithms in the later Chapters.

Recall that the physical consistency conditions on $\phi_b \in \mathbb{R}^{10}$ can be identified with the requirement that $P_b = P(\phi_b) \in \mathcal{P}(4)$ be positive-definite. The manifold \mathcal{M} of the set of physically consistent inertial parameters for a single rigid body is defined as follows:

$$\begin{aligned} \mathcal{M} &\simeq \{\phi_b \in \mathbb{R}^{10} : P(\phi_b) \succ 0\} \subset \mathbb{R}^{10} \\ &\simeq \{P_b \in \mathcal{S}(4) : P_b \succ 0\} = \mathcal{P}(4), \end{aligned}$$

where \simeq denotes an isomorphic relationship between sets. The elements can be identified in both \mathbb{R}^{10} and $\mathcal{P}(4)$, also for different choices of body-fixed reference frame $\{b\}$.

Alternatively, instead of ϕ_b one can use the coordinate (m, p_b, Σ_b^C) , where p_b locates the CoM and

$$\Sigma_b^C = \int_{\Omega} (x_b - p_b)(x_b - p_b)^T \rho(x_b) dV_b = \Sigma_b - m p_b p_b^T \quad (3.3.17)$$

is the density-weighted second moment matrix at the CoM. However, the coordinate representation (m, p_b, Σ_b^C) is unlike P_b not linear with respect to ϕ_b .

3.3.1 Riemannian Distance Metric

Taking advantage of the fact that the manifold \mathcal{M} of physically consistent inertial parameters of each link is equivalent to $\mathcal{P}(4)$, the Riemannian manifold structure of \mathcal{M} can naturally be inherited from that of $\mathcal{P}(4)$ defined in Section 2.2 of the previous Chapter 2.

A differential metric on \mathcal{M} is chosen as

$$ds^2 = \langle d\phi_b, d\phi_b \rangle_{\phi_b} \quad (3.3.18)$$

$$\triangleq \frac{1}{2} \text{tr} \left((P(\phi_b)^{-1} P(d\phi_b))^2 \right) \quad (3.3.19)$$

$$\begin{aligned} &= \frac{1}{2} \left(\frac{dm}{m} \right)^2 + \frac{1}{2} \text{tr} \left(\left((\Sigma_b^C)^{-1} d\Sigma_b^C \right)^2 \right) \\ &\quad + m \cdot dp_b^T (\Sigma_b^C)^{-1} dp_b, \end{aligned} \quad (3.3.20)$$

matching the affine-invariant Riemannian metric (2.2.5) on $\mathcal{P}(4)$ (see Appendix A.1, Proposition A.1 for a proof of (3.3.20)). Note that the inertial parameters consist of diverse physical quantities (e.g., m [mass], p_b [length], Σ_b^C [mass·length²]). Inspecting each of the terms in (3.3.20), it can be seen that the affine-invariant Riemannian metric offers a physical unit- and scale-free way of measuring distances on the space of inertial parameters.

We now show that this distance metric is coordinate-invariant.

Proposition 3.1. *The Riemannian metric defined in (3.3.19) is invariant with respect to choice of body-fixed reference frame and physical unit/scale.*

Proof. Under a change of body-fixed reference frame from $\{b\}$ to $\{a\}$ and physical units associated with length and mass, three dimensional position vector $r \in \mathbb{R}^3$ and volume element dV transform according to the following rules:

$$\begin{aligned} r_a &= \text{diag}(d_{ab}^x, d_{ab}^y, d_{ab}^z) \cdot (R_{ab}r_b + t_{ab}) \\ &= F_{ab}r_b + g_{ab} \end{aligned} \quad (3.3.21)$$

$$\rho_a(\cdot)dV_a = c_{ab} \cdot \rho_b(\cdot)dV_b, \quad (3.3.22)$$

where $d_{ab}^k \in \mathbb{R}^+$ and $c_{ab} \in \mathbb{R}^+$ are the change in physical scale/unit of length and mass, respectively, and $R_{ab} \in SO(3)$, $t_{ab} \in \mathbb{R}^3$ represent the rigid body transformation between frames $\{a\}$ and $\{b\}$. Using the homogeneous representation of a 3-D vector, (3.3.21) can be rewritten as

$$\begin{bmatrix} r_a \\ 1 \end{bmatrix} = H_{ab} \cdot \begin{bmatrix} r_b \\ 1 \end{bmatrix}, \quad (3.3.23)$$

where $H_{ab} \triangleq \begin{bmatrix} F_{ab} & g_{ab} \\ 0 & 1 \end{bmatrix} \in GL(4)$. Now from (3.2.5), (3.3.22), (3.3.23), the pseudo-inertia matrix representation of the inertial parameters admits the following coordinate

transformation rule in the form of a $GL(4)$ group action * (2.2.4):

$$\begin{aligned}
P(\phi_a) &= \int_{\Omega} \begin{bmatrix} r_a \\ 1 \end{bmatrix} \begin{bmatrix} r_a \\ 1 \end{bmatrix}^T \rho_a(r_a) dV_a \\
&= c_{ab} \cdot H_{ab} \left(\int_{\Omega} \begin{bmatrix} r_b \\ 1 \end{bmatrix} \begin{bmatrix} r_b \\ 1 \end{bmatrix}^T \rho_b(r_b) dV_b \right) H_{ab}^T \\
&= c_{ab} \cdot H_{ab} P(\phi_b) H_{ab}^T \\
&= S_{ab} P(\phi_b) S_{ab}^T = S_{ab} * P_b,
\end{aligned} \tag{3.3.24}$$

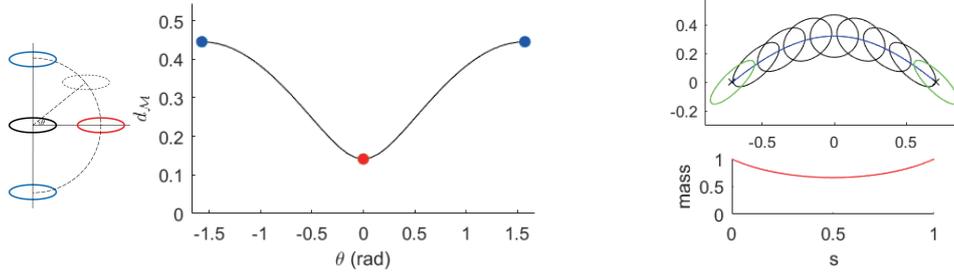
where $S_{ab} = \sqrt{c_{ab}} \cdot H_{ab} \in GL(4)$. Since the affine-invariant Riemannian metric defined in (3.3.19) is invariant under the $GL(4)$ group action, it is thus invariant to the change of reference frame and physical unit/scale. \square

On the other hand, the standard Euclidean metric for $\phi_b \in \mathbb{R}^{10}$, i.e., $ds_{\text{Euc}}^2 = d\phi_b^T d\phi_b = dm^2 + dp_b^T dp_b + (dI_b^{xx})^2 + (dI_b^{yy})^2 + \dots$, satisfies none of these invariance properties. The metric does not have meaningful physical units, unless each of the components are normalized by appropriate constants. In adopting the Riemannian metric, we now omit all the coordinate frame subscripts for expressing the inertial parameters as ϕ and P .

Remark 3.2. *The Riemannian metric (3.3.19) is also closely related to the log-determinant barrier function on $\mathcal{P}(4)$. Consider $P, P+dP \in \mathcal{P}(4)$. From a Taylor series expansion,*

$$\begin{aligned}
-\log|P+dP| &= -\log|P| - \text{tr}(P^{-1}dP) \\
&\quad + \frac{1}{2} \text{tr}\left((P^{-1}dP)^2\right) + o(\|dP\|^2).
\end{aligned}$$

The differential metric on \mathcal{M} can therefore be recognized as following from the Hessian of the log-determinant barrier [31].



(a) Plot of geodesic distances between a pair of mass density functions (b) Geodesic path between two mass density functions

Figure 3.2: Minimal geodesic path between two mass density functions (green ellipses) under the Riemannian metric 3.3.19. Each mass density function is visualized by an ellipse whose center and shape matches the center of mass and covariance matrix, respectively. 3.2(a) shows a plot of the distances from the fixed mass density function (black ellipse) to ones with the same mass and covariance, but constant deviation in positions of center of mass. 3.2(b) shows the geodesic path .

Under the metric 3.3.19, the distance between two arbitrary physically consistent inertial parameters of a single rigid body ${}^1\phi_b, {}^2\phi_b \in \mathcal{M}$ can be defined with the geodesic distance $d_{\mathcal{P}(4)}$ (2.2.10) as

$$d_{\mathcal{M}}({}^1\phi_b, {}^2\phi_b)^2 = d_{\mathcal{P}(4)}({}^1P_b, {}^2P_b)^2, \quad (3.3.25)$$

where ${}^iP_b = P({}^i\phi_b) \in \mathcal{P}(4)$ for $i = 1, 2$. For an n -link multibody system, the distance metric can be defined via direct summation as

$$d_{\mathcal{M}^n}({}^1\Phi_b, {}^2\Phi_b)^2 = \sum_{i=1}^n d_{\mathcal{M}}({}^1\phi_{b_i}, {}^2\phi_{b_i})^2. \quad (3.3.26)$$

We note how the proposed metric measures distances and produces geodesic paths between mass density functions depending on the location of the center of mass p and

covariance Σ . Figure 3.2 shows that deviations in the center of mass along directions of shorter radii of the covariance tend to produce larger distances between density functions. As shown in Figure 3.2(a), the pair of density functions whose centers of masses are aligned along the major principal axes produces the shortest distance. Moreover, in Figure 3.2(b), the initial and final directions of the minimal geodesics are both aligned toward the major principal axes. Such tendencies are actually a highly desirable property for treating inertial parameter data and an important distinction from using the Euclidean metric. For instance, if a given rigid body is known to have the shape of a long bar with unknown density distribution, then it is desirable for the estimated center of mass to deviate more along the stretched direction of the bar than the perpendicular direction. With the present distance metric, we can naturally take this into account in the estimation, while the Euclidean metric fails to do so.

Lastly, observe that the distance metric (3.3.25) is well-defined on the space $\mathcal{P}(4)$; that is, when an element in $\mathcal{P}(4)$ approaches the boundary, i.e., as it becomes positive semi-definite, its distance to any fixed element of $\mathcal{P}(4)$ grows to infinity. This behavior is due to the metric being defined via a logarithmic function. This property implies that the Riemannian distance serves as barrier function that inherently enforces the physical consistency of the estimated inertial parameters without imposing it as a hard constraint.

Remark 3.3. *We further note the close relation between the proposed metric and the Fisher information metric on multivariate normal distribution [32]. Let us define a submanifold \mathcal{N} of \mathcal{M} whose total mass m is fixed to 1. Then the manifold \mathcal{N} is equivalent to the statistical manifold of multivariate Gaussian distributions. The line element ds^2 induced onto the submanifold \mathcal{N} can be expressed under $m = 1$ and $dm = 0$ as*

$$ds_{\text{induced}}^2 = dp^T (\Sigma^C)^{-1} dp + \frac{1}{2} \text{tr} \left(\left((\Sigma^C)^{-1} d\Sigma^C \right)^2 \right). \quad (3.3.27)$$

The induced metric (3.3.27) on \mathcal{N} is exactly the Fisher information metric on the statistical manifold of multivariate normal distributions, widely used in information geometry as a coordinate-invariant Riemannian metric. Although \mathcal{N} turns out to be non-geodesic submanifold of \mathcal{M} , it is of interest to note that the characteristics of the metric described in Figure 3.2 correspond to the well-known ones of the Fisher information metric [33].

3.3.2 Entropic Divergence

Recall that the (Bregman) log-det divergence defined on $\mathcal{P}(n)$ (2.2.14) is a second-order approximation of the affine-invariant Riemannian distance. Using the one-to-one mapping from ϕ to $P = P(\phi)$, we can define a divergence measure on \mathcal{M} as

$$d_{\mathcal{M}}({}^1\phi || {}^2\phi)^2 \triangleq d_h({}^1P || {}^2P), \quad (3.3.28)$$

where ${}^iP = P({}^i\phi) \in \mathcal{P}(4)$ for $i = 1, 2$. Since the log-det divergence (2.2.14) is affine invariant, and pseudo inertia P exhibits a change of coordinate in a form of $GL(4)$ group action $*$, it follows that this divergence measure is also coordinate invariant.

Similar to the Riemannian distance (3.3.25), the divergence is well-defined on $\mathcal{P}(4)$ and tends to infinity as arguments approach the boundary of the positive semi-definite cone. This behavior is likewise due to the metric being defined via a logarithmic function and implies that the divergence also serves as a barrier function.

More explicitly, in the coordinates (m, p, Σ^C) , the divergence metric defined in (3.3.28) can be decomposed as (c.f. (3.3.20))

$$\begin{aligned} d_{\mathcal{M}}({}^1\phi || {}^2\phi)^2 = & \underbrace{d_h({}^1m || {}^2m)}_{\text{mass}} + \underbrace{d_h({}^1\Sigma^C || {}^2\Sigma^C)}_{\text{shape and orientation}} \\ & + \underbrace{{}^1m ({}^1p - {}^2p)^T ({}^2\Sigma^C)^{-1} ({}^1p - {}^2p)}_{\text{position (CoM)}} \end{aligned} \quad (3.3.29)$$

The proof is given in Appendix A.1, Proposition A.2. Note that the mass m and second moment matrix Σ^C are elements of $\mathcal{P}(1)$ and $\mathcal{P}(3)$, respectively, on which the divergence (2.2.14) is well-defined. As is evident from (3.3.29), the proposed divergence measure simultaneously encodes the scale-free difference for both the total mass and the geometric distribution of the mass density, similar to the Riemannian distance (3.3.25).

Remark 3.4. *In (3.3.29), when both masses are fixed to one, the proposed divergence measure coincides with the Kullback-Leibler (KL) divergence [34] on multivariate Gaussians, which is known to capture the geometric (e.g., position, orientation and shape) difference of two normalized densities. (See Appendix A.1, Proposition A.3 for proof.) We use the term “entropic” for our divergence measure since the KL divergence can be derived as the Bregman divergence associated with the negative entropy of a normalized density function [35]. Our entropic divergence can be viewed as a particular case of a generalized KL divergence for unnormalized densities like mass densities.*

3.3.3 Constant Pullback Metric

In addition to the entropic divergence, another direct approximation of the geodesic distance (3.3.25) can be given by considering a constant Riemannian metric evaluated at some representative inertial parameter value ${}^0\phi \in \mathbb{R}^{10}$ near the region of interest, e.g., the prior estimate; that is, from (3.3.20), a simplified differential metric on \mathcal{M} can be defined as

$$ds_0^2 \triangleq \frac{1}{2} \text{tr} \left((P({}^0\phi)^{-1} P(d\phi))^2 \right) \quad (3.3.30)$$

$$= d\phi^T g({}^0\phi) d\phi, \quad (3.3.31)$$

where a constant positive definite matrix $g({}^0\phi)$ can be straightforwardly constructed considering the fact that (3.3.30) is a quadratic form with respect to $d\phi$. Formally, $g(\cdot)$

is the pullback of the affine-invariant Riemannian metric on $\mathcal{P}(4)$ to \mathbb{R}^{10} under the mapping $P(\cdot)$. An expression for $g(\cdot)$ is provided in Appendix B.1. With this definition, the constant pullback distance metric on $\phi \in \mathbb{R}^{10}$ can be defined as

$$d_0({}^1\phi, {}^2\phi)^2 \triangleq ({}^1\phi - {}^2\phi)^T g({}^0\phi) ({}^1\phi - {}^2\phi), \quad (3.3.32)$$

or equivalently,

$$d_0({}^1\phi, {}^2\phi)^2 = \frac{1}{2} \text{tr} \left((P({}^0\phi)^{-1} P({}^1\phi - {}^2\phi))^2 \right). \quad (3.3.33)$$

Like the Riemannian distance and entropic divergence, it is straightforward to check from (3.3.33) that the constant pullback distance is dimensionless, and coordinate and scale invariant.

Remark 3.5. *Mathematically, with a constant Riemannian metric $g({}^0\phi)$ (3.3.31), the ensuing space \mathcal{M} is isometric to a flat Euclidean space; considering the reparametrization $\phi' = g({}^0\phi)^{1/2}\phi$, the differential metric is given by $ds_0^2 = d\phi'^T d\phi'$. Meanwhile, unlike the naive standard Euclidean metric on ϕ , the constant pullback metric allows a reasonable approximation of the curved Riemannian structure near ${}^0\phi$; approximations far from ${}^0\phi$ can be inaccurate, however. To demonstrate, the quadratic distance (3.3.32) has a finite value on the boundary of \mathcal{M} , while the Riemannian distance (and also entropic divergence) have infinite values.*

3.3.4 Distribution Awareness of Geometric Distances

Figure 3.3 and 3.4 present multiple cases of perturbations to rigid-body mass distributions and compares how the perturbations are captured using different distance measures. As shown in Figure 3.3(a), the geometric distance measures—the Riemannian distance, entropic divergence and constant pullback distance—produce the shortest distances for the pair of mass distributions whose centers of mass are aligned along the major principal axes; physically, the densities maximally overlap in such configurations. As also

discussed in Figure 3.2, the proposed geometric distance measures can naturally penalize the estimation of CoM to lie inside the body through regularization. In contrast, the Euclidean distance shows little to no meaningful deviation, as the Euclidean distance on the first moment h is constant without regard to changes in the configuration (θ), and changes in the moment of inertia are encoded at a much smaller magnitude than other quantities under the Euclidean metric in SI units.

In Figure 3.4(b), the inertial parameters of a body are deliberately perturbed by adding or subtracting multiple point masses at uniformly distributed positions inside the body (see Appendix B.2 for further details). The magnitude of the mass perturbations are set proportional to the mass of the object. It can be observed from Figure 3.4(b) that the geometric distance measures are insensitive to scale changes in the target mass, while the Euclidean distance monotonically increases as the absolute magnitude of the point mass noise increases. This phenomenon can be understood in terms of the scale-invariance of geometric distances (see the first term in (3.3.20) or (3.3.29)).

Remark 3.6. *Observe from Figure 3.3(a), 3.4(a), 3.3(b) that the Euclidean metric shows conspicuously different distance evaluations just by changing the system of units, while geometric distances are invariant with respect to unit/scale choices.*

Moreover, in Figure 3.4(a), as the body shrinks to a point mass, the Riemannian distance and entropic divergence to the target inertia diverge to infinity. More specifically, since the effective volume of the density is captured in the determinant of the second moment matrix Σ^C , the log-det function in the second term of (3.3.29) makes the distance approach infinity as the volume shrinks. In contrast, as alluded to earlier, distances to the zero volume inertia (located on the boundary of \mathcal{M}) are quantified as finite values for the quadratic Euclidean and constant pullback distances.

Remark 3.7. *The scale-invariant property together with the logarithmic landscape of the*

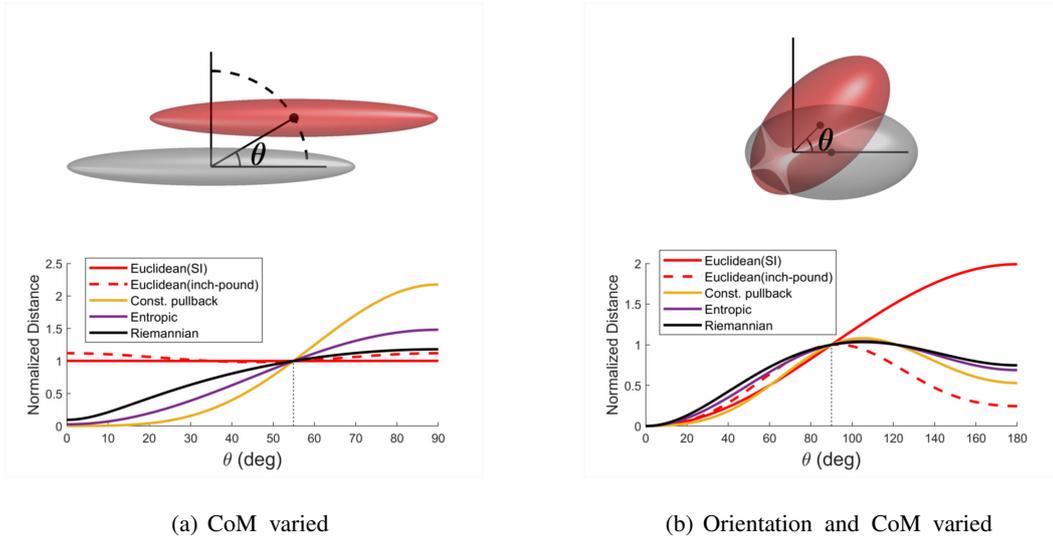


Figure 3.3: Comparisons of squared distances between inertial parameters of ellipsoidal bodies, each of uniform density and drawn to the same scale. In each plot, the distances are normalized with the value at a particular configuration of the two bodies (marked on the plots with black vertical dashed lines). Two unit systems are used to evaluate the distances: kg-m (SI unit) and inch-pound. The inertial parameters of the target bodies (grey ellipsoids) are used to evaluate the pullback form of the Riemannian metric for the constant pullback metric. The distance from the fixed target density (grey ellipsoid) to the perturbed ones (red ellipsoids) are compared while varying the position and orientation, but with total mass fixed.

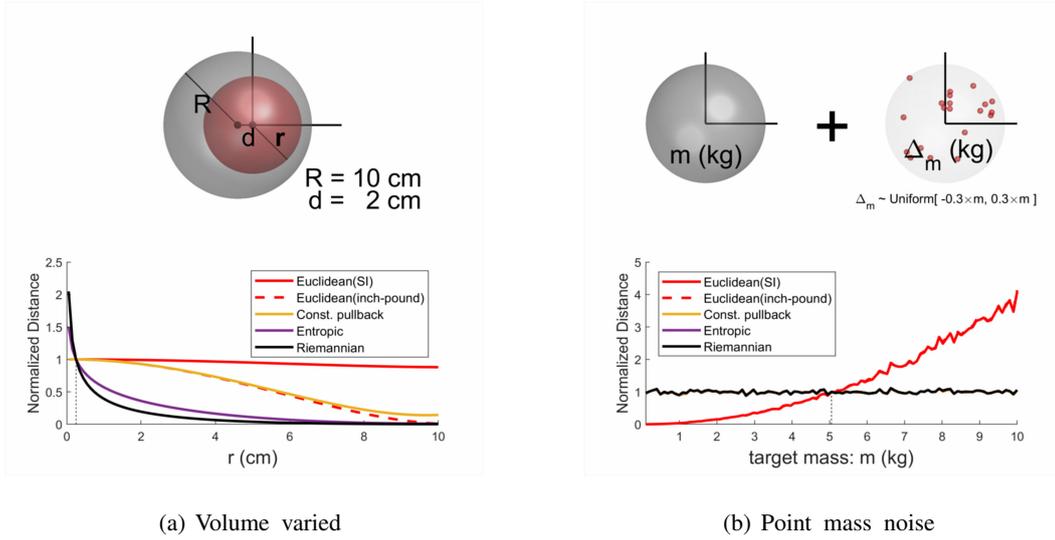


Figure 3.4: Figures 3.4(a) compares the distance from the fixed target density (grey ellipsoid) to the perturbed ones (red ellipsoids) while varying the total volume, but with total mass fixed. Figure 3.4(b) compares the distance from the target body (grey sphere) with varied total mass m to the identical body perturbed by adding uniformly distributed point mass noise inside the body whose total mass Δ_m is also drawn from the uniform distribution on $[-0.3 \cdot m, 0.3 \cdot m]$. For each case of the target body with total mass m , the mean of the distances to 500 samples of the perturbed bodies is evaluated; the deviation between the “normalized” Euclidean (SI, inch-pound) distances is very small so that their plots in Figure 3.4(b) almost overlap each other. The “normalized” geometric (const. pullback, entropic and Riemannian) distances likewise have minimal deviation from one another.

distance can naturally prevent the estimation of inertial parameters from being overly sensitive to lighter/smaller links relative to heavier/larger ones.

3.4 Geometric Identification with Geodesic Least Squares

In this section, we frame the inertial parameter identification problem as an optimization problem on the Riemannian manifold \mathcal{M}^n . By exploiting the natural affine-invariant Riemannian structure of \mathcal{M}^n , a physically meaningful objective function can be formulated and optimized with iterative cyclic procedures of taking geodesic projections and a geometric mean.

3.4.1 Intrinsic Riemannian Error Criterion

Recall identifying the physically consistent inertial parameters of a multibody system reduces to finding $\Phi \in \mathcal{M}^n \subset \mathbb{R}^{10n}$ that best fits the system of linear equations (3.2.15), $A\Phi = b$. Referring to Figure 3.5, the problem can be equivalently restated in a geometrical way as finding Φ that is closest to each of the hyperplanes

$$\mathcal{H}_i \triangleq \{x \in \mathbb{R}^{10n} : a_i^T x = b_i\}, \quad i = 1, \dots, m, \quad (3.4.34)$$

where $a_i \in \mathbb{R}^{10n}, b_i \in \mathbb{R}$ are the rows of A and b , respectively, and at the same time resides in \mathcal{M}^n . Viewed from this perspective, ordinary least-squares minimizes the weighted sum of squares of the Euclidean distances from Φ to each hyperplane and the prior value ${}^0\Phi$, i.e., referring to (3.2.16),

$$\min_{\Phi, \{\gamma_i\}_{i=1}^m} \sum_{i=1}^m \gamma_i \|\Phi - {}^i\Phi\|^2 + \gamma \|\Phi - {}^0\Phi\|^2 \quad (3.4.35)$$

$$\text{s.t. } a_i^T \Phi = b_i, \quad i = 1, \dots, m, \quad (3.4.36)$$

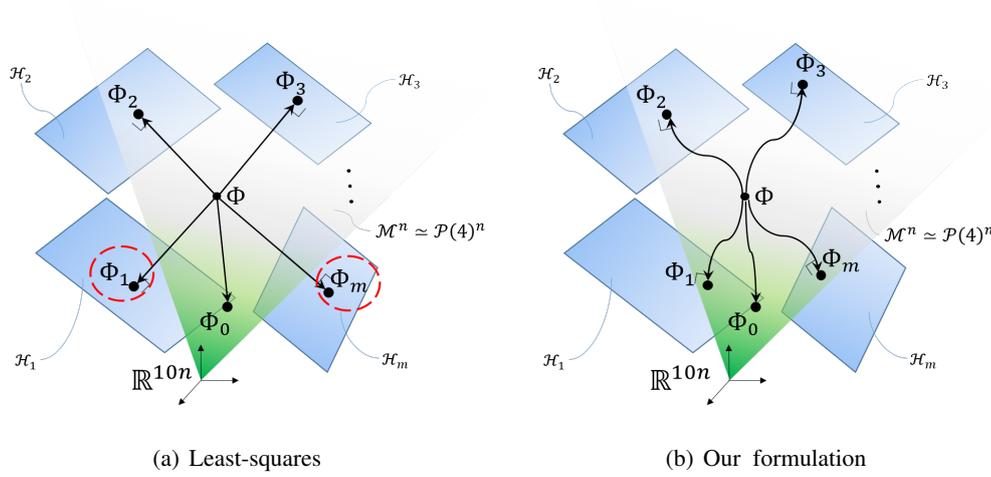


Figure 3.5: Pictorial description of the geometric meaning of the physically consistent inertial parameter identification problem. The green cone-shaped shaded region indicates the manifold $\mathcal{M}^n \simeq \mathcal{P}(4)^n$ embedded in \mathbb{R}^{10n} . The blue hyperplanes each represent the linear constraints in $A\Phi = b$. Each of the black arrows indicates the projection of Φ to the hyperplanes whose paths depend on the metric. Red dashed circles denote the projections of Φ to physically inconsistent values outside the green shaded region.

where $\gamma_i = \|a_i\|^2/\sigma_i^2$ and the ${}^i\Phi$ are a set of slack variables used in the equality-constrained optimization, physically corresponding to the projected points from Φ to each of the hyperplanes \mathcal{H}_i . Note however that the notion of projecting a point to the hyperplane depends on the metric defined on the space. As can be seen in (3.4.35) and (3.4.36), the least-squares formulation adopts an Euclidean metric on the parameter space; this can cause the projected points ${}^i\Phi$, marked as red dashed circles in Figure 3.5(a), to be physically inconsistent. Since Φ eventually converges to the weighted arithmetic mean of ${}^i\Phi$ s and ${}^0\Phi$, any physical inconsistencies in ${}^i\Phi$ can adversely affect the estimation of Φ .

The key idea in our approach is to use the natural metric on the physically consistent parameter space \mathcal{M}^n . The corresponding optimization problem formulation can be achieved by substituting the Euclidean distance $\|(\cdot) - (\cdot)\|$ with the natural geodesic distance $d_{\mathcal{M}^n}(\cdot, \cdot)$:

$$\min_{\Phi, \{\mathcal{H}_i\}_{i=1}^m} \sum_{i=1}^m \gamma_i \cdot d_{\mathcal{M}^n}(\Phi, \mathcal{H}_i)^2 + \gamma \cdot d_{\mathcal{M}^n}(\Phi, \mathcal{H}_0)^2 \quad (3.4.37)$$

$$\text{s.t.} \quad \sum_{j=1}^n \text{tr}({}^i P_j X_{ij}) = b_i, \quad i = 1, \dots, m, \quad (3.4.38)$$

where ${}^i P_j = P({}^i \phi_j) \in \mathcal{P}(4)$, and the constraint $\mathcal{H}_i \in \mathcal{H}_i$ is expressed in matrix form as (3.4.38) with $X_{ij} \in \mathcal{S}(4)$, which is equivalent to vector constraints (3.4.36) (See Appendix B.1 for a construction of the matrix W_{ij}). With the proposed formulation, physical consistency can be guaranteed without explicitly constraining the parameters to be physically consistent, due to the well-defined metric-based projection and geometrical means on \mathcal{M}^n .

Remark 3.8. *The squared measurement error in (3.2.16) is invariant to the linear transformations $\Phi \rightarrow M^{1/2}\Phi$ and $a_i \rightarrow M^{-1/2}a_i$ for any positive-definite matrix M . Therefore a constant dimensionless choice of $\gamma_i = a_i^T M_0^{-1} a_i / \sigma_i^2$ with unique (pullback metric) M_0 satisfying $d\Phi^T M_0 d\Phi \equiv \sum_{j=1}^n \langle P(d\phi_j), P(d\phi_j) \rangle_{P({}^0 \phi_j)}$ is preferred over the choice $\gamma_i = \|a_i\|^2 / \sigma_i^2$.*

Remark 3.9. *Directly regularizing the linear least squares error, i.e.,*

$$\min_{\Phi} \|A\Phi - b\|_{\Sigma^{-1}}^2 + \gamma \cdot d_{\mathcal{M}^n}(\Phi, \mathcal{H}_0)^2, \quad (3.4.39)$$

is also possible and coordinate invariant. Convex relaxations of this form is to be adopted in Section 3.5 for allowing fast, computationally efficient algorithms with guaranteed convergence. Still, the formulation (3.4.37), (3.4.38) provides a more geometrically intuitive algorithm with a robust feature for rejecting anomaly data as we argue below.

3.4.2 Cyclic Optimization Algorithm

The objective function (3.4.37) is minimized by an iterative cyclic optimization procedure, alternatively optimizing the slack variables $\{^i\Phi\}_{i=1}^{mk}$ and the inertial parameters Φ with an initial guess ${}^0\Phi$; geometrically this corresponds to repeated projection, and finding the means of the projected points. We now explain these two steps in more detail.

- **Optimizing $\{^i\Phi\}_{i=1}^m$: Projection of Φ to \mathcal{H}_i**

Given fixed Φ , each of the $^i\Phi$ can be optimized in parallel, each $^i\Phi$ updated via the following subproblem:

$$\begin{aligned} {}^i\Phi &= \arg \min_{\hat{\Phi}} d_{\mathcal{M}^n}(\Phi, \hat{\Phi})^2 \\ \text{s.t. } \hat{\Phi} &\in \mathcal{H}_i \end{aligned}$$

which is exactly the definition of point-set projection in the metric space. However, it turns out that there exist multiple ($\sim O(2^{4n})$) of extremum points. Rather than exhaustively finding all possible extrema, the computational burden can be reduced by restricting the search space to the following neighborhood of Φ :

$$\mathcal{B}_{\Phi} = \{\hat{\Phi} \in \mathcal{M}^n : \lambda_{\max}(P(\phi_i)^{-1}P(\hat{\phi}_i)) \leq e, \forall i = 1, \dots, n\}.$$

It can be shown that there exists at most a single local minimum in \mathcal{B}_{Φ} . In practice, most of the local minima $^i\Phi \in \mathcal{B}_{\Phi}$ are found to exist, except for cases where the measurement noise is excessive. We now propose the following point (Φ)-set $(\mathcal{H} \cap \mathcal{B}_{\Phi})$ projection rule on \mathcal{M}^n :

Proposition 3.2. *The natural projection of $\Phi \in \mathcal{M}^n \subset \mathbb{R}^{10n}$ onto $\mathcal{H} \cap \mathcal{B}_{\Phi}$, where the hyperplane \mathcal{H} is defined by $\mathcal{H} = \{\Phi \in \mathbb{R}^{10n} : a^T\Phi = b\} \simeq \{\{P_j\}_{j=1}^n \in$*

$\mathcal{S}(4)^n : \sum_{i=1}^n \text{tr}(P_j X_j) = b$, is uniquely determined if and only if the monotonically decreasing function

$$\mathcal{C}(\lambda) = \sum_{j=1}^n \sum_{k=1}^4 \sigma_j^k e^{-W(\lambda \sigma_j^k)} - b \quad (3.4.40)$$

has a unique root $\hat{\lambda}$ on the interval $[-g(\sigma_{max}), g(-\sigma_{min})]$, where $W(\cdot)$ is an inverse function of $w : [-1, \infty) \rightarrow \mathbb{R}$, $w(x) = xe^x$,

$$g(\sigma) = \begin{cases} 1/(e \cdot \sigma), & \text{if } \sigma > 0 \\ +\infty, & \text{otherwise} \end{cases}$$

σ_j^k are the eigenvalues of $\bar{X}_j = P_j^{1/2} X_j P_j^{1/2}$, and σ_{max} , σ_{min} are the largest and smallest eigenvalues of $\{\bar{X}_j\}_{j=1}^n$.

Then the unique projected point $\hat{\Phi}$ is given by

$$\hat{\phi}_j = P^{-1}(\hat{P}_j) = P^{-1}(P_j^{1/2} e^{Q_j} P_j^{1/2}) \quad (3.4.41)$$

where $Q_j = R_j \cdot \text{diag}(-W(\hat{\lambda} \sigma_j^k)) \cdot R_j^T$ and $R_j \in SO(4)$ is obtained from the eigendecomposition of $\bar{X}_j = R_j \Sigma R_j^T$.

Proof. The proof is given in the Appendix A.2. □

Note that because of the monotonicity of \mathcal{C} , existence of the unique solution $\hat{\lambda}$ can be verified by simply evaluating \mathcal{C} at the boundary of the specified interval. If the solution exists, the root is found numerically. Otherwise, the hyperplane \mathcal{H} lies outside \mathcal{M}^n , or the projected point lies outside \mathcal{B}_Φ , in which case the corresponding constraint is simply discarded at the initial stage.

- **Optimizing Φ : Geometric Mean of $\{^i\Phi\}_{i=1}^m$**

The optimal Φ given fixed $\{^i\Phi\}_{i=1}^n$ is obtained by solving an unconstrained optimization problem of the form

$$\Phi = \arg \min_{\hat{\Phi}} \sum_{i=1}^{mk} \gamma_i \cdot d_{\mathcal{M}^n}(\hat{\Phi}, ^i\Phi)^2 + \gamma \cdot d_{\mathcal{M}^n}(\hat{\Phi}, ^0\Phi)^2, \quad (3.4.42)$$

which geometrically corresponds to the point that minimizes the weighted sum of squares of geodesic distances from $\{^i\Phi\}_{i=1}^n$ and $^0\Phi$. Therefore, the optimal Φ can be thought of as an extended version of the mean defined on the manifold; this same definition has been studied in the literature, known as a geometric mean [16] or Karcher mean [36]. The Karcher mean is known to uniquely exist for $\mathcal{P}(n)$. Moreover there exist gradient-based algorithms for numerically finding the Karcher mean. We use the Matlab function *manopt* [37] to solve (3.4.42).

3.5 Geometric Identification with Convex Relaxations

In this section, we propose a convex programming approach to the geometric identification of inertial parameters by considering entropic divergence and constant pull-back distance metric for regularization. With the proposed approach, one can realize distribution-aware identification of physically consistent inertial parameters in a well-conditioned manner while guaranteeing fast convergence to the unique global solution using semi-definite programming solvers.

Since the residual error term $\|A\Phi - b\|_{\Sigma^{-1}}^2$ in (3.2.16) or (3.4.39) is clearly independent on the coordinate choice, the distance metric in the regularization term needs to be coordinate invariant. Convex geometric regularization can be accomplished by replacing the regularization term in (3.2.16) with convex geometric distances; for instance, using

the entropic divergence,

$$\min_{\Phi} \|A\Phi - b\|_{\Sigma^{-1}}^2 + \gamma \cdot d_{\mathcal{M}^n}(\Phi \| {}^0\Phi)^2, \quad (3.5.43)$$

where $d_{\mathcal{M}^n}(\Phi \| {}^0\Phi)^2 \triangleq \sum_{i=1}^n d_{\mathcal{M}}(\phi_i \| {}^0\phi_i)^2$.

Remark 3.10. From the definition (3.3.28),

$$d_{\mathcal{M}^n}(\Phi \| {}^0\Phi)^2 = \sum_{i=1}^n d_h(P(\phi_i) \| P({}^0\phi_i)) \quad (3.5.44)$$

holds. Since d_h is convex with respect to its first argument and $P(\cdot)$ is a linear mapping, the entropic divergence regularizer (3.5.44) is also convex with respect to $\Phi = [\phi_1^T, \dots, \phi_n^T]^T$.

As noted in Remark 3.7, the entropic regularization term acts as a barrier function that inherently enforces the physical consistency of the estimated inertial parameters without imposing it as a hard constraint. Moreover, the convex formulation (3.5.43) can also be extended to include additional convex inequality constraints, e.g., the \mathcal{E} -density realization condition (3.2.12), by imposing $\text{tr}(P(\phi_i) Q_i) \geq 0$, $i = 1, \dots, n$.

The problem (3.5.43) can be reformulated in coordinates $(P_1, \dots, P_n) = (P(\phi_1), \dots, P(\phi_n))$ with \mathcal{E} -density realizability constraints as

$$\begin{aligned} \min_{\{P_i\}_{i=1}^n} \sum_{j=1}^m \left[\left(\sum_{i=1}^n \text{tr}(P_i X_{ij}) - b_j \right)^2 / \sigma_j^2 \right] \\ + \gamma \sum_{i=1}^n (-\log |P_i| + \text{tr}({}^0P_i^{-1} P_i)) \end{aligned} \quad (3.5.45)$$

$$\text{s.t.} \quad \text{tr}(P_i Q_i) \geq 0, \quad i = 1, \dots, n. \quad (3.5.46)$$

The quantities $X_{ij} \in \mathcal{S}(4)$ are the unique 4×4 symmetric matrices satisfying $a_j^T \Phi = \sum_{i=1}^n \text{tr}(P_i X_{ij})$, where $a_j^T \in \mathbb{R}^{10}$ is the j -th row vector in matrix $A \in \mathbb{R}^{m \times 10n}$, b_j is

the j -th scalar component in vector $b \in \mathbb{R}^m$, and σ_j^2 is the j -th diagonal entry in the diagonal matrix $\Sigma \in \mathbb{R}^{m \times m}$ (See Appendix B.1 for a construction of the matrix X_{ij}).²

Quadratic error regularization with the constant pullback metric can be formulated as

$$\min_{\Phi} \|A\Phi - b\|_{\Sigma^{-1}}^2 + \gamma \cdot d_0(\Phi, {}^0\Phi)^2 \quad (3.5.47)$$

$$\text{s.t. } P(\phi_i) \succ 0, \quad (3.5.48)$$

$$\text{tr}(P(\phi_i)Q_i) \geq 0, \quad i = 1, \dots, n, \quad (3.5.49)$$

where

$$d_0(\Phi, {}^0\Phi)^2 \triangleq (\Phi - {}^0\Phi)^T G({}^0\Phi) (\Phi - {}^0\Phi) \quad (3.5.50)$$

and the pullback $G({}^0\Phi) \in \mathcal{P}(10n)$ for the full parameter set is given by

$$G({}^0\Phi) = \text{diag}(g({}^0\phi_1), \dots, g({}^0\phi_n)) . \quad (3.5.51)$$

The use of this constant pullback metric can be practically appealing as a convex regularizer since it is quadratic in the vector representation Φ . Meanwhile, as the quadratic error distance (3.5.50) fails to naturally serve as a barrier function for physically consistent parameters, the inclusion of physical consistency constraints (3.5.48) is required in the optimization. This drawback is similar to the Euclidean metric but in contrast to the Riemannian distance metric and entropic divergence.

3.5.1 Provable Comparative Analysis Scheme

Comparing the generalizability performance of different regularizers raises a challenge to ensure fairness. Maintaining the same scalar regularization factor γ in (3.2.16) across

²For a non-diagonal covariance matrix Σ , let the eigendecomposition be $\Sigma = R\Lambda R^T$. Then, one can simply substitute the variables as $A \leftarrow R^T A$, $b \leftarrow R^T b$, $\Sigma \leftarrow \Lambda$ to always make Σ a diagonal matrix without changing the formulation.

different regularization types may fail to equally balance data fit with fit to the prior. For example, when the magnitude of the regularization error increases by switching the regularizer, naively fixing the value of γ can lead to increased bias toward the prior value. To avoid this issue, γ should be selected differently for each regularizer. For example, one could select γ to ensure a certain level of error in the training dataset, or to ensure optimal generalizability via cross-validation strategies.

In this section, we provide a methodology to compare the performance of different regularizers by guaranteeing a fixed residual error on the training dataset. The key idea rests on the fact that the regularized formulation (3.2.16) can be reformulated as a constrained point (${}^0\Phi$)-to-set (\mathcal{X}_c) projection problem:

$$\begin{aligned} \min_{\Phi} d(\Phi, {}^0\Phi)^2 & \quad (3.5.52) \\ \text{s.t. } \Phi \in \mathcal{X}_c & \triangleq \{\hat{\Phi} : \|A\hat{\Phi} - b\|_{\Sigma^{-1}}^2 \leq c\}, \end{aligned}$$

where the scalar c depends on the regularization weight γ . The proof that such a reformulation is always possible is given in Appendix A.3, Proposition A.4.

Here, we again highlight the importance of using the natural regularizer for ill-posed identification problems from an alternative geometric perspective. From the equivalence of the formulation (3.2.16) and (3.5.52), regularizing the least-squares error can be understood as finding the closest estimate to the nominal value ${}^0\Phi$ among the set \mathcal{X}_c with bounded least-squares error. When the regressor matrix A is highly ill-conditioned and rank-deficient, this property implies that the set \mathcal{X}_c will be highly anisotropic and unbounded in some directions. As a result, achieving a physically natural projection of the nominal value ${}^0\Phi$ to \mathcal{X}_c depends largely on the choice of distance metric. Figure 3.6 shows a conceptual sketch of projecting a prior inertial parameter ${}^0\Phi$ to \mathcal{X}_c using different distance metrics. A naive Euclidean projection can easily result in physically infeasible estimates, for which the estimation parameters are projected outside the

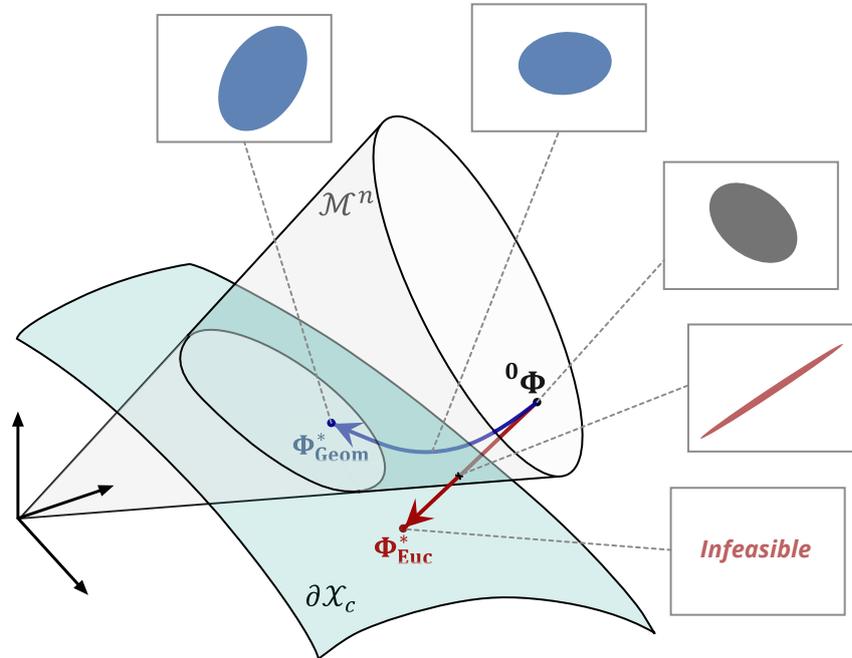


Figure 3.6: Conceptual sketch of solving a constrained point(${}^0\Phi$)-to-set(\mathcal{X}_c) projection problem for least-squares identification of inertial parameters with different distance measures for regularization. The grey cone represents the region of physically consistent inertial parameters, while the green set represents the boundary of the set \mathcal{X}_c with bounded least-squares error. Φ_{Euc}^* and Φ_{Geom}^* are the projected values of the prior value ${}^0\Phi$ to the set \mathcal{X}_c with Euclidean distance and geometric distance, e.g., entropic divergence or Riemannian distance, respectively. The ellipsoids depicted inside the grey boxes represent the visualization of the equivalent mass densities of the inertial parameters along the paths of projections.

positive-definite cone \mathcal{M}^n or converge to a point on the boundary of \mathcal{M}^n when the physical consistency condition is explicitly enforced. With this alternative constrained formulation in mind, below we directly present the proposed methodology followed by a rigorous justification of the method.

Given R different convex regularizers $d_1(\cdot, \cdot), \dots, d_R(\cdot, \cdot)$, the method proceeds as follows:

- (i) Select an arbitrary regularizer, without loss of generality $d_1(\cdot, \cdot)$, and a value for the regularization factor γ_1 .
- (ii) Solve a convex regularized least-squares problem (e.g., such as (3.2.16)) for the regularizer $d_1(\cdot, \cdot)$ with regularization factor γ_1 , and obtain the optimized residual error loss c .
- (iii) For $i = 2, \dots, R$, solve the corresponding convex bounded residual error formulation (e.g., such as (3.5.52)) with regularizer $d_i(\cdot, \cdot)$ and residual error bound c .

From Proposition A.5 in Appendix A.3, one can guarantee that the alternative optimizations in step iii have solutions that are equivalent to regularized formulations for some choice of factors γ_i . Further, each of these solutions provides equivalent training error. This means that each solution in step iii is always located at the boundary of \mathcal{X}_c . Moreover, the method is applicable to the case where additional optimization variables are included, e.g., friction parameters, or additional convex constraints are imposed, e.g., \mathcal{E} -density realizability (3.2.12), non-negativity constraints on friction parameters, etc.

In addition to convex regularized identification formulations of the form (3.2.16), there are other optimization formulations for inertial parameter identification that merit comparison. First, comparison to the non-convex Riemannian regularized identification

formulation may indicate how much is practically gained or lost by using the entropic divergence as a convex approximation. This formulation is given by

$$\min_{\Phi} \|A\Phi - b\|_{\Sigma^{-1}}^2 + \gamma \cdot d_{\mathcal{M}^n}(\Phi, {}^0\Phi)^2. \quad (3.5.53)$$

While the non-convexity of the Riemannian regularization prevents transforming this problem into an equivalent point-to-set projection problem, we instead rely on (3.5.53) with γ set to the value used for entropic divergence regularization. Since the Riemannian distance is approximated by the entropic divergence up to second order, using the same γ is reasonable.

Another formulation for identifying physically plausible inertial parameters was proposed in [38]. The mass distribution of each body of a robot is approximated by a finite number of point masses at fixed positions inside an a priori known region of support. Then, the inertial parameters $\Phi = [\phi_1^T, \dots, \phi_n^T] \in \mathbb{R}^{10n}$ can be represented linearly using a vector $\mu \in \mathbb{R}_{\geq 0}^{N_\rho}$ of N_ρ mass values at these locations. Mathematically, $\Phi = R\mu$ for some matrix $R \in \mathbb{R}^{10n \times N_\rho}$ that depends on the positions of the point masses. One of the core benefits of this reparametrization is that the physical consistency constraint on the pseudo-inertia (3.2.7) and \mathcal{E} -density realizability condition on Φ can be effectively approximated by simply enforcing all the virtual point masses to be non-negative, i.e., $\mu \geq 0$. In [38] the authors adopt a Euclidean metric on μ for regularized inertial parameter identification:

$$\begin{aligned} \min_{\mu} \quad & \|AR\mu - b\|_{\Sigma^{-1}}^2 + \gamma \cdot \|\mu - {}^0\mu\|^2 \\ \text{s.t.} \quad & \mu \geq 0, \end{aligned} \quad (3.5.54)$$

where the identified value of the inertial parameters can be retrieved by the relation $\Phi = R\mu$ after the optimization.

Remark 3.11. *Variables in the redundant specification of point masses μ can be chosen to have identical physical units of mass. Thus, unlike the original Euclidean metric on inertial parameters Φ , adopting the standard Euclidean metric on point values is physically more sensible.*

The point mass formulation (3.5.54), like our geometric formulation, also provides a means of distribution-aware regularization at the expense of introducing redundant point mass parameters. Comparison to this point mass regularization approach allows one to assess how much is practically gained or lost by regularizing with our geometric distance measures directly on a more compact representation of inertial parameters.

Since (3.5.54) is clearly a convex problem, the corresponding bounded least square error formulation

$$\min_{\mu \geq 0} \|\mu - {}^0\mu\|^2 \quad (3.5.55)$$

$$\text{s.t. } \|AR\mu - b\|_{\Sigma^{-1}}^2 \leq c \quad (3.5.56)$$

also has a unique global minimum solution. Like the aforementioned convex formulations, the least-square error constraint (3.5.56) is guaranteed to be active at any optimal solution if the prior value is designated such that $R^0\mu = {}^0\Phi$.

Remark 3.12. *The proposed geometric distances can simultaneously measure differences both in the geometric distribution of mass and in the densities. By comparison, fixed mass locations in the above point-mass formulation limit its flexibility to measure changes in the geometric distribution. Beyond this high-level benefit, we will show that geometric regularizers capture the main performance benefits of the point-mass formulation, while only requiring a prior for the inertial parameters.*

3.6 Experiments

This section benchmarks the proposed convex geometric distances for regularized identification alongside the Euclidean parameter metric, Euclidean point-mass metric and full Riemannian metric. We first consider identification of the AMBIDEX manipulator (fixed base), then the MIT Cheetah 3 (fixed-base identification of a floating-base system), and finally a human (full floating-base identification from low-cost measurement devices). This battery of tests allows us to characterize the effects of data scarcity and quality on the accuracy of the models identified and the plausibility of the inertial parameters that are found.

3.6.1 AMBIDEX Robot Manipulator

Identification experiments were conducted on the AMBIDEX, a seven-dof anthropomorphic robot manipulator jointly developed by KoreaTech and NAVER LABS. The core mechanical design of AMBIDEX is based on the LIMS manipulator [39]. Ten rigid body links are articulated by seven cables wired across the bodies. Each cable is actuated by an electric motor. The cable wiring results in a 1:30 torque amplification from motors to joints. Owing to the unique mechanical coupling (rolling joint mechanism) between three pairs of bodies, the system can be reduced to a seven-dof fully actuated system [39]. Denoting by $q \in \mathbb{R}^7$ the motor angles and by $\theta \in \mathbb{R}^{10}$ the joint angles, the dynamic equations for the motor torques $\tau \in \mathbb{R}^7$ are given by

$$\tau = J(q)^T \tau_\theta + \text{diag}(\ddot{q}) I_r + \text{diag}(\text{sign}(\dot{q})) f_c + \text{diag}(\dot{q}) f_v, \quad (3.6.57)$$

where $I_r = [I_{r_1}, \dots, I_{r_7}]^T \in \mathbb{R}^7$ is the vector of inertias of the rotors about the motor axes, $f_c \in \mathbb{R}^7$ and $f_v \in \mathbb{R}^7$ are the Coulomb and viscous friction coefficients, and $J(q) \in \mathbb{R}^{10 \times 7}$ is the full-rank Jacobian of the nonlinear mapping between the motor

angles and the joint angles. The joint torque vector $\tau_\theta \in \mathbb{R}^{10}$ is given by

$$\tau_\theta = M(\theta)\ddot{\theta} + C(\theta, \dot{\theta})\dot{\theta} + g(\theta) = \Gamma(\theta, \dot{\theta}, \ddot{\theta})\Phi, \quad (3.6.58)$$

where $\Phi = [\phi_1^T, \dots, \phi_{10}^T]^T \in \mathbb{R}^{100}$ is the vector of inertial parameters of the ten bodies.

The motor torque τ has a linear relation with respect to the inertial parameters $\bar{\Phi} = [\Phi^T, I_r^T] \in \mathbb{R}^{107}$ and also friction parameters $\bar{f} = [f_v^T, f_c^T] \in \mathbb{R}^{14}$:

$$\tau = \Gamma_{\bar{\Phi}}(q, \dot{q}, \ddot{q}) \cdot \bar{\Phi} + \Gamma_{\bar{f}}(\dot{q}) \cdot \bar{f}, \quad (3.6.59)$$

where

$$\Gamma_{\bar{\Phi}} = [J^T \Gamma_\theta, \text{diag}(\ddot{q})] \in \mathbb{R}^{7 \times 107}, \quad \text{and}$$

$$\Gamma_{\bar{f}} = [\text{diag}(\text{sign}(\dot{q})), \text{diag}(\dot{q})] \in \mathbb{R}^{7 \times 14}.$$

Data was gathered at a sampling rate of 100Hz from 47 sets of random point-to-point movements, each lasting 1.3 seconds. After postprocessing the raw kinematic data samples q, \dot{q} with a fifth-order Savitzky-Golay filter of window size 35, acceleration measurements \ddot{q} were obtained by numerical differentiation of the filtered \dot{q} . The first 4,100 samples were used for identification paired with raw motor torque measurements τ . The next 2,000 samples were used for the validation with filtered motor torques using a fifth-order Savitzky-Golay filter of window size 35.

Concatenating the regressor matrices $\Gamma_{\bar{\Phi}}$ and $\Gamma_{\bar{f}}$ and motor torques τ for $n_s = 4,100$ time samples, the regularized identification formulation is given by

$$\min_{\bar{\Phi}, \bar{f}} \|A_{\bar{\Phi}}\bar{\Phi} + A_{\bar{f}} \cdot \bar{f} - b\|^2 + \gamma \cdot d(\bar{\Phi}, {}^0\bar{\Phi})^2 \quad (3.6.60)$$

$$\text{s.t. } P(\phi_i) \succ 0, \quad (3.6.61)$$

$$\text{tr}(P(\phi_i)Q_i) \geq 0, \forall i \in \{1, \dots, 10\}, \quad (3.6.62)$$

$$I_r \geq 0, \quad (3.6.63)$$

$$\bar{f} \geq 0, \quad (3.6.64)$$

where $A_{\bar{\Phi}} = [\Gamma_{\bar{\Phi}}(t_1)^T, \dots, \Gamma_{\bar{\Phi}}(t_{n_s})^T]^T \in \mathbb{R}^{7n_s \times 107}$, $A_{\bar{f}} = [\Gamma_{\bar{f}}(t_1)^T, \dots, \Gamma_{\bar{f}}(t_{n_s})^T]^T \in \mathbb{R}^{7n_s \times 14}$ and $b = [\tau(t_1)^T, \dots, \tau(t_{n_s})^T]^T \in \mathbb{R}^{7n_s}$. CAD data was used to set the prior inertial parameter values ${}^0\bar{\Phi} = [{}^0\Phi^T, {}^0I_r]$ for both bodies and rotors as well as the bounding-ellipsoid parameters Q_i . The following choices of distance metrics $d(\bar{\Phi}, {}^0\bar{\Phi})^2$ are considered for comparison:

$$\left\{ \begin{array}{ll} \|\bar{\Phi} - {}^0\bar{\Phi}\|^2 & \text{(Euclidean)} \\ d_0(\bar{\Phi}, {}^0\bar{\Phi})^2 + \sum_{i=1}^7 \left(\frac{I_{r_i} - {}^0I_{r_i}}{{}^0I_{r_i}} \right)^2 & \text{(Constant Pullback)} \\ d_{\mathcal{M}^{10}}(\bar{\Phi} \| {}^0\bar{\Phi})^2 + \sum_{i=1}^7 d_F(I_{r_i} \| {}^0I_{r_i}) & \text{(Entropic)} \\ d_{\mathcal{M}^{10}}(\bar{\Phi}, {}^0\bar{\Phi})^2 + \sum_{i=1}^7 \frac{1}{2} \log \left(\frac{I_{r_i}}{{}^0I_{r_i}} \right)^2 & \text{(Riemannian)} \end{array} \right.$$

Remark 3.13. While the inertial parameters ϕ_i of the bodies reside in $\mathcal{P}(4) \simeq \mathcal{M}$, the rotor inertias about each motor axis I_{r_i} can be treated as elements of $\mathcal{P}(1)$ (the set of positive scalars), for which the affine-invariant Riemannian distance and its approximations can be obtained as above.

The point-mass parametrization method [38] is also considered as

$$\min_{\bar{\mu}, \bar{f}} \|A_{\bar{\Phi}} R \bar{\mu} + A_{\bar{f}} \cdot \bar{f} - b\|^2 + \gamma \cdot \|\bar{\mu} - {}^0\bar{\mu}\|^2 \quad (3.6.65)$$

$$\text{s.t. } \bar{\mu} \geq 0 \quad (3.6.66)$$

$$\bar{f} \geq 0, \quad (3.6.67)$$

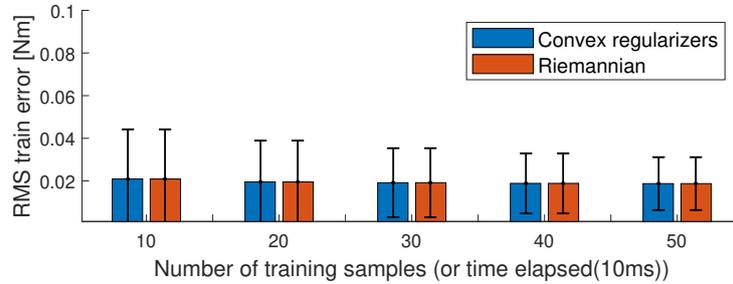
where the inertial parameters $\bar{\Phi} = R \bar{\mu}$ are reparametrized with $N_\rho = 653$ point masses $\bar{\mu} \in \mathbb{R}^{N_\rho}$ uniformly distributed across the bodies and rotors.

The convex optimization problems (3.6.60) and (3.6.65) were solved using MOSEK [40] and CVX [41] in MATLAB. The non-convex formulation of (3.6.60) with Riemannian distance regularization was solved using the interior-point method in MATLAB with the built-in function `fmincon`. Analytic gradients of the objective function

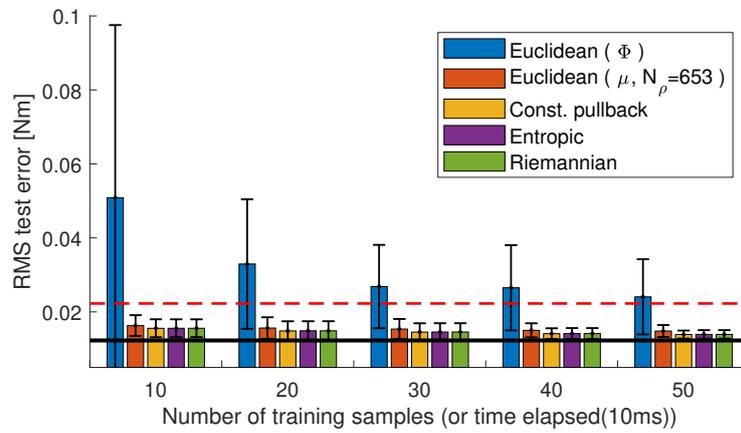
were provided. The identification with the entire 4,100 training samples took 1.0-1.2s, 0.9-1.0s, 10-11s, and 18-19s for Euclidean, constant pullback, entropic regularizers and point mass regularization respectively on an Intel Core i7-6700 desktop computer. As the training data samples are sufficiently rich, the test error discrepancies across different convex regularizers were minor; the solutions converged to provide the RMS joint torque errors on the 2,000 test samples to be around 0.0122 ± 0.001 Nm with minimal regard to the regularization effect.

The performance and computation time for the Riemannian distance regularized formulation depended largely on the initial guess and conditions provided to the solver. When the friction parameters were set to be fixed to the nominal value and the initial guess for $\bar{\Phi}$ was set to the prior value ${}^0\bar{\Phi}$, the solution converged to a local minimum having RMS test error of 0.0156 Nm (slightly larger than the others). However, when the friction parameters were optimized together with the inertial parameters, the solver converged to a bad local minimum having a large RMS test error of 0.6 Nm. When the initial guess was set to the optimal solution obtained from the entropic or the constant pullback regularization, the `fmincon` solver took 7-10s to converge to a very nearby local minimum solution having almost the same test error. Convex approximation methods thereby not only improve computational efficiency and stability of the geometric inertial parameter identification but also increase the flexibility to more easily handle additional model parameters like friction.

Figure 3.7 shows the plots of identification results when the training samples were provided only for a short duration of time. Performance is measured in terms of both mean and standard deviation of test errors on 30 randomly sampled sets of training data. The residual errors for the training samples maintain the same values across different regularization types, which was realized by the method provided in Sec. 3.5.1 (the regularization factor was set as $\gamma = 0.03$ for the entropic divergence regularizer). Under

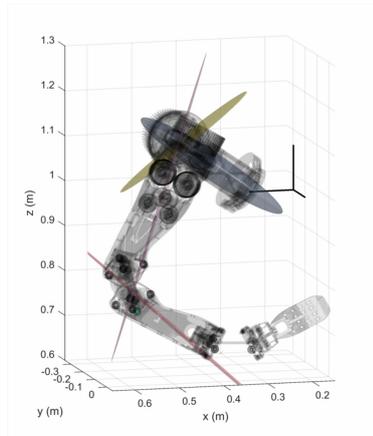
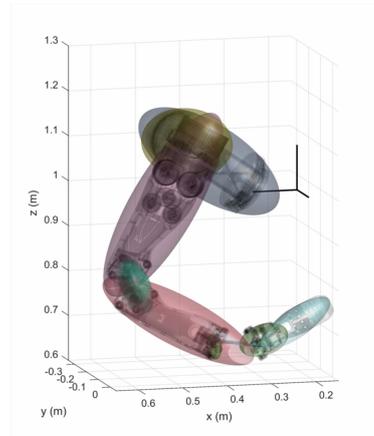
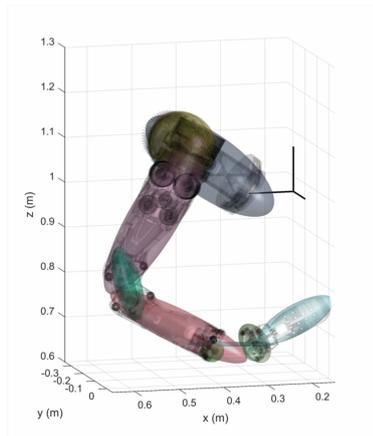


(a) RMS motor torque error on the training samples

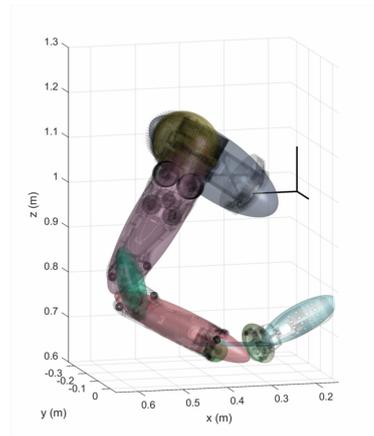


(b) RMS motor torque error on the test samples for validation

Figure 3.7: Inertial parameter identification results for the AMBIDEX robot manipulator with different regularizers and varying the number of training samples. Friction parameters are fixed to the values identified from the full sample set. The red dashed horizontal line indicates the RMS test error with the prior value ${}^0\bar{\Phi}$ and the black bold horizontal line indicates the level of RMS test error using full set of 4,100 training samples for identification.

(a) Euclidean (Φ)(b) Euclidean (μ)

(c) Constant pullback



(d) Entropic

Figure 3.8: Visualization of the identified inertial parameters of the AMBIDEX manipulator with 30 training samples using different distance measures for regularization. The ellipsoids represent uniform mass densities uniquely realizable from the inertial parameters ϕ_i of each link. The ellipsoids for point mass parameters $\bar{\mu}$ are obtained by the transformation rule $\bar{\Phi} = R\bar{\mu}$.

identical training errors, geometric regularization on $\bar{\Phi}$ and Euclidean point-mass regularization considerably surpass the generalization performance of the Euclidean regularizer on $\bar{\Phi}$. These regularizers, using only a small number of samples, effectively capture the distributional information of the prior inertial parameters and generalize comparably to the identification results using the full sample set. In contrast, observe that the Euclidean regularizer on the inertial parameters shows larger test error than the test error evaluated with using prior inertial parameters. It also converges slowly. Figure 3.8 shows the identified inertial parameters using 30 samples obtained during a 0.3s interval. It can be clearly observed that the naive Euclidean regularization on $\bar{\Phi}$ does not make use of the distributional information of the prior inertial parameters and shows physically unrealistic results.

This particular experiment sheds further light on the potential use of distribution-aware regularization methods for online identification [42] or direct adaptive control [43, 44] of robots, wherein generalizable estimation with limited information is desirable for fast convergence. Considering both the computational cost of online optimization relative to the resultant performance, quadratic regularization on the compact representation $\bar{\Phi}$ with the constant pullback metric may be practically appealing over the point-mass parametrization.

Remark 3.14. *Least squares with a quadratic regularizer can be efficiently implemented with online recursive algorithms. However, these recursive least squares methods do not in general guarantee a system's stability [45]. Our recent results [43] (Chapter 5) show that the entropic divergence can be used to develop efficient geometric direct adaptation laws while also guaranteeing the stability of the system. Nevertheless, as direct methods are generally known to suffer from parameter drift problem [45], it would still be interesting to investigate geometric approaches in, e.g., composite methods [46].*

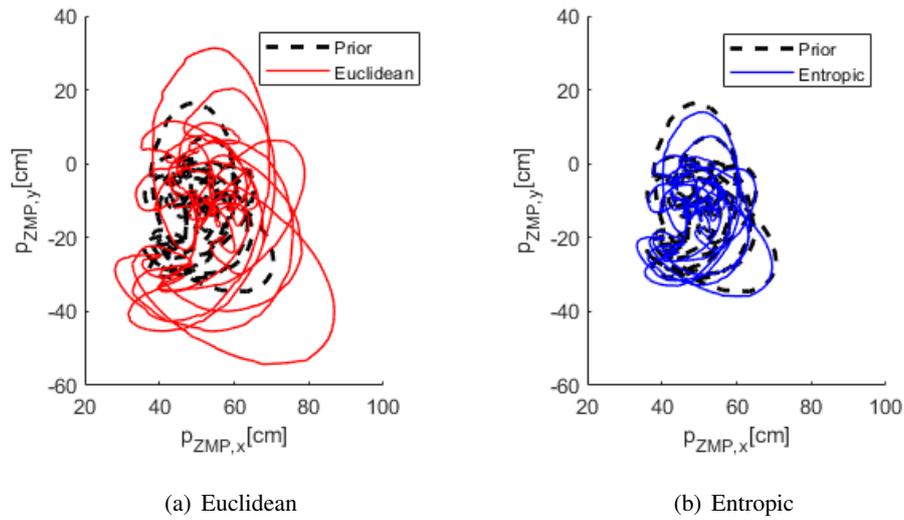


Figure 3.9: Comparison of the ground ZMP prediction trajectories with the identified inertial parameters and prior value.

Lastly, Figure 3.9 shows the plot of the predicted Zero Moment Point(ZMP) [47] on the ground. The predicted ZMP trajectory of the Euclidean deviates more than 5 cm in average from the ones of both prior and entropic. Although the AMBIDEX manipulator is currently fixed to the ground and thereby does not need an extra information, like ZMP, for active balancing control, this result foreshadows the potential of using geometric identification in floating-base systems, which shall be discussed in the following Section.

3.6.2 MIT Cheetah 3 Robot

Extensive identification experiments were conducted on a leg of the MIT Cheetah 3 robot [48]. In particular, we examine the generalizability of the fixed-base identification to both fixed-base and floating-base motion trajectories.

3.6.2.1 Fixed-base Identification and Validation

Each 3-DoF leg of the MIT Cheetah 3 robot is driven by three proprioceptive actuators [49], each including a high-inertia rotor coupled to the joint by a 10.6 : 1 gearbox. To address actuator effects, the leg was treated as a system of 6 bodies (3 links and 3 rotors). To account for transmission losses, viscous and Coulomb friction coefficients $f_v \in \mathbb{R}^3$ and $f_c \in \mathbb{R}^3$ were included in the joint space dynamics as [28],

$$\begin{aligned}\tau &= \Gamma(q, \dot{q}, \ddot{q})\Phi + \text{diag}(\dot{q})f_v + \text{diag}(\text{sign}(\dot{q}))f_c \\ &= \Gamma(q, \dot{q}, \ddot{q})\Phi + \Gamma_{\bar{f}}\bar{f},\end{aligned}$$

where $\Phi = [\phi_1^T, \dots, \phi_6^T] \in \mathbb{R}^{60}$ is the vector of inertial parameters of the 6 bodies, $\bar{f} = [f_v^T, f_c^T]^T \in \mathbb{R}^6$ is the concatenated vector of the friction parameters, and $\Gamma_{\bar{f}} = [\text{diag}(\dot{q}), \text{diag}(\text{sign}(\dot{q}))] \in \mathbb{R}^{3 \times 6}$. Data was gathered from a leg swinging experiment conducted in [28] with a sampling rate of 1kHz. The samples obtained during the first 10s of motion were used for identification and the ones from the next 10s of motion were used for the validation. The raw kinematic data samples were postprocessed with a 4-th order Savitzky-Golay filter of window size 101. Moreover, in account of the large error in the Coloumb friction model observed during slow motion, the data samples having joint velocities less than 5deg/s were excluded for both identification and validation. We also note that we use the raw torque measurements for identification, while a 4-th order Savitzky-Golay filter of window size 101 is used to filter out the high-frequency noise in the torque measurements only for the validation (test) set.

Concatenating the regressor matrices and motor torques τ for $n_s = 7,379$ training

samples, the regularized identification formulation can be given by

$$\min_{\Phi, \bar{f}} \|A\Phi + A_{\bar{f}}\bar{f} - b\|^2 + \gamma \cdot d(\Phi, {}^0\Phi)^2 \quad (3.6.68)$$

$$\text{s.t. } P(\phi_i) \succ 0, \quad (3.6.69)$$

$$\text{tr}(P(\phi_i)Q_i) \geq 0, \forall i \in \{1, \dots, 6\}, \quad (3.6.70)$$

$$\bar{f} \geq 0, \quad (3.6.71)$$

where $A = [\Gamma(t_1)^T, \dots, \Gamma(t_{n_s})^T]^T \in \mathbb{R}^{3n_s \times 60}$, $A_{\bar{f}} = [\Gamma_{\bar{f}}(t_1)^T, \dots, \Gamma_{\bar{f}}(t_{n_s})^T]^T \in \mathbb{R}^{3n_s \times 6}$ and $b = [\tau(t_1)^T, \dots, \tau(t_{n_s})^T]^T \in \mathbb{R}^{3n_s}$. The prior inertial parameter value ${}^0\Phi$ and the bounding-ellipsoid parameters Q_i were set using CAD data. Convex optimization problems (3.6.68) with Euclidean, constant pullback, and entropic regularizations were solved as well as the point mass regularization formulation,

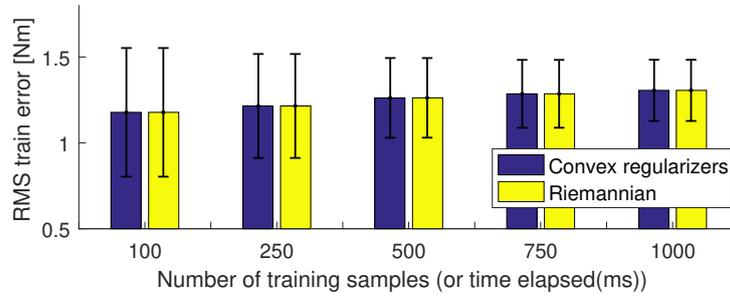
$$\min_{\mu, \bar{f}} \|AR\mu + A_{\bar{f}} \cdot \bar{f} - b\|^2 + \gamma \cdot \|\mu - {}^0\mu\|^2 \quad (3.6.72)$$

$$\text{s.t. } \mu \geq 0 \quad (3.6.73)$$

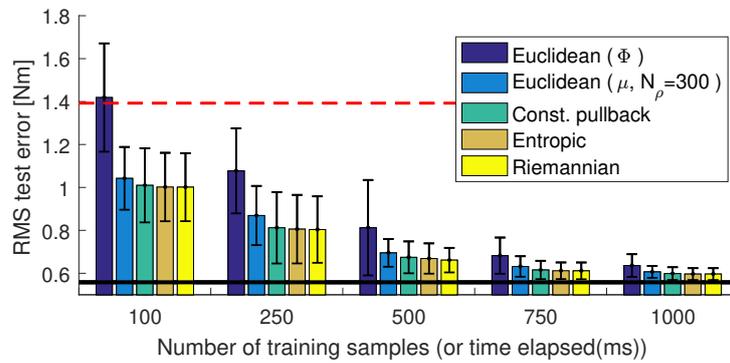
$$\bar{f} \geq 0, \quad (3.6.74)$$

with $\mu \in \mathbb{R}^{N_\rho}$ for $N_\rho = 300$ point masses. All problems were solved using MOSEK [40] and CVX [41] in MATLAB. The identification with the entire 7,379 training samples took 0.7-0.8s, 0.6-0.7s, 4-5s, and 3-4s for the Euclidean, constant pullback, entropic, and point mass regularizers respectively on an Intel Core i7-6700 desktop computer. Riemannian distance regularization was again solved using the `fmincon` interior-point solver with an initial guess from the solution of the entropic regularization.

Figure 3.10 shows plots of the identification results while varying the number of training samples. Performance is measured in terms of both the mean and standard deviation of test errors on 35 randomly sampled sets of training samples. The residual errors for the training samples maintain the same values across different regularization



(a) RMS joint torque error on the training samples



(b) RMS joint torque error on the test samples

Figure 3.10: Identification results on the MIT Cheetah 3 leg with different regularizers and varying the number of training samples. Friction parameters are fixed to the values identified from the full samples. The red dashed horizontal line indicates the RMS test error with the prior value ${}^0\Phi$ while the black bold horizontal line indicates the level of RMS test error using the full 7,379 training samples for identification.

types, which was again realized by the method provided in Sec. 3.5.1 (the regularization factor is explicitly set as $\gamma = 0.2$ only for the entropic regularizer). As can be seen throughout Figure 3.10, and also as shown previously in the experiments on the AMBIDEX, for small training samples, distribution-aware regularizers enjoy better generalization errors over Euclidean regularization on Φ . As can be expected based on Figure 3.10, when using all 7,379 training data samples, the test error discrepancies across different regularization types are minor. However, this observation does not imply that the parameters converge to the same values.

Figure 3.11 shows the visualization of the identified parameters when using the full set of data samples. It can be observed that regularization with the Euclidean metric on Φ does not converge to a physically meaningful set of values, although the parameters predict joint torque values comparably to other results. Such a result can be understood from the fact that only reduced linear combinations of the parameters, so-called base parameters, affect the joint space dynamics. In the following section, we consider how these errors in the identified parameters may affect the accuracy of the full floating-base dynamics of the Cheetah.

3.6.2.2 Fixed-Base Identification with Floating-Base Validation

Due to safety issues and poor excitability, the identification of inertial parameters for legged-robots with floating-base motions is restricted in practical situations. Instead, isolated fixed-base identification is commonly conducted one leg at a time with the body fixed in place. Although fixing the base makes these approaches viable, the associated downside is that the entire parameter set is not maximally excitable [50]. In this section, we conduct an identification study with simulation data to demonstrate the effects of regularization in fixed-base identification on the accuracy of the full floating-base model.

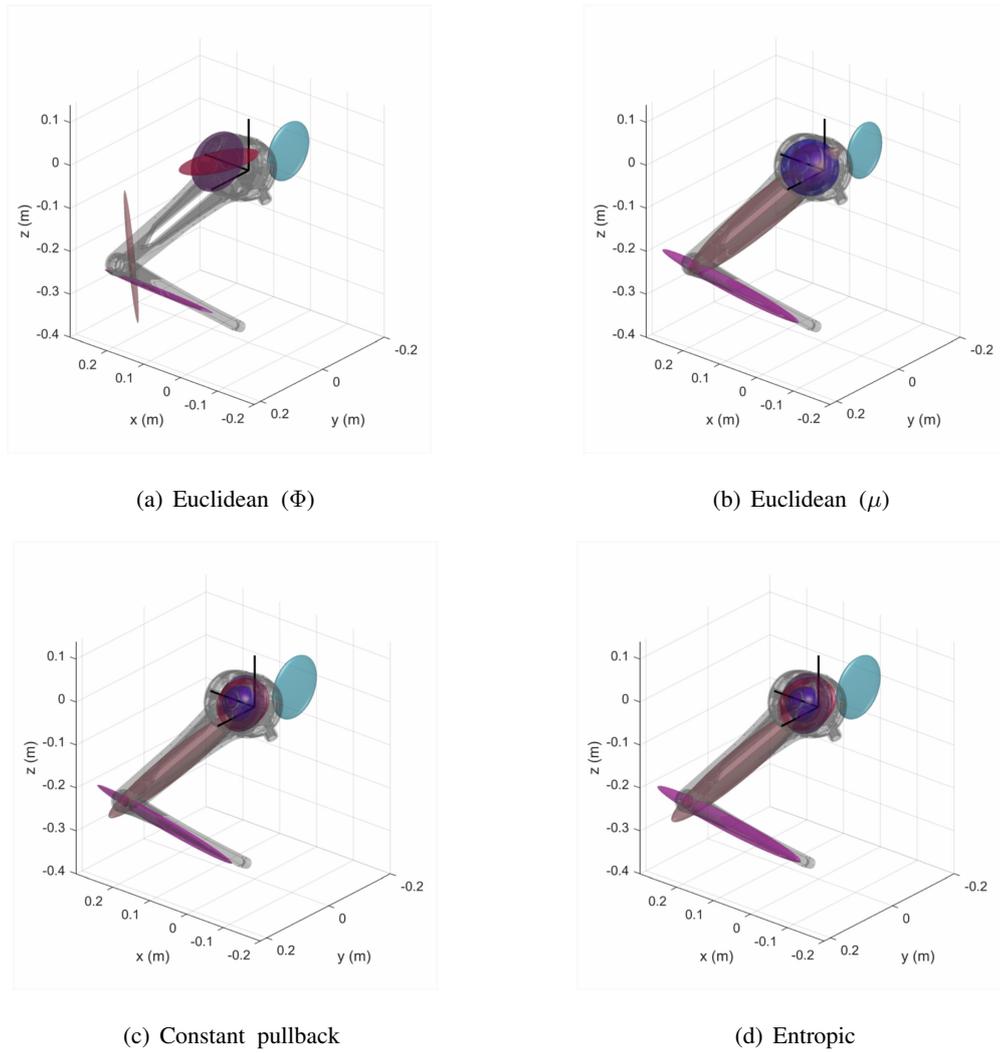


Figure 3.11: Visualization of the identified inertial parameters of the MIT Cheetah 3 leg with 30 training samples using different distance measures for regularization. The ellipsoids represent uniform mass densities uniquely realizable from the inertial parameters ϕ_i of each link. The ellipsoids for the point mass parameters μ are obtained by the transformation rule $\Phi = R\mu$.

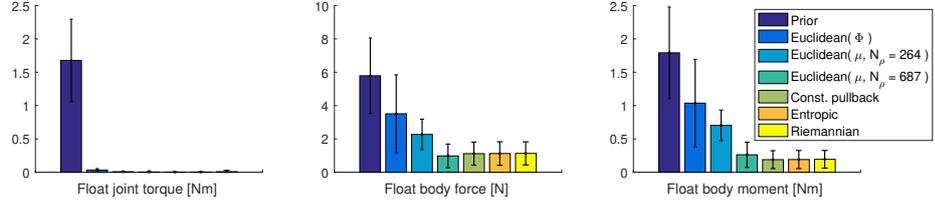


Figure 3.12: Identification results on the MIT Cheetah 3 leg with Euclidean, constant pullback, and entropic distance based regularizers validated on floating-base galloping motion trajectory. According to the rule (B.2.11), the noise condition $\delta=0.5$, $N_m=10$ is used for perturbing the ground truth inertial parameters from the fixed prior value ${}^0\Phi$. The mean and variance of RMS prediction errors on the joint torque, body force, and body moment are presented.

To generate data, we consider the Cheetah 3 executing a dynamic transverse gallop in simulation [51]. Ground-truth data for the galloping motion includes the configuration $q_{\text{full}} \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \times SE(3)$ of four legs and a floating base, generalized velocity v , generalized acceleration \dot{v} , and generalized force during this galloping motion. The generalized force includes effects from both active joint torques as well as effects from ground reaction forces. Collecting this dataset would be impractical in experiments, as it would require force plates at precise locations during galloping.

A fixed-base identification of the front-left leg was considered by mimicking the table-top setup in the previous section. Samples of the front-left (FL) leg configuration $q_{FL} \in \mathbb{R}^3$, velocity \dot{q}_{FL} , and acceleration \ddot{q}_{FL} from the galloping dataset were used with ground truth inertial parameters Φ_{GT} to obtain the required joint torques, i.e., $\tau_{FL} = \Gamma(q_{FL}, \dot{q}_{FL}, \ddot{q}_{FL}) \cdot \Phi_{GT} \in \mathbb{R}^{60}$. This synthetic data was used to identify the leg inertial parameters $\Phi \in \mathbb{R}^{60}$ (3 bodies and 3 rotors) by solving the following problem similar

to the previous section:

$$\min_{\Phi} \|A\Phi - b\|^2 + \gamma \cdot d(\Phi, {}^0\Phi)^2 \quad (3.6.75)$$

$$\text{s.t. } P(\phi_i) \succ 0, \quad (3.6.76)$$

$$\text{tr}(P(\phi_i)Q_i) \geq 0, \forall i \in \{1, \dots, 6\}. \quad (3.6.77)$$

We also considered the point mass parametrization formulation:

$$\min_{\mu} \|AR\mu - b\|^2 + \gamma \cdot \|\mu - {}^0\mu\|^2 \quad (3.6.78)$$

$$\text{s.t. } \mu \geq 0. \quad (3.6.79)$$

In this experiment, we additionally considered the effects of varying the number of point mass parameters ($N_\rho = 264$ and $N_\rho = 687$).

It should be emphasized that the synthetic data in this example is noise-free; therefore, a near-zero regularization factor is chosen as $\gamma = 10^{-11} \cdot \text{tr}(A^T A)$ (for the Euclidean regularizer) for accurate identification of the fixed-base dynamic model. However, as mentioned earlier, we focus on the generalizability of the regularized solutions to the floating-base model, which is affected by a larger number of parameters.

To generate the ground truth inertial parameters of the simulated model, we adopt a physically sensible approach that samples different values of Φ_{GT} : for each rigid body, we deliberately perturb the fixed prior value ${}^0\Phi$ by adding or subtracting N_m point masses at uniformly distributed positions inside a prescribed bounded region (See Appendix B.2 for details of the implementation). This intuitive sampling strategy reflects realistic situations where extra parts such as cables or mechanical units are added to the robot without full modelling.

Figure 3.12 shows plots of joint torques errors and body force/moment prediction errors during the floating-base galloping motion. The geometric regularizations are shown

to generalize significantly better compared to the Euclidean regularization on Φ . Moreover, it is interesting to observe that using a larger number of point mass parameters leads to better generalization ability, although in practice, the achievable performance by increasing the number of point mass parameters may be largely limited by the measurement conditions (as observed in the previous experiments).

Remark 3.15. *For the Cheetah robot, the external force prediction errors shown above may not be as significant within whole-body control, as a large portion of the total mass is concentrated in the trunk [51] and the legs are relatively lighter. Still, these results demonstrate the potential of geometric identification methods for various types of floating-base structures, c.f., [52], [53], where heavier limbs will increase the importance of proper regularization.*

3.6.3 Human with Low-Cost Affordable Sensors

In this section, regularized identification methods are compared for a human subject with data recorded using publicly affordable sensor equipment. A low-cost force plate (Wii balance board) was used to measure the vertical ground reactive force $f_z \in \mathbb{R}$ and position of the Center of Pressure (CoP), or equivalently the Zero Moment Point (ZMP) $(x_{\text{ZMP}}, y_{\text{ZMP}}) \in \mathbb{R}^2$. A Kinect depth camera was used to obtain point cloud data samples of human motions. The poses of 16 links of a human were tracked by fitting a SCAPE articulated parametric human mesh model to the point cloud streams. This data was then filtered, with spatial velocities and accelerations obtained by a finite difference method on $SE(3)$. The fitted mesh model was also used to obtain the prior inertial parameter values ${}^0\Phi$ by assuming a constant mass density and known total mass. An outer-bounding ellipsoid for each link was also derived from the mesh geometry.

Whole-body momentum dynamics were used to generate a system of linear equations for the human inertial parameters. Each motion lasted 10 seconds, with 300 samples used for identification. Ten sets of 100 samples from short motions (3 seconds each) excited different parts of the body and are used for validation purposes. With this data, a convex regularized inertial parameter identification problem for a human subject is formulated as

$$\min_{\Phi} \|A\Phi - b\|_{\Sigma^{-1}}^2 + \gamma \cdot d(\Phi \| {}^0\Phi) \quad (3.6.80)$$

$$\text{s.t. } P(\phi_i) \succ 0, \quad (3.6.81)$$

$$\text{tr}(P(\phi_i)Q_i) \geq 0, \forall i \in \{1, \dots, 16\}, \quad (3.6.82)$$

$$\sum_{i=1}^{16} m_i = M, \quad (3.6.83)$$

where the last linear equality constraint (3.6.83) assures that the total mass is a constant known value M . The corresponding point mass parametrization formulation can be realized as

$$\min_{\mu} \|AR\mu - b\|_{\Sigma^{-1}}^2 + \gamma \cdot \|\mu - {}^0\mu\|^2 \quad (3.6.84)$$

$$\text{s.t. } \mu \geq 0, \quad (3.6.85)$$

$$\mu^T \cdot \mathbf{1}_{N_\rho} = M, \quad (3.6.86)$$

where $\mathbf{1}_{N_\rho} \in \mathbb{R}^{N_\rho}$ is a constant vector of ones. We use $N_\rho = 1351$ points uniformly distributed over the mesh geometry.

Remark 3.16. *Note that if not for the convex formulation of the problem, even including a simple linear equality constraint (3.6.83) would be less straightforward and more computationally intensive when conducting nonlinear optimization on a manifold. For example, the MATLAB `fmincon` interior-point solver fails to converge to a feasible point*

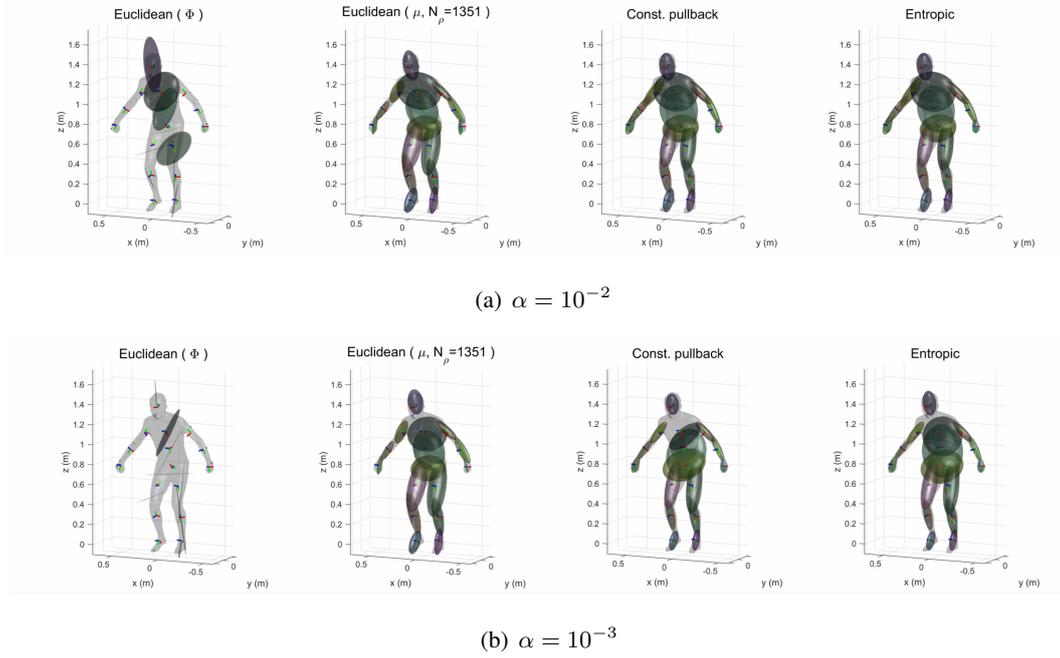


Figure 3.13: Human inertial parameter identification results while varying the regularization factor $\gamma = \alpha \cdot \text{tr}(A^T A)$. The ellipsoids represent uniform mass densities uniquely realizable from the inertial parameters ϕ_i of each link. The ellipsoids for point mass parameters μ are obtained by the transformation rule $\Phi = R\mu$.

that satisfies the equality constraint with the Riemannian distance regularizer. Optimization solvers originally developed for vector valued variables, like fmincon, are not fully equipped for optimization on manifolds [54].

The data in this case is inherently noisy due to a low data sampling frequency of 20-30 Hz combined with the poor precision of the low-cost sensor. In this case, the baseline regularization factor is set to $\gamma = \alpha \cdot \text{tr}(A^T A)$, $\alpha = 10^{-2}$ for the Euclidean regularizer. Again, the regularization factors for other regularizers are implicitly determined according to the residual error loss of the Euclidean-regularized solution, using

Table 3.1: Identification Results on a Human Subject : RMS Error of Predicted f_z [N], x_{ZMP} [cm], y_{ZMP} [cm]

	Train			Test		
	f_z	x_{ZMP}	y_{ZMP}	f_z	x_{ZMP}	y_{ZMP}
Prior	13.79	2.09	1.58	13.32	2.33	1.01
Euclidean(Φ)	12.89	1.61	0.82	13.09	1.87	0.95
Euclidean(μ)	12.47	1.64	0.81	13.19	1.83	0.76
Const. pullback	12.52	1.66	0.81	13.28	1.87	0.78
Entropic	12.51	1.66	0.81	13.28	1.86	0.78

the method explained in Sec. 3.5.1.

Table 3.1 shows the ZMP and vertical ground reactive force f_z prediction errors on the validation motion sets. Differences in the prediction errors for the ZMP position and ground reaction force f_z are difficult to discern across different regularization types. However, as shown in Figure 3.13(a), when comparing the actual identified parameters visualized with equivalent uniform ellipsoids, the ones obtained from the Euclidean regularizer are mostly physically implausible. These results imply that while the reduced linear combinations of identifiable parameters from the f_z and ZMP measurements may be identified properly, the full set of parameters is not. The precision of the inertial parameters of each link may be less important in terms of the net momentum dynamics, yet they can be important when considering local features like joint forces/torques acting at, e.g., the wrist, shoulder, spine, pelvis, etc, which can be of special interest in biomechanics analysis.

Remark 3.17. *Similar to the table-top identification of a quadruped leg from joint torque measurements, identification with partial data or with low-cost sensors requires appropriate regularization for results to generalize to an accurate full model. The incorporation of a geometric regularizer is effective for making this process less sensitive to the types and sources of observations.*

Lastly, a qualitative assessment of the identification results is more viable with distribution-aware regularizations. As shown in Figure 3.13, a reasonable range for the regularization factor α can be deduced to be greater than 10^{-3} . The range of $\alpha < 10^{-3}$ would imply overfitting to the sensor noise or possibly unmodelled non-rigid effects, e.g., muscle contractions, fluid flow, etc, which are known to be non-trivial for accurate modeling of human body dynamics [55].

3.7 Discussion

Through a wide range of identification experiments with various conditions on the sample sufficiency and quality, we have demonstrated that geometric identification methods lead to much more generalizable estimation of the inertial parameters of multibody systems. In fact, in many cases of multibody systems identification, even the structurally identifiable set of parameters may not be effectively identifiable from data, and the estimation can easily become highly sensitive depending on the measurement conditions. Nevertheless, it is worth emphasizing again that inertial parameters are physical quantities which can in many cases be provided with a reasonable prior estimate, e.g., from CAD data or, if not available, even through a rough guess. The proposed geometric approach allows one to incorporate the prior information from a distributional perspective

on the inertial parameters, and effectively regularize the ill-posed inertial parameter estimation problems in a computationally efficient manner. As was observed in the preceding experiments, naive non-physical regularization can easily lead to large information loss in the estimation.

With this view, although redundant point mass parameters are auxiliary for estimating the inertial parameters, they do provide an alternate powerful way to incorporate the prior information. In fact, there may be situations where identifying the mass distribution itself is important, e.g., for biomechanics application. These point mass parametrization techniques can potentially provide more information as a prior, especially for bodies with highly irregular mass densities or with complicated shapes. In this regard, it will be interesting to further investigate the general class of divergence metrics on the space of positive measures, to be used within regularized identification; for instance, the generalized KL divergence [35] between the discretized mass densities $\mu = [\mu_1, \dots, \mu_n] \in \mathbb{R}_+^{N_\rho}$ and $q = [\nu_1, \dots, \nu_n] \in \mathbb{R}_+^{N_\rho}$ is defined by

$$D_{\text{GenKL}}(\mu||\nu) \triangleq \sum_{i=1}^{N_\rho} \mu_i \log \left(\frac{\mu_i}{\nu_i} \right) + \nu_i - \mu_i \geq 0, \quad (3.7.87)$$

which is, like the Euclidean distance $\|\mu - \nu\|^2$, also a convex function with respect to both μ and ν .

Remark 3.18. *We also note that for bodies with complicated shapes and densities, rather than relying entirely on the point-mass discretization, the choice of more compact discretized mass primitives, e.g., the combination of multiple mass ellipsoids for each body, could be more effective in terms of computational cost while not sacrificing too much on the representation capability for the prior.*

Overall, this metric may provide an opportunity to unify point-mass and geometric

regularization strategies. That being said, the current paper has shown the power of geometric regularizers to capture the main benefits of point-mass formulations while requiring only a prior for the parameters, and without fixing the spatial distribution of mass. These findings support the central claim that geometric regularizers implicitly enable regularization toward a prior mass distribution, and not simply toward prior parameter values.

3.8 Conclusion

Our main contribution lies in reformulating the classical linear least squares problem of inertial parameter identification to a nonlinear one, by exploiting the Riemannian geometry of the manifold of physically consistent inertial parameters. The robustness of our approach is traceable in large to the use of the natural distance metric in lieu of the standard Euclidean metric. The proposed natural distance metric more effectively captures the distance between the underlying mass distributions of two rigid bodies.

We have also presented a convex programming approach to geometric inertial parameter identification. The entropic divergence and constant pullback distance provide a convex, second-order approximation to the exact Riemannian distance while retaining the desired coordinate invariance properties. Further, its structure enables the use of semi-definite optimization for coordinate-free regularization. Extensive hardware and simulation experiments, supported by a provable comparative analysis scheme, demonstrate that our approach enjoys improved generalizability and computational efficiency over existing Euclidean counterparts.

While dynamic model-based control methods in robotics are becoming mainstream for ever more complex systems like humanoids and multi-legged robots, the sensor measurements for full precise identification are often impractical in terms of data sufficiency,

noise, and parameter excitability. We would argue that our contribution is made more meaningful to the practitioner by offering methods that work robustly even with the limited source and quality of measurements.

4

Geometric Criteria for Excitation Trajectory Optimization

4.1 Introduction

Alongside with the vast literature on dynamic identification [15, 20, 21, 22, 23], the problem of generating optimal excitation trajectories has been studied extensively [56, 24]. It is generally stated as a trajectory optimization of some optimality criterion so as to reduce the estimator variance or well condition the regressor matrix subject to various physical constraints, e.g., joint limits, torque limits, and whole-body balancing constraints for floating-base systems like humanoids and multi-legged robots [57], [26].

Most of the existing approaches [56, 24] have been based on using classical optimal experimental design criteria [58, 59]. A class of optimality criterion for optimal experimental design, e.g., alphabet-optimality and condition number, are defined by the symmetric functions of eigenvalues of the estimator variance or information matrix. Owing to the linear decomposition of dynamic parameters in robot dynamic equations, least

square estimator variance of the parameters can be derived in an explicit closed-form matrix. This greatly eases the evaluation of any variance-related optimality criterion and their gradients for conducting excitation trajectory optimization [24, 57].

Meanwhile, it is generally the case [58, 60, 61, 62, 63] that the pure classical optimal design criteria may still come to a poor estimation of parameters when the parameters are badly scaled; estimations of small parameters are prone to be erroneous with respect to their scale whereas the absolute value of error is often acceptable. Indeed, inertial parameters for multibody systems are typically badly scaled, not only due to the mass and size scale difference across bodies but also from the fact that they, e.g., mass, rotational inertia tensor, and even friction parameters, consist of diverse physical quantities with different units. [64] is the only work to address this issue in the context of inertial parameter identification. Authors normalize the estimator covariance matrix using the given nominal value so as to balance the respective accuracy of the parameters. Yet, the criteria can refer to intrinsically different normalizations depending on the choice of coordinate (this includes body-fixed reference frames and linear reparametrization of the observable parameters or so-called *base parameters*). This is essentially due to coordinate dependency of the normalized criterion which is only scale invariant. Perhaps the most complicating issue in defining a meaningful scale-normalization lies in the fact that there is no canonical choice of coordinates in which each of the base parameter values has own physical meaning.

In this Chapter, by exploiting the intrinsic geometric structure of the inertial parameters as introduced in Chapter 3, a coordinate-free framework in which to formulate optimal trajectory excitation criteria is presented. Our main focus will be on a criterion in which the information matrix is weighted with the Riemannian metric evaluated at the nominal value, with symmetric functions f of its eigenvalues admitting physically meaningful interpretations while preserving coordinate-invariance. The proposed

coordinate-free framework frees the designer to focus solely on the intrinsic aspects of the problem, e.g., choosing a physically meaningful Riemannian metric and symmetric function f . Moreover, structural similarities between our proposed criterion and existing criteria allows for the use of the same numerical optimization algorithms for, e.g., evaluating the gradient of the criterion [57], by performing simple linear transformations of the variables via the chain rule.

Finally, we show how the same geometric framework can be applied to infer the optimal reduced set of effectively identifiable parameters from the given trajectory samples. Such reduction of the model parameters can be useful not only for finding optimal excitation trajectories, but also parameter identification for complex high-dimensional systems like humanoid robots, in which the execution of reference trajectories that sufficiently excite the entire structurally identifiable set of parameters is difficult or impractical to achieve.

4.2 Preliminaries

In this section, we revisit various conventional optimality criteria for excitation trajectories used for mechanical systems identification. We also discuss through both examples and analysis how the choice of reference frames, base parameter set, and physical units explicitly enters into formulations of existing criteria. This discussion serves to motivate a new set of coordinate-invariant criteria to be proposed in the next section.

4.2.1 Optimal Design of Experiments

Consider m samples of observations $\{a_i, b_i\}_{i=1}^m$, $a_i \in \mathbb{R}^n$, $b_i \in \mathbb{R}$, under the linear observation model given by

$$b_i = a_i^T x + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, \sigma_i^2), \quad i = 1, \dots, m \quad (4.2.1)$$

, where $x \in \mathbb{R}^n$ is a model parameter vector and ϵ_i is a zero mean Gaussian noise. Linear equations in (4.2.1) can be expressed in a vector form as,

$$b = Ax + \epsilon, \quad \epsilon \sim \mathcal{N}(0, \Sigma) \quad (4.2.2)$$

, where $b = [b_1, \dots, b_m]^T \in \mathbb{R}^m$, $A = [a_1, \dots, a_m]^T \in \mathbb{R}^{m \times n}$, and $\Sigma = \text{diag}(\sigma_1^2, \dots, \sigma_m^2)$. Then, least squares estimation of x finds the minimum weighted residual error, or equivalently the maximum log likelihood, solution, i.e.

$$\min_x \|\epsilon\|_{\Sigma^{-1}}^2 = \|Ax - b\|_{\Sigma^{-1}}^2. \quad (4.2.3)$$

When the regressor matrix A has maximal rank, the unique closed-form solution is given by

$$x^* = (A^T \Sigma^{-1} A)^{-1} A^T \Sigma^{-1} b \quad (4.2.4)$$

Since the observation b is a random vector following the Gaussian distribution $\mathcal{N}(Ax, \Sigma)$, the least square estimator x^* (4.2.4) is also a Gaussian random vector with mean and covariance matrix explicitly given by $\mathbb{E}[x^*] = x$ and

$$\text{var}[x^*] = (A^T \Sigma^{-1} A)^{-1}. \quad (4.2.5)$$

It can be observed that x^* is always an unbiased estimator of x . Moreover, the estimator is efficient in the sense that the estimator covariance matrix achieves the Cramer-Rao lower bound, i.e.,

$$\text{var}[x^*] = M^{-1}, \quad (4.2.6)$$

where $M \in \mathbb{R}^{n \times n}$ is the Fisher information matrix given by

$$M = \mathbb{E} [\nabla_x \log p(b|x) \cdot \nabla_x \log p(b|x)^T] \quad (4.2.7)$$

$$= A^T \Sigma^{-1} A. \quad (4.2.8)$$

Observe that both the information matrix (4.2.8) and the estimator covariance matrix (4.2.5) are directly influenced by the regressor matrix A . This implies that some care could be taken to effectively sample the observation set $\{a_i\}_{i=1}^m$ in order to obtain a desired estimator variance. Sampling the observations so as to maximize the information or equivalently to minimize the estimator variance can be regarded as a problem in *optimal design of experiments* [58]. Optimality criteria are typically defined by symmetric functions $f : \mathbb{R}_+^n \rightarrow \mathbb{R}$ of eigenvalues $\lambda = (\lambda_1, \dots, \lambda_n)$ of the inverse of estimator covariance, or information matrix, i.e.,

$$J = f(\lambda(M)) = f(\lambda(\text{var}[x^*]^{-1})) = f(\lambda(A^T \Sigma^{-1} A)). \quad (4.2.9)$$

Some of the most popular widely used criteria are

- A-optimality: $f(\lambda) = \sum_{i=1}^n 1/\lambda_i = \text{tr}((A^T \Sigma^{-1} A)^{-1})$
- D-optimality: $f(\lambda) = -\sum_{i=1}^n \log \lambda_i = -\log |A^T \Sigma^{-1} A|$
- E-optimality: $f(\lambda) = \lambda_{\max}((A^T \Sigma^{-1} A)^{-1})$
- Condition number : $f(\lambda) = \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}} = \text{cond}(\Sigma^{-1/2} A)$

Minimizing the above criteria typically correspond to estimates that minimize variance or maximize information. The exception is the condition number, which attempts to balance uncertainties across the estimation parameters, i.e., $\text{var}[x^*] \simeq \alpha \mathbb{I}$ for some positive scalar α (here \mathbb{I} denotes the identity matrix).

Except for D-optimality, traditional optimality criteria that are concerned with the variance of the estimation parameters, e.g., A-, E-, T-optimality and condition number, are generally not coordinate invariant; that is, the landscape of the criterion over different sets of samples changes by arbitrary linear reparametrizations of the parameters, i.e., $x \mapsto Sx$ with nonsingular square matrix S [58].

4.2.2 Excitation Criteria for Multibody Systems

Recall that inertial parameter identification problem for multibody systems always reduces to a parameter estimation problem on linear observation models; following the notations in (3.2.13), consider the linear regression model

$$\tau(t_i) = Y(q(t_i), \dot{q}(t_i), \ddot{q}(t_i))\Phi + \epsilon, \epsilon \sim \mathcal{N}(0, \Sigma_\epsilon), \quad (4.2.10)$$

where the observation τ is subject to additive Gaussian noise with covariance $\Sigma_\epsilon \in \mathbb{R}^{n \times n}$. We assume that observations are taken at m time instances t_1, \dots, t_m along some reference trajectory $q(t)$. Then, the least-squares estimation problem for Φ is formulated as

$$\min_{\Phi} \|A\Phi - b\|_{\Sigma}^2. \quad (4.2.11)$$

where $\Sigma = \text{diag}(\Sigma_\epsilon, \dots, \Sigma_\epsilon) \in \mathbb{R}^{mn \times mn}$ is a block diagonal matrix of Σ_ϵ , and

$$A = [Y(t_1)^T, \dots, Y(t_m)^T]^T \in \mathbb{R}^{nm \times 10n}, \quad (4.2.12)$$

$$b = [\tau(t_1)^T, \dots, \tau(t_m)^T]^T \in \mathbb{R}^{nm}. \quad (4.2.13)$$

Meanwhile, one of the unique properties of inertial parameter identification problem of coupled rigid body systems is that the concatenated regressor matrix A is inherently rank deficient. This implies that, in principle, the reduced set or some linear combinations of inertial parameters are only identifiable purely from the observation data (without incorporating prior information through, e.g., regularization); this restriction is in fact due to the kinematic structure of the mechanism that restricts admissible spatial motion for each body in space [50]. Specifically, there exists some constant fat matrix $B \in \mathbb{R}^{n_B \times 10n}$ ($n_B < 10n$) such that the regressor matrix $Y(q, \dot{q}, \ddot{q})$ is explicitly decomposable as

$$Y(q, \dot{q}, \ddot{q}) = Y_B(q, \dot{q}, \ddot{q})B \quad (4.2.14)$$

for all q, \dot{q}, \ddot{q} . Hence, reduced linear combinations of Φ given as

$$\Phi_B = B\Phi \in \mathbb{R}^{n_B}, \quad (4.2.15)$$

so-called *base parameters*, are only identifiable.

Remark 4.1. *Note however that the representation of the base parameters is not unique, as different choice of basis vectors can be used through, e.g., $\Phi_{\hat{B}} = S\Phi_B = (SB)\Phi = \hat{B}\Phi$ and $Y_{\hat{B}} = Y_B S^{-1}$ with any nonsingular square matrix $S \in \mathbb{R}^{n_B \times n_B}$.*

To this end, least square estimation for the base parameter vector is well-posed as,

$$\min_{\Phi_B} \|A_B \Phi_B - b\|_{\Sigma^{-1}}^2, \quad (4.2.16)$$

where $A_B = AB^\dagger \in \mathbb{R}^{nm \times n_B}$, and B^\dagger is a left Moore-Penrose pseudoinverse of B .

When the sampled trajectory is fully excited so that the reduced regressor matrix A_B satisfies full rank condition, the least squares solution is given by $\Phi_B^* = (A_B^T \Sigma^{-1} A_B)^{-1} A_B^T \Sigma^{-1} b$, with variance

$$\text{var}[\Phi_B^*] = (A_B^T \Sigma^{-1} A_B)^{-1}, \quad (4.2.17)$$

and the Fisher information matrix

$$M_B = (A_B^T \Sigma^{-1} A_B) \quad (4.2.18)$$

Now one can resort to the standard methods of optimal experimental design introduced in Section 4.2.1 in order to achieve a well-conditioned and low-variance least squares estimate of the base parameter Φ_B^* :

$$\min_{q(t)} f(\lambda(A_B^T \Sigma^{-1} A_B)) \quad (4.2.19)$$

Various design criteria have been used, e.g., condition number [56, 65, 24], D-optimality [24], E-optimality [56, 66]. More explicit formulations and optimization algorithms for generating excitation trajectories are to be presented in Section 4.4.

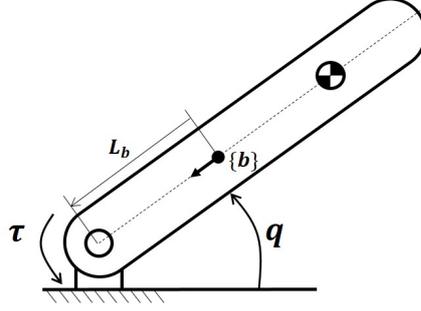


Figure 4.1: 1-link, 1-dof robot

4.2.3 Coordinate Invariance and Normalization

Here we show explicitly how formulating a measure of information through criteria of the form (4.2.9) can be affected by the particular coordinates used to characterize the dynamic model. To illustrate, consider a simple 1-dof robot shown in Figure 4.1. The dynamic equations are given by

$$\tau = \underbrace{\begin{bmatrix} L_b^2 \ddot{q} - gL_b \cos(q) & -2L_b \dot{q} + g \cos(q) & \ddot{q} \end{bmatrix}}_{=Y(q,\dot{q},\ddot{q}) \in \mathbb{R}^{1 \times 3}} \underbrace{\begin{bmatrix} m \\ h_b \\ I_b \end{bmatrix}}_{=\Phi \in \mathbb{R}^3}, \quad (4.2.20)$$

where $L_b \in \mathbb{R}$ is the location of the joint, and $m \in \mathbb{R}^+$, $h_b \in \mathbb{R}$, and $I_b \in \mathbb{R}^+$ are the mass, first, and second moments of the 1-d mass distribution of the link, respectively, all described with respect to the body-fixed frame $\{b\}$. More compactly,

$$\tau = \underbrace{\begin{bmatrix} -2L_b \dot{q} + g \cos(q) & \ddot{q} \end{bmatrix}}_{=Y_B(q,\dot{q},\ddot{q}) \in \mathbb{R}^{1 \times 2}} \underbrace{\begin{bmatrix} h_b - mL_b \\ I_b - mL_b^2 \end{bmatrix}}_{=\Phi_B \in \mathbb{R}^2}, \quad (4.2.21)$$

where

$$\Phi_B = \underbrace{\begin{bmatrix} -L_b & 1 & 0 \\ -L_b^2 & 0 & 1 \end{bmatrix}}_{=B \in \mathbb{R}^{2 \times 3}} \Phi. \quad (4.2.22)$$

Consider a family of sinusoidal excitation trajectories of the form

$$q(t; A, q_0) = A \sin(2\pi t) + q_0, \quad 0 \leq t \leq 2, \quad (4.2.23)$$

parametrized by the amplitude A and offset q_0 . The infinite-dimensional optimization problem (4.2.19) can then be reduced to a two-dimensional parameter optimization problem on $(A, q_0) \in \mathbb{R}^2$. Figure 4.2 shows the contour map of the A-optimality criterion (which is to be minimized for richer/lower information/variance). It can be observed that just by relocating the position of the body-fixed frame $\{\mathbf{b}\}$ or changing physical units, the shape of the objective function changes dramatically, resulting in substantially different optimal excitation trajectories.

More generally, as noted in Remark 4.1, coordinate choices for describing the multi-body dynamic model include not only reference frames and physical units of measurement, but also linear reparametrizations of the base parameter vector Φ_B . To illustrate, let B_1 and B_2 be two arbitrary choices of matrix B . As long as the identifiable set of parameters is consistent without regard to its representation, there should exist some nonsingular square matrix $S_{12} \in \mathbb{R}^{N_B \times N_B}$ such that $B_1 = S_{12}B_2$ holds. The Fisher information matrix M_B then exhibits a linear transformation with respect to the change of base parameter representation, i.e.,

$$M_{B_2} = S_{12}^T M_{B_1} S_{12}. \quad (4.2.24)$$

Since the set of eigenvalues is not invariant to such linear transformations of the Fisher information matrix, i.e.,

$$\lambda(M_{B_1}) \neq \lambda(S_{12}^T M_{B_1} S_{12}), \quad (4.2.25)$$

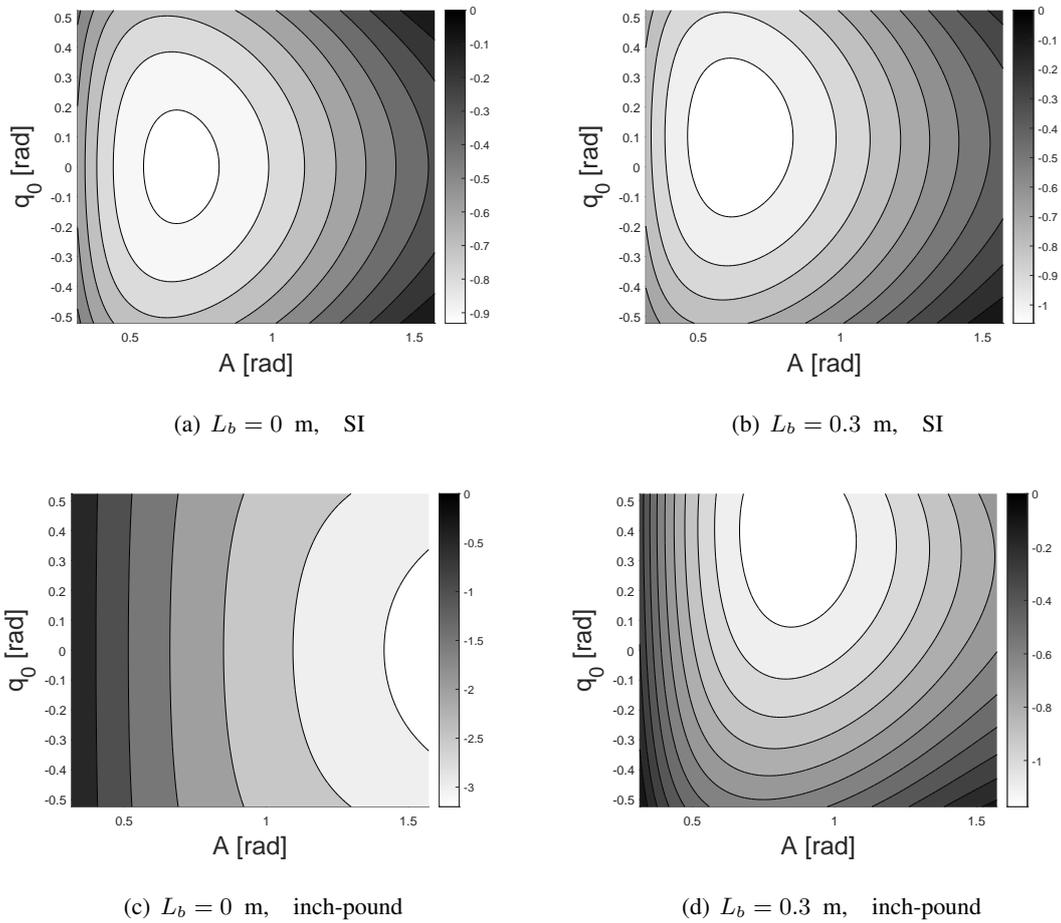


Figure 4.2: Contour maps of A-optimality excitation criterion for the one-dof robot, shown at logarithmic scale and evaluated with different coordinate choices

the choice of base parameter representation also affects conventional experimental designs (4.2.9).

In the work of [64], which is most relevant to ours, although the issue of coordinate invariance is not explicitly raised, it is pointed out that standard ways of specifying a multibody dynamic model (with respect to a particular coordinate system) can more often than not lead to poorly scaled values of the model parameters. This affects the characterization of the excitation criteria in a way that makes certain parameters overly sensitive to measurement noise. To circumvent this problem, a normalized form of the Fisher information matrix is proposed so as to balance the contribution of each entry in M_B . The proposed criterion is of the form

$$J = f(\lambda(M_B \cdot [\text{diag}(\Phi_B^0)]^2)) \quad (4.2.26)$$

$$= f(\lambda(\text{var} [[\text{diag}(\Phi_B^0)]^{-1} \cdot \Phi_B^*]^{-1})) \quad (4.2.27)$$

$$= f(\lambda(A_B^T \Sigma^{-1} A_B \cdot [\text{diag}(\Phi_B^0)]^2)), \quad (4.2.28)$$

where Φ_B^0 is some nominal parameter vector, which for mechanical systems can be easily obtained from, e.g., CAD models.

While this can alleviate the aforementioned difficulties to some extent, we claim that by appealing to geometric methods, normalization can in fact be achieved in a much more physically meaningful way while also ensuring coordinate invariance of the criteria.

4.3 Geometric Excitation Criteria

In this section we propose a new framework for normalizing excitation criterion given nominal values of the parameters. The criterion is concisely formulated to be coordinate-invariant, while the relative sensitivities or variabilities of the estimation parameters are

quantified in a geometrically natural manner. As a result, many of the aforementioned difficulties of existing coordinate-based criteria can be mitigated.

4.3.1 Motivation

The general experimental design problem as reviewed in the previous section is an intrinsically multivariate problem ([58]): the matrix-valued information of the estimation parameter vector is reduced to a scalar-valued criterion to be optimized, with the traditional criterion as formulated in (4.2.9) one possible solution. We instead construct a scalar representation of this matrix-valued information from a geometric coordinate-invariant perspective.

The information matrix given by the inverse of the parameter covariance matrix, i.e., $\text{var}[\Phi_B^*]^{-1} = A_B^T \Sigma^{-1} A_B$, can be viewed as a matrix-valued quantity that represents the local variabilities of the estimated parameter Φ_B^* accruing from the observation noise. Defining a scalar representation of the information matrix can be thought of as quantifying some representative scalar-valued variability of the estimation parameter Φ_B^* on the parameter space. Toward this end, it is noted that any geometric quantifications, e.g., local displacement or volume element, require a well-defined metric predefined on the space.

4.3.2 Pushforward Metric on Observable Parameters

In this section, we define a natural metric on the space of base parameters $\Phi_B = B\Phi$. In order to accomplish this we demonstrate how the natural metric on \mathcal{M}^n defined in Section 3.3.1 can be naturally endowed to the base parameter space \mathcal{N} under the surjective linear mapping $B : \mathcal{M}^n \rightarrow \mathcal{N}$.

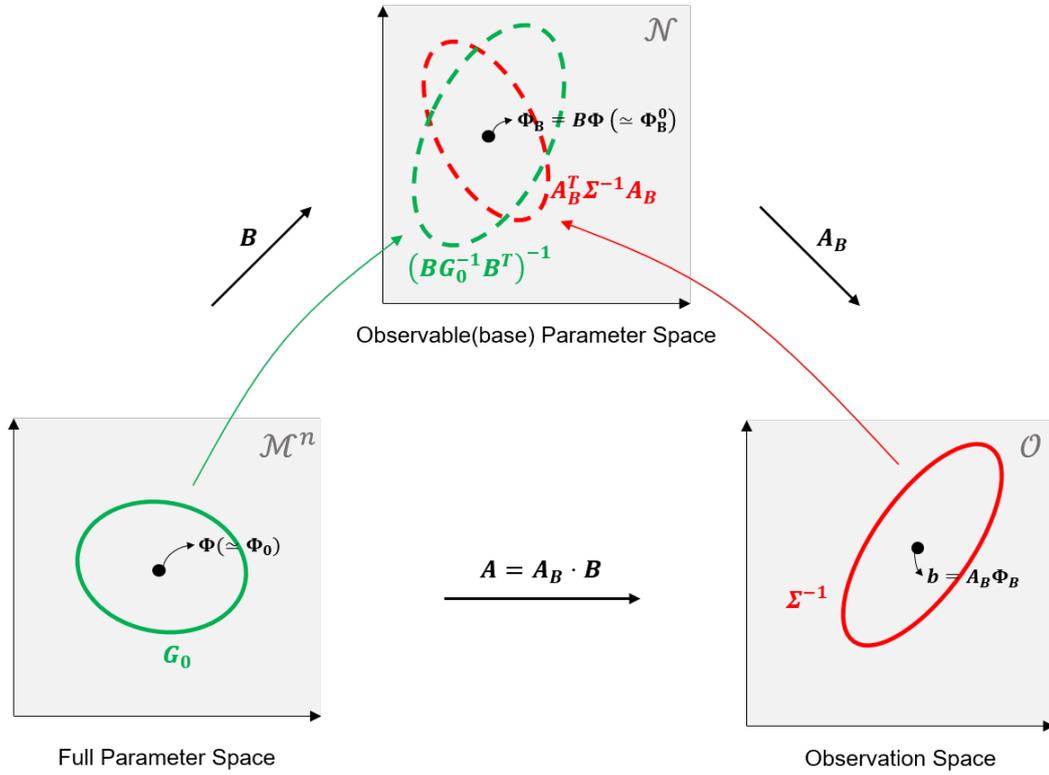


Figure 4.3: Geometric distortion view of the proposed excitation criterion. The metrics characterized with (green and red) bold ellipses are predefined respectively on the full parameter space and observation space. The (green and red) dashed ellipses represent respectively the pushforward and pullback of the metrics to the observable(base) parameter space under the linear mappings B and A_B .

Recall, a natural Riemannian metric on \mathcal{M} is given by,

$$ds^2 = \frac{1}{2} \text{tr}(P^{-1}dPP^{-1}dP) = d\phi^T g(\phi)d\phi, \quad (4.3.29)$$

where the metric $g(\phi) \in \mathbb{R}^{10 \times 10}$ is the pullback of the affine-invariant metric defined on the space of symmetric positive definite matrices, $\mathcal{P}(4)$, under the one-to-one linear mapping P . Also, for n -link multibody system, Riemannian metric on the whole set of inertial parameters \mathcal{M}^n is defined as,

$$ds^2 = \sum_{i=1}^n d\phi_i^T g(\phi_i)d\phi_i = d\Phi^T G(\Phi)d\Phi, \quad (4.3.30)$$

$$G(\Phi) = \text{diag}(g(\phi_1), \dots, g(\phi_n)) \in \mathbb{R}^{10n \times 10n}. \quad (4.3.31)$$

We denote

$$G_0 \triangleq G(\Phi_0) \quad (4.3.32)$$

as the Riemannian metric evaluated at some given nominal inertial parameter value Φ_0 . As discussed in Remark 3.5, with the constant Riemannian metric G_0 , the ensuing space \mathcal{M}^n is isometric to a flat Euclidean space; considering the reparametrization $\Phi' = G_0^{1/2}\Phi$, the metric is given by the standard Euclidean metric, i.e., $ds^2 = d\Phi'^T d\Phi'$. Assuming that the region of interest is near the nominal value Φ_0 , the metric G_0 allows a reasonable approximation of the curved Riemannian structure with Riemannian metric G . Also, we again note that the constant metric G_0 is a well-defined coordinate-invariant Riemannian metric on \mathcal{M}^n .

A natural consequence for endowing a metric on the base parameter space \mathcal{N} would be to enforce isometry with respect to \mathcal{M}^n , i.e.,

$$d\Phi_B^T H_0 d\Phi_B \doteq d\Phi G_0 d\Phi, \quad (4.3.33)$$

where $H_0 \in \mathbb{R}^{n_B \times n_B}$ is a constant Riemannian metric on \mathcal{N} to be defined. If the mapping B is not just surjective but one-to-one, the isometry is trivially achieved by letting

$H_0 = B^{-T}G_0B^{-1}$. For general surjective mapping B , the displacement $d\Phi_B$ on \mathcal{N} actually imposes a redundancy on $d\Phi$ with the relation $Bd\Phi = d\Phi_B$. The following intuitive minimal principle can be enforced to propose a unique choice of metric on \mathcal{N} :

$$\begin{aligned} d\Phi_B^T H_0 d\Phi_B &\doteq \min_{d\Phi \in \{x: Bx=d\Phi_B\}} d\Phi G_0 d\Phi \\ &= d\Phi_B^T (BG_0^{-1}B^T)^{-1} d\Phi_B. \end{aligned} \quad (4.3.34)$$

Consequently, the constant Riemannian metric H_0 , *pushforwarded* from \mathcal{M}^n to \mathcal{N} , is given by

$$H_0 \triangleq (BG_0^{-1}B^T)^{-1}. \quad (4.3.35)$$

We use the following coordinate-invariant inner product notation,

$$\langle x, y \rangle \triangleq x^T H_0 y, \quad (4.3.36)$$

where $x, y \in \mathbb{R}^{n_B}$ are formally the coordinate values of the elements in the tangent space of the linear manifold \mathcal{N} .

Remark 4.2. *The inherited metric of the form $(BG_0^{-1}B^T)^{-1}$ in (4.3.35) has the following alternative interpretation. Let the given nominal value Φ_0 be a random vector whose precision is proportional to the natural metric, i.e.,*

$$\text{var}[\Phi_0]^{-1} = G_0 \quad (4.3.37)$$

(Inverse covariance matrix $\text{var}[\cdot]^{-1}$ which corresponds to the information matrix also represents the precision of the estimate). Then the precision of the nominal base parameter Φ_B^0 is straightforwardly given by

$$\begin{aligned} \text{var}[\Phi_B^0]^{-1} &= \text{var}[B\Phi_0]^{-1} \\ &= (B\text{var}[\Phi_0]B^T)^{-1} \\ &= (BG_0^{-1}B^T)^{-1}, \end{aligned}$$

which turns out to be identical to the metric in (4.3.35) derived from solving (4.3.34).

4.3.3 Coordinate-invariant Criterion

We now present our coordinate-invariant formulation of the general excitation criterion.

Proposition 4.1 (Geometric Excitation Criterion). *Symmetric functions f of eigenvalues of a normalized information matrix defined as,*

$$J(A_B) = f \left(\lambda \left(A_B^T \Sigma^{-1} A_B \cdot H_0^{-1} \right) \right), \quad (4.3.38)$$

or equivalently,

$$J(A) = f \left(\lambda_{\neq 0} \left(A^T \Sigma^{-1} A \cdot G_0^{-1} \right) \right), \quad (4.3.39)$$

is invariant to the choice of base parameter representations, body-fixed reference frames, and physical units ($\lambda_{\neq 0}(X)$ denotes the nonzero eigenvalues of the symmetric positive semi-definite matrix X).

Proof. The proof is given in the Appendix A.4. □

From the following identity,

$$\begin{aligned} J(A_B) &= f \left(\lambda \left(\text{var}[\Phi_B^*]^{-1} H_0^{-1} \right) \right) \\ &= f \left(\lambda \left(H_0^{-1/2} \text{var}[\Phi_B^*]^{-1} H_0^{-1/2} \right) \right) \\ &= f \left(\lambda \left(\text{var} \left[H_0^{1/2} \Phi_B^* \right]^{-1} \right) \right), \end{aligned} \quad (4.3.40)$$

the proposed optimality criteria be interpreted as normalized versions of traditional optimality criteria; the estimation parameters are normalized via $\Phi_B^* \leftarrow H_0^{1/2} \Phi_B^*$.

As alluded to earlier, geometrically the proposed criteria is constructed so as to encode the intrinsic local distortion of the regressor mapping $A_B : \mathcal{N} \rightarrow \mathcal{O}$ in a coordinate-invariant way. Referring to Figure 4.3, one can regard the information matrix

of the parameter estimate $A_B^T \Sigma^{-1} A_B$ as the *pullback* of the observation space metric Σ^{-1} under the linear regressor mapping A_B . Then, the set of generalized eigenvalues of the pullback metric $A_B^T \Sigma^{-1} A_B$ relative to the natural metric H_0 become the invariant measures that capture how much the local geometry gets distorted from the mapping A_B in comparison to the natural geometry predefined on the parameter space \mathcal{N} with metric H_0 .

Remark 4.3. *Under our geometric framework, the normalization given in [64] can be understood as adopting a diagonal scale-normalizing (but coordinate-dependent) metric $[\text{diag}(\Phi_B^0)]^{-2}$, on the parameter space.*

Below is a detailed comparison of the proposed criterion with their unnormalized counterparts.

- A-optimality: The traditional A-optimality criterion is defined by the expected parameter error using the standard Euclidean metric, i.e.,

$$\begin{aligned} \text{tr}((A_B^T \Sigma^{-1} A_B)^{-1}) &= \text{tr}(\text{var}[\Phi_B^*]) \\ &= \mathbb{E}[(\Phi_B^* - \Phi_B)^T (\Phi_B^* - \Phi_B)], \end{aligned}$$

In contrast, the proposed normalized A-optimality criteria is defined to be the expected quadratic parameter error but in coordinate-invariant form, i.e.,

$$\mathbb{E}[(\Phi_B^* - \Phi_B)^T H_0 (\Phi_B^* - \Phi_B)]. \quad (4.3.41)$$

- D-optimality: D-optimal designs are already invariant to reparametrizations of the parameters. Indeed, the normalized D-optimality criterion turns out to be same as the original D-optimality criterion shifted by the constant value $\log |H_0|$:

$$\begin{aligned} -\log |A_B^T \Sigma^{-1} A_B H_0^{-1}| \\ = -\log |A_B^T \Sigma^{-1} A_B| + \log |H_0| \end{aligned} \quad (4.3.42)$$

- E-optimality: The E-optimality criterion is defined by the maximum eigenvalue of the estimator covariance matrix:

$$\begin{aligned}\lambda_{\max}((A_B^T \Sigma^{-1} A_B)^{-1}) &= \lambda_{\max}(\text{var}[\Phi_B^*]) \\ &= \max_{x^T x=1} x^T \text{var}[\Phi_B^*] x \\ &= \max_{x^T x=1} \text{var}[x^T \Phi_B^*],\end{aligned}$$

which as derived above can geometrically be understood as the largest variance of the estimated parameter projected along the unit-normalized directional vector x . Similarly, the proposed normalized E-optimality criterion

$$\lambda_{\max}((A_B^T \Sigma^{-1} A_B)^{-1} H_0) \quad (4.3.43)$$

is equivalent to

$$\max_{\langle x, x \rangle=1} \text{var}[\langle x, \Phi_B^* \rangle], \quad (4.3.44)$$

where the Euclidean inner products are replaced by $\langle \cdot, \cdot \rangle$ defined in (4.3.36).

- Condition number: Traditional condition number optimization seeks to find the estimator with isotropic variance, i.e. $\text{var}[\Phi_B^*] \sim \mathbb{1}$, while the proposed normalized condition number

$$\sqrt{\frac{\lambda_{\max}(A_B^T \Sigma^{-1} A_B H_0^{-1})}{\lambda_{\min}(A_B^T \Sigma^{-1} A_B H_0^{-1})}} \quad (4.3.45)$$

is minimized when the precision of the estimator is proportional to the metric H_0 , i.e.,

$$\text{var}[\Phi_B^*]^{-1} = A_B^T \Sigma^{-1} A_B \sim H_0.$$

4.4 Optimal Excitation Trajectory Generation

Finding optimal excitation trajectories is an infinite-dimensional variational problem, and in this section we describe direct numerical methods for finding finite-dimensional approximate solutions. For the conventional (i.e., non-geometric) criteria proposed in the literature, the variational formulation of the problem is of the form

$$\min_{q(t)} J = f(\lambda(M_B)), \quad (4.4.46)$$

where $M_B = A_B^T \Sigma^{-1} A_B = \sum_{i=1}^m Y_B(t_i)^T \Sigma_\epsilon^{-1} Y_B(t_i)$, typically subject to dynamic equations of the form (3.2.13) and other physical constraints such as joint angle, velocity, and torque limits. Direct methods parametrize the trajectory $q(t)$ in terms of a finite set of parameters $P = \{p_1, \dots, p_{n_p}\}$, e.g., B-spline control points. The parametrized trajectory $q(t, P)$ and its time derivatives $\dot{q}(t, P)$, $\ddot{q}(t, P)$ are now explicit functions of P , and the variational problem reduces to the following finite-dimensional parameter optimization:

$$\min_P J = f \left(\lambda \left(\sum_{i=1}^m Y_B(t_i, P)^T \Sigma_\epsilon^{-1} Y_B(t_i, P) \right) \right). \quad (4.4.47)$$

Note that the regressor $Y_B(q(t_i), \dot{q}(t_i), \ddot{q}(t_i))$ is also an explicit function of P .

Efficient and reliable numerical optimization of J in (4.4.47) requires analytic gradients with respect to the parameters P ; the accumulation of numerical errors resulting from finite-difference estimates of the true gradient more often than not leads to ill-conditioning and poor convergence. Gradients of J can be analytically obtained via application of the chain rule [57]:

$$\frac{\partial J}{\partial p} = \sum_{i=1}^{n_B} \frac{\partial f}{\partial \lambda_i} \text{tr}(\nabla_{M_B} \lambda_i(M_B)^T \frac{\partial M_B}{\partial p}), \quad (4.4.48)$$

where the notation ∇_X denotes the derivative operator on a function with respect to a

matrix X , i.e., $(\nabla_X g)_{ij} \triangleq \frac{\partial g}{\partial X_{ij}}$. The term $\frac{\partial M_B}{\partial p}$ is given by

$$\frac{\partial M_B}{\partial p} = \sum_{i=1}^m \left[\frac{\partial Y_B(t_i, P)^T}{\partial p} \Sigma_\epsilon^{-1} Y_B(t_i, P) + Y_B(t_i, P)^T \Sigma_\epsilon^{-1} \frac{\partial Y_B(t_i, P)}{\partial p} \right], \quad (4.4.49)$$

where

$$\frac{\partial Y_B(t_i, P)}{\partial p} = \frac{\partial Y(t_i, P)}{\partial p} B^\dagger, \quad (4.4.50)$$

can be computed in a systematic and efficient manner using the recursive algorithm for the inverse dynamics computation of multibody systems [67, 68, 69]. Lie-group based recursive algorithms derived for computing the gradient of the general types (joint torque/ external wrench) of regressor matrices for tree-structured multibody systems are provided in the Appendix B.3.

Gradients for the geometric trajectory excitation criterion proposed in this paper can be obtained by a slight modification of the above procedure. First, evaluating the cost is trivial via the following substitution for M_B :

$$M_B \leftarrow M_B \cdot H_0^{-1}. \quad (4.4.51)$$

For the computation of the gradient, the following substitution is made in place of (4.4.48):

$$\frac{\partial M_B}{\partial p} \leftarrow \frac{\partial M_B}{\partial p} \cdot H_0^{-1}, \quad (4.4.52)$$

since H_0 is a given constant matrix.

4.5 Determination of Effectively Identifiable Parameter Set

For fairly low-dimensional mechanical systems, e.g., standard industrial manipulators with six or fewer dofs, generating sufficiently rich sample trajectory through, e.g., excitation trajectory optimization, and finding reliable estimate for the entire set of base parameters is more often than not feasible. This is generally not the case for more complex, high-dof structures.

First of all, hard physical/safety constraints typically imposed on high-dof floating-base systems, e.g., humanoids and multi-legged robots, can largely restrict the admissible dynamic motion; in this type of systems the safety issue is mainly concerned with having to strictly ensure whole-body balance and not to fall over. For this reason, relatively smaller gain can be made from having optimal exciting motions compared to, e.g., fixed-base manipulators. Moreover, available sensor measurements for identification are usually more restricted for these high-dof robotic systems, as torque sensors may not be provided for every joints. This limits the amount of information gained within the fixed time interval and may require inefficiently long duration of excitation for reliable identification purely from the observation data.

Furthermore, some inherent kinematic features of the robot and sensors, e.g., kinematic complexity of the mechanism design, sensor placement locations relative to the bodies, etc, may also badly influence the relative sensitivities of the dynamic parameters, as they directly pertain to the characterization of the regressor matrix Y_B as relevant kinematic parameters (these include screw axes of the joints, relative displacements of the bodies, etc).

4.5.1 Reduced Identification

As the aforementioned aspects all come to the effect in a compositive way for complex high-dof systems, pursuing for the entire set of parameters to be sufficiently excited for identification can easily be inefficient or may not even be viable in practice. In this section, we propose a coordinate-invariant framework that rather captures the reduced set of parameters that can reliably be identified with a reasonable precision from the given feasible excitation trajectories. The key is in comparing the respective sensitivities of the parameter based on the normalized covariance matrix. The proposed reduced identification strategy is proceeded as follows:

- a.1) Concatenate all the regressor matrices over the given feasible trajectories to calculate the estimation covariance matrix $C = (A_B^T \Sigma^{-1} A_B)^{-1}$ for the base paramters.
- a.2) Evaluate the eigenvalues $0 \leq \lambda_1 \leq \dots \leq \lambda_{n_B}$ and corresponding eigenvectors u_1, \dots, u_{n_B} of the normalized estimation covariance matrix $H_0^{1/2} C H_0^{1/2}$.
- a.3) Transform $v_i = H_0^{-1/2} u_i$, $\forall i = 1, \dots, n_B$, so that v_i represent the direction of the i-th largest projected estimation variance characterized by the corresponding eigenvalue,

$$\lambda_i = \text{var}[\langle v_i, \Phi_B^* \rangle], \quad (4.5.53)$$

under the unit-norm and orthogonality constraints, $\langle v_i, v_i \rangle = 1$ and $\langle v_i, v_j \rangle = 0$, which holds from $u_i^T u_i = 1$ and $u_i^T u_j = 0$.

a.4) Set a parameter estimation variance threshold $\epsilon_\lambda > 0$ and define matrices

$$V_- = [v_1, \dots, v_r], \quad (4.5.54)$$

$$V_+ = [v_{r+1}, \dots, v_{n_B}], \quad (4.5.55)$$

$$\Lambda_- = \text{diag}(\lambda_1, \dots, \lambda_r), \quad (4.5.56)$$

$$\Lambda_+ = \text{diag}(\lambda_{r+1}, \dots, \lambda_{n_B}), \quad (4.5.57)$$

such that λ_r is the largest eigenvalue less than or equal to ϵ_λ .

a.5) Identify only the low estimation variance parameters by solving,

$$\min_{\Phi_B} \|A_B \Phi_B - b\|_{\Sigma^{-1}}^2 \quad (4.5.58)$$

$$\text{s.t.} \quad V_+^T H_0 (\Phi_B - \Phi_B^0) = 0, \quad (4.5.59)$$

, which has a closed-form solution,

$$\Phi_B^* = V_- \Lambda_- V_-^T A_B^T \Sigma^{-1} b + V_+ V_+^T H_0 \Phi_B^0. \quad (4.5.60)$$

The reduced set of sufficiently excited parameters are given by,

$$\langle v_i, \Phi_B^* \rangle = \lambda_i v_i^T A_B^T \Sigma^{-1} b, \quad (4.5.61)$$

for $i = 1, \dots, r$.

The proposed method essentially updates only the reduced set of parameters with a certain level of confidence as specified by the eigenvalues λ_i , reflecting the sensitivities (or variances) of the projected parameters $\langle v_i, \Phi_B \rangle$. We point out that the presented parameter reduction strategy is not merely a combinatoric selection of the entries of the vector-valued parameters, but determines non-trivial linear combinations in a coordinate-free manner. Also, determining the legitimate scale of estimation variance threshold value ϵ_λ in step 4 is eased by the fact that all the eigenvalues λ_i of the normalized covariance matrix are coordinate invariant and dimensionless.

4.5.2 Reduced Optimal Excitation

The present framework can also be extended to readily generate optimal excitation trajectory for the reduced set of effectively excitable parameters. Parameters in the largest estimation variance directions are disregarded in a sequential fashion, until the optimized excitation criterion concerned with the reduced set of parameters reduces to a reasonably small dimensionless value. The algorithm is proceeded as follows:

- b.1) Initialize $r = n_B$ and decide a threshold ϵ_J for the optimality criterion J .
- b.2) Optimize the exciting trajectory with respect to the criterion $J = f(\lambda_1, \dots, \lambda_r)$, where f should be some fixed form of symmetric function of r number of positive scalar values.
- b.3) Evaluate the optimized objective J_{opt} . If $J_{opt} > \epsilon_J$, substitute $r \leftarrow r - 1$ and repeat from step b.2). Otherwise, end the iteration.
- b.4) With the optimized exciting trajectory, run the algorithm a.1)-5) for $\epsilon_\lambda = \lambda_r$ to identify Φ_B^* .

Again determining the threshold values ϵ_J and ϵ_λ for J and λ_i are both eased by the coordinate invariance and dimensionlessness of J and λ_i .

4.6 Simulation Study

We now demonstrate the advantages of our coordinate-invariant natural criterion through extensive numerical parameter identification experiments. The main claims that we experimentally verify through our case studies are: (i) normalization of the excitation criterion improves the accuracy of the poorly scaled parameters more uniformly; (ii) the quality of excitation trajectories using the existing criteria can conspicuously vary

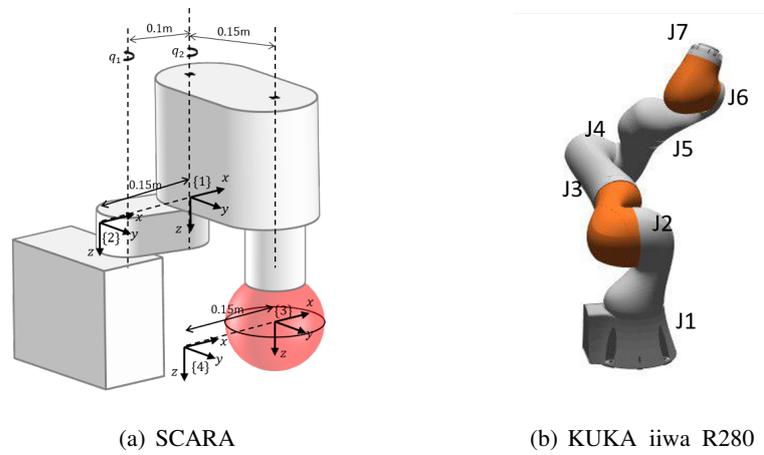


Figure 4.4: The figures show 2-dof SCARA robot with unknown payload (red solid sphere) and reference frames considered, and joint specifications of 7-dof KUKA iiwa R280 industrial manipulator.

depending on different choices of coordinates; (iii) the proposed parameter reduction method can significantly improve the robustness and generalizability in high-dimensional systems identification, specifically when not all the structurally identifiable set of parameters (base parameters) can be effectively identifiable from the given trajectory samples.

4.6.1 SCARA with Unknown Payload

Optimal excitation of a 2-dof SCARA robot for identifying the unknown payload is considered. The simple, yet non-trivial, case is made to facilitate the presentation and comparison of the results more intuitive and concrete; owing to the low dimensionality of the problem, a canonical representation of the identifiable set of parameters can be easily chosen and compared with no complex compositions of the parameters belonging to different links.

Figure 4.4(a) shows the kinematic specifications of the SCARA robot and unknown

(uniform sphere) payload loaded at the end-effector link. Joint torque values are observed with additive zero-mean Gaussian noise of standard deviation 0.03 Nm. Here we only excite and identify the inertial parameters of the payload while assuming the parameters associated with the rest of the robot dynamics are known. In this setting, there are four number of identifiable parameters associated with the rigid body dynamics of the payload which are given by mass m , x - y components of the first moment h^x, h^y , and rotational inertia about the z -axis I^{zz} .

Least square parameter identification results are compared across using the pure non-normalized condition number and its normalized counterparts for optimal excitation trajectory generation. Moreover, we examine the effect of using different body-fixed reference frames in adopting coordinate-dependant criterion that includes both the non-normalized one and the normalized one of [64]. Four body-fixed coordinate frames $\{1\}$, $\{2\}$, $\{3\}$, $\{4\}$ are considered as shown in Figure 4.4(a). In particular, the frame $\{3\}$ is exactly located at the center of mass of the payload.

50 number of joint kinematics samples, i.e., $P = \{q_i, \dot{q}_i, \ddot{q}_i\}_{i=1}^{50}$, used for identification are directly optimized using the SQP solver of MATLAB built-in function `fmincon` with a recursive analytic gradient computation. Initial random excitation is set by uniformly sampling the values q, \dot{q}, \ddot{q} in their pre-specified limits.

Table 4.1 compares the RMS errors of the identified parameters depending on the coordinate frames and normalization techniques used for excitation trajectory generation (the identified parameters h^x, h^y, I^{zz} are transformed to be the estimated x - y position of center of mass p^x, p^y and length of the moment arm r^z respectively). Identified parameters from the initial random samples result in significant errors for p^y and r^z . Such ill-conditioned identification result is alleviated to some extent from using samples with traditional condition number optimization. Meanwhile, the results from optimization of condition number with normalization technique of [64] vary conspicuously depending

on the coordinate choices. This can in fact be deduced from the fact that specifically the coordinates $\{1\}$, $\{3\}$ are positioned in locations which ill define the normalization; that is, the prior value of the x component of the mass center is evaluated to be near zero $p^x = 0.02$ [mm] in coordinate $\{1\}$ and $\{3\}$, and this makes normalized criteria unnecessarily pay too much attention to h^x . In contrast, our proposed approach for normalization is completely free of the aforementioned downside arising from coordinate dependency, while also showing improved balance in errors of identified parameters.

Table 4.1: 2-dof SCARA robot with unknown payload: identification results with “condition number” optimization

Standardization	Coordinate (frame)	RMS error of the identified parameters (in coordinate {3})			
		m [kg]	$p^x = \frac{h^x}{m}$ [mm]	$p^y = \frac{h^y}{m}$ [mm]	$r^z = \sqrt{\frac{I_{zz}}{m}}$ [mm]
None	1	0.03	0.78	1.27	5.24
	2	0.03	0.85	1.56	5.33
	3	0.03	0.70	1.40	5.14
	4	0.03	0.79	1.30	4.92
Diagonal	1	1.53	0.41	317.90	119.67
	2	0.04	1.07	1.39	2.87
	3	0.79	0.18	19.11	20.46
	4	0.07	2.31	2.19	4.15
Proposed	(Invariant)	0.06	0.60	1.09	1.56
Initial random excitation		0.04	0.16	149.82	89.02
Prior value		4.58	0.02	1.94	86.07
Ground truth value		4.00	0.00	0.00	81.65

4.6.2 KUKA iiwa R280 Manipulator

Optimal exciting trajectory generation for KUKA iiwa R820 manipulator is conducted followed by least square parameter identification. Joint angle trajectories $q_i(t), i = 1, \dots, 7$ are parametrized with Fourier basis functions,

$$q_i(t) = \sum_{j=1}^{n_f} a_{ij} \cos(j \cdot \omega t) + b_{ij} \sin(j \cdot \omega t), \quad (4.6.62)$$

with number of modes $n_f = 5$, base frequency $\omega = 0.3\pi$. The fixed duration of time for excitation is set to ten seconds while samples are drawn uniformly at the sampling frequency of 1kHz. The optimization parameters are the coefficients of the basis given by $P = \{\{a_{ij}, b_{ij}\}_{i=1}^7\}_{j=1}^{n_f}$, and joint angle and velocity limits are appropriately specified as constraints. Torque values are observed with additive zero-mean Gaussian noise of standard deviation 0.1 Nm. True inertial parameter values of each link are determined from CAD data, while the prior values used for the normalization of the excitation criteria are obtained from the effective mass, center of mass and rotational inertia of the collection of finite number of point masses randomly drawn from the groundtruth mass distribution of each link.

Parameter identification results are compared across using the pure unnormalized E-optimality and its normalized counterparts for optimal excitation trajectory generation. In particular, we examine the effect of using different base parameter representations in using the existing coordinate-dependant criterion. The base parameter representation, $\Phi_{B_1} = B_1 \Phi$, for “Base 1” is characterized using numerical SVD of a regressor matrix given randomly sampled joint trajectories as in [15]; 43 number of base parameters are found. Then, for each of the other three base parameter representations, random sampled nonsingular square matrix S_{rand} is multiplied to the “Base 1” base parameter representation, e.g., $\Phi_{B_2} = S_{\text{rand}} \Phi_{B_1}$. We again use the SQP optimizer in MATLAB

built-in function `fmincon` with a recursive analytic gradient computation to conduct parameter optimization for obtaining excitation trajectories.

Table 4.2 shows the comparative identification results that include dimensionless quadratic parameter error, $\sqrt{\frac{1}{n_B} \langle \Phi_B^* - \Phi_B, \Phi_B^* - \Phi_B \rangle}$, RMS joint torque prediction errors $\tau - \hat{\tau} = \tau - \Gamma_B \Phi_B^*$, and also RMS acceleration errors $M(q)^{-1}(\tau - \hat{\tau})$ which better reflects the effective disturbance acting at the joints accruing from the model error when using model-based tracking control [70, p. 299]. Trajectory samples used for the validation of the prediction errors are obtained by uniformly sampling 10,000 instances of joint kinematics information q, \dot{q}, \ddot{q} within their limits.

Remark 4.4. *To be more precise in terms of the joint acceleration error, the closed loop dynamics for the computed torque control law, $\tau = M(q, \Phi_B^*)\{\ddot{q}_d - K_v \dot{e} - K_p e\} + C(q, \dot{q}, \Phi_B^*)\dot{q} + g(q, \Phi_B^*)$, is given by*

$$M(q)^{-1}(\tau - \hat{\tau}) = \ddot{e} + K_v \dot{e} + K_p e, \quad (4.6.63)$$

where $e = q - q_d$ is the joint tracking error and K_p, K_v are the user-defined positive definite gain/impedance matrices. Therefore, the torque prediction error weighted by the inverse mass matrix can be viewed as a disturbance to the closed-loop joint space dynamics accruing from the model error.

It can be observed that the exciting trajectories from normalized criterion come in improved identification accuracies and torque/acceleration predictions. In particular, the acceleration errors of the last joint from unnormalized criterion are highly erroneous, whereas the absolute values of torque errors of the corresponding joint seem comparably reasonable. This is in fact due to highly low inertia concerned with the last joint and the unnormalized criterion fails to take into an account the relative range of precision of the low inertias. While the overall performance indices are improved using

the scale-normalization technique [64], the results are to some extent variable depending on the choice of coordinates. In contrast, the proposed standardization results in the best parameter accuracy, to highlight again, with an added advantage of guaranteed invariance to the coordinate choices.

Table 4.2: KUKA iiwa R280: identification results with “E-optimality” optimization

Standardization	Coordinate (base)	Param. error	RMS torque error $\tau - \hat{\tau}$ [Nm]				RMS acc. error $M^{-1}(\tau - \hat{\tau})$ [rad/s ²]			
			J 1,2	J 3,4	J 5,6	J 7	J 1,2	J 3,4	J 5,6	J 7
None	1	2.967	0.023	0.021	0.012	0.007	0.048	0.295	1.263	20.843
	2	2.757	0.020	0.018	0.011	0.006	0.040	0.234	1.143	18.366
	3	2.356	0.018	0.016	0.009	0.006	0.037	0.226	1.022	16.678
	4	2.724	0.020	0.017	0.011	0.006	0.040	0.243	1.131	18.934
Diagonal	1	0.307	0.011	0.009	0.002	0.001	0.026	0.164	0.288	3.324
	2	1.705	0.012	0.011	0.007	0.004	0.023	0.120	0.756	12.361
	3	1.288	0.013	0.011	0.006	0.003	0.027	0.166	0.617	9.448
	4	1.986	0.014	0.012	0.008	0.005	0.026	0.147	0.810	13.938
Proposed	(Invariant)	0.184	0.010	0.008	0.002	0.0004	0.025	0.151	0.275	1.278
Initial random excitation		4.350	0.070	0.052	0.019	0.010	0.146	0.914	2.140	29.890

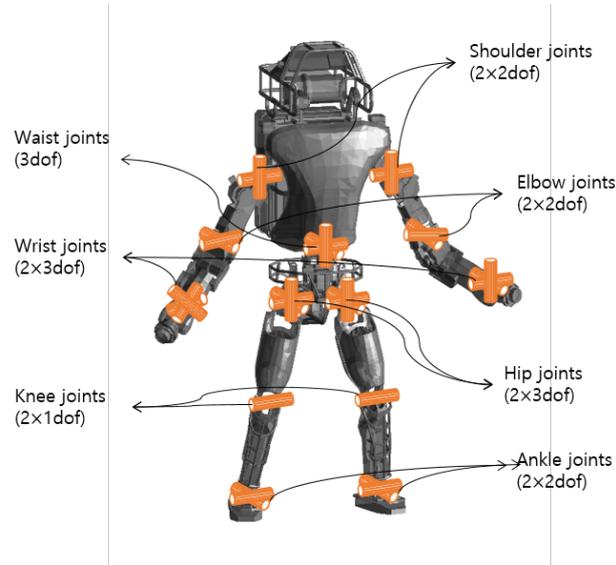


Figure 4.5: Atlas V5 joint specifications

4.6.3 Atlas V5 Humanoid Robot

Inertial parameter identification of Atlas V5 humanoid model having 29 joints and 30 body links is considered (See Fig. 4.5). We consider the case for which the humanoid is in dynamic balance with a single foot placed onto the flat ground without slip; then one can regard the underactuated full dynamics as a fixed-base tree-structure system, provided that ZMP always lies in the interior of the support polygon and vertical reactive force is positive. Rather than having torque sensors at all the joints, we adopt a single standard force plate placed under the support foot which measures 6-axis external Ground Reaction (GR) force and moment [23]. The force and moment outputs are observed with additive zero-mean Gaussian noise of standard deviations 0.5 N and 1.5 Nm respectively.

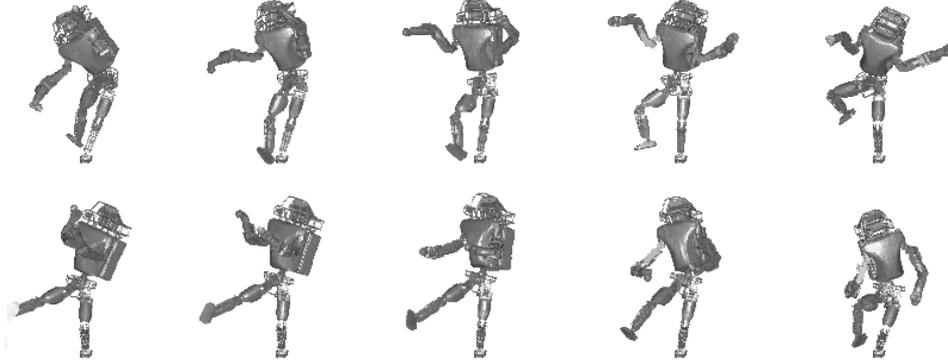


Figure 4.6: Snapshots of 10-th (top) and 20-th (bottom) balancing motion trajectories used for (reduced) dynamic identification and validation

As introduced in section 4.5, we examine the efficacy of using the proposed reduced identification strategy given a database of random generated feasible trajectories. The random trajectories are sampled by, first, sampling the feasible static initial and final joint configurations examining the COM positions. Then each pair of these boundary configurations are connected with B-spline trajectory, whose control points are repeatedly random-sampled until the dynamic balance constraints as well as joint angle, velocity limit constraints are satisfied at the time sampling frequency of 15Hz. Out of 20 random generated feasible trajectories (Fig. 4.6) (each lasting 3 seconds), 10 motions are used for identification and the remaining 10 motions are used for validation (the trajectory samples are drawn at the sampling frequency of 300Hz). True inertial parameter values are determined from CAD data, and prior values are obtained through the same procedure as in the previous section. The scale-free variance threshold value ϵ_λ is set as $\epsilon_\lambda = 10^{-2}$.

70 number of sufficiently excited parameters out of 204 base parameters (Fig. 4.7) are identified from the data using the reduced identification strategy (4.5.60). Table

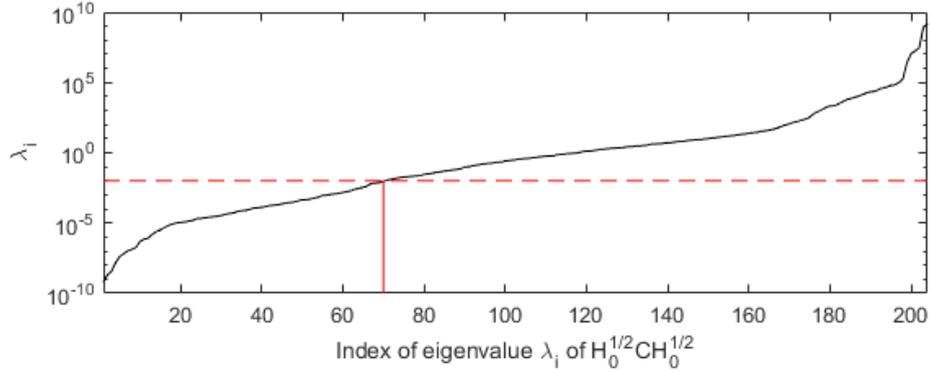


Figure 4.7: Plot of eigenvalues of the normalized covariance matrix $H_0^{1/2}CH_0^{1/2}$ for Atlas V5 balancing trajectories. The horizontal dashed line represents the threshold value $\epsilon_\lambda = 10^{-2}$ used for parameter selection.

Table 4.3: Identification results on Atlas V5 humanoid with parameter reduction: Output validation

	Param. error	GR force	GR moment
Prior	0.269	61.768	11.770
Full identification	2048.3	0.017	0.141
Reduced identification	0.154	0.009	0.024

4.3 compares the identification result to the pure least squares estimation of the entire base parameters. The full-identified parameters are highly erroneous as presented with the normalized quadratic error metric. Indeed, as shown in Fig. 4.7, large portion of the estimation parameters are subject to exceedingly large variance. While the GR force/moment prediction accuracy (in Table 4.3) for the full identification is fairly comparable to the reduced identification, Table 4.4 shows that the full identification comes in a poor prediction accuracy in terms of joint space dynamics; some particular set of joints, e.g., hip, ankle, wrist, are shown to result in large acceleration error for the full identification case. In contrast, the reduced identification allows much more generalizable result also for joint space dynamics. Note that the nominal/prior values used for the reduced identification are fairly rough estimates, as it shows the worst overall prediction accuracy.

Table 4.4: Identification results on Atlas V5 humanoid with parameter reduction: Validation on joint space dynamics

		RMS torque error $\tau - \hat{\tau}$ [Nm]						
	hip (2×3dof)	knee (2×1dof)	ankle (2×2dof)	waist (3dof)	shoulder (2×2dof)	elbow (2×2dof)	wrist (2×3dof)	
Prior	11.279	14.834	10.101	11.717	1.011	0.304	0.062	
Full identification	0.131	0.132	0.119	0.135	0.053	0.034	0.015	
Reduced identification	0.019	0.017	0.034	0.018	0.009	0.008	0.002	

		RMS acc. error $M^{-1}(\tau - \hat{\tau})$ [rad/s ²]						
	hip (2×3dof)	knee (2×1dof)	ankle (2×2dof)	waist (3dof)	shoulder (2×2dof)	elbow (2×2dof)	wrist (2×3dof)	
	4.584	4.399	8.304	4.767	2.908	3.643	22.351	
	1.197	0.799	5.820	0.400	0.238	1.062	18.494	
	0.120	0.119	0.818	0.076	0.035	0.123	2.122	

4.7 Conclusion

We have proposed new normalized criteria for optimal excitation of multibody mechanical systems. Our criteria more naturally capture the variance of the multivariate estimation parameters in a coordinate-invariant way, respecting the geometry of the parameter space. Practical benefit of using the proposed normalized criterion for optimal excitation trajectory generation lies not only in achieving improved robustness of parameter identification, but also in making the ensuing numerical optimization provably invariant to the coordinate choices. Moreover, simple linear relation to the traditional criterion allows directly applying any existing gradient-based optimization method for excitation trajectory generation without significant additional computation.

The proposed natural criteria is also generally useful for assessing the effective excitability of the parameters for complex high-dimensional systems. In particular, parameter reduction method which examines the respective sensitivities of the parameters based on the geometrically normalized covariance matrix is shown to be effective for robust identification of a humanoid robot with limited source of sensor measurements.

5

Geometric Robust Adaptive Control of Robot Manipulators

5.1 Introduction

In order to attain a high-performance model-based control of robots, accurate dynamic parameters, e.g., inertial, friction parameters, are essential. These parameters can be obtained in advance from CAD models or through offline system identification. However, there can quite often be the case where these dynamic parameters change from the initial values during the task and need to be re-identified in an online manner. Especially for many of the recent robot arms which come in lightweight designs, loading moderately heavy object or tool at the end-effector can largely change the dynamic behavior of the robot.

One important issue that arises in online parameter estimation for model-based control is on the stability of the system. For instance, completely modularizing the parameter estimator from the model-based controller does not in general guarantee the

system's stability. Adaptive control which exactly addresses this type of problem has a long history in robotics. The two most important classes of globally convergent (stable) adaptive controllers are *adaptive inverse dynamics control* (or *adaptive computed torque control*), proposed by Craig et al [71], and *passivity-based adaptive control*, proposed by Slotine and Li [72]. The distinguishing feature of these two methods compared to previous works is that they are proven to be globally convergent without relying on any linear approximation of the dynamics; instead, they utilize the explicit linear decomposition of the inertial parameters in the robot dynamics.

An underlying assumption of adaptive inverse dynamics control [71] and other similar approaches [73], [74], [75] in this category is the requirement of a uniformly positive-definite estimated mass matrix, since the inverse of the estimated mass matrix is explicitly used in the parameter adaptation law. An appealing advantage of passivity-based adaptive control methods [72], [46] is that uniform positive-definiteness of the estimated mass matrix is not required in their implementation.

The requirement that the estimated mass matrix be positive-definite is directly related to the physical consistency of the estimated inertial parameters of a robot. That is, it can be shown that the necessary condition for physical consistency of the inertial parameters (i.e., all links have positive mass, and positive-definite rotational inertia matrix) guarantees that the corresponding robot mass matrix is positive-definite. Exploiting this property together with the convex characterization of the set of physically consistent inertial parameters, projection-based algorithms [76], or parameter resetting algorithms, have been applied for adaptive robot controllers; these essentially switch the parameter update law on the boundary of the feasible set to enforce the estimated parameters to be feasible, which in turn produces a nonsmooth adaptive control when the parameter values lie on the user-specified boundary of the feasible set [71], [77], [78], [79].

Another critical liability that direct adaptive robot control methods based on [71] and

[72] have in common is the requirement that users must choose a valid initial adaptation gain matrix Γ . This can be a time-consuming process requiring repeated trial and error, as the number of constant adaptation gains varies as the square of the number of adaptation parameters, which is problem-dependent (e.g., adapting the entire set, or a subset of the link inertial parameters, or only the end-effector link for compensating unknown payloads).

In this Chapter we propose a new adaptive control law that mitigates many of the above difficulties, in the form of a parameter adaptation law that guarantees physical consistency of the estimated parameters in a smooth manner without relying on any parameter projection or resetting procedure. Whereas existing adaptation laws can be viewed as a gradient update law on a flat Euclidean space with constant metric Γ , our method can be viewed as a natural gradient-like update law on a curved space endowed with a Riemannian metric [80]. It does not require additional computation beyond existing methods, and is directly applicable to any globally convergent Lyapunov-based adaptive controller that exploits the linear decomposition of the inertial parameters in the dynamics.

Our geometric adaptation law, which like [80] is coordinate-invariant and further respects the underlying Riemannian geometry of the space of inertial parameters, considerably reduces the degree to which engineering choices must be made in the constant adaptation gain; only the choice of a scalar constant gain γ is required for adjusting the speed of adaptation, without regard to the number of parameters for adaptation.

Toward deriving our geometric adaptation law, we first extend the previous framework of characterizing Riemannian manifold structure on the space of physically consistent inertial parameters to general mechanical parameters defined on an arbitrary convex set; the affine-invariant Riemannian geometry defined on the inertial parameter space

[81, 82] can be regarded as a special case. More specifically, if one can impose a particular convex barrier function that induces a Hessian manifold structure [83] on the convex set, then a Lyapunov-stable parameter update law can be naturally derived using the associated Bregman divergence [84] as a Lyapunov function candidate. We also show how non-smooth projection strategies, so-called robust adaptation laws, can be naturally defined within our geometric framework. Our methods and claims are validated via extensive simulation experiments involving a seven-dof robot manipulator and a real experiment on a seven-dof AMBIDEX cable-driven robot manipulator.

5.2 Preliminaries

In this Section, we briefly revisit the two main classes of Lyapunov-based globally convergent adaptive controllers developed for robot manipulators in [71] and [72]. The dynamic equations for a general n -dof open chain manipulator are of the form

$$M(q, \Phi)\ddot{q} + C(q, \dot{q}, \Phi)\dot{q} + g(q, \Phi) + k(\dot{q}, \mu) = u, \quad (5.2.1)$$

where $q \in \mathbb{R}^n$ is the vector of joint angles, $M(\cdot) \in \mathbb{R}^{n \times n}$, $C(\cdot) \in \mathbb{R}^{n \times n}$, and $g(\cdot) \in \mathbb{R}^n$ respectively denote the mass matrix, Coriolis matrix, and the gravitational force vector, $u \in \mathbb{R}^n$ is the motor torque input, $\Phi = [\phi_1^T, \dots, \phi_n^T]^T \in \mathbb{R}^{10n}$ is the complete set of inertial parameters for the n links. $k(\cdot)$ denote the joint friction force, and the vector μ represents the friction parameters. For instance, joint Coulomb-viscous friction model is given by

$$k(\dot{q}, \mu) = \text{diag}(\text{sgn}(\dot{q}))\mu_c + \text{diag}(\dot{q})\mu_v, \quad (5.2.2)$$

where $\mu = [\mu_c^T, \mu_v^T]^T \in \mathbb{R}^{2n}$ with $\mu_c \in \mathbb{R}_+^n$ and $\mu_v \in \mathbb{R}_+^n$ respectively the positive values of joint Coulomb and viscous friction parameters. We would also often use the

following abbreviated form,

$$M(q, p)\ddot{q} + b(q, \dot{q}, p) = u, \quad (5.2.3)$$

where $b(\cdot) \in \mathbb{R}^n$ represents forces or torques arising from Coriolis, gravitational and frictional forces, and parameter vector $p = [\Phi^T, \mu^T]^T$ is simply the collection of inertial and friction parameters.

The technical objective of adaptive control is to design a control input u that ensures global convergence of the trajectory tracking error in the presence of model uncertainties. As the term “adaptive” implies, time-varying estimates of the true model parameters are obtained with the trajectory tracking controller, distinct from the class of robust controllers that make use of fixed parameter estimates with known uncertainty bound. Robust extensions of the adaptive control laws, so-called *robust adaptive control* [76], are to be discussed in Section 5.4.2.

5.2.1 Adaptive Control of Robot Manipulators

5.2.1.1 Adaptive Computed Torque Control

The control input for adaptive computed torque control proposed by [71], also referred to as adaptive inverse dynamics, is given by

$$u = M(q, \hat{p})\{\ddot{q}_d - K_v\dot{\tilde{q}} - K_p\tilde{q}\} + b(q, \dot{q}, \hat{p}), \quad (5.2.4)$$

where $K_p \in \mathbb{R}^{n \times n}$ and $K_v \in \mathbb{R}^{n \times n}$ are diagonal matrices of positive gains, q_d is the given reference trajectory, $\tilde{q} = q - q_d$, and $\hat{p} = [\hat{\Phi}^T, \hat{\mu}^T]^T$ is the estimate of the true dynamic parameter vector p , whose update law is clarified later. Using the property that dynamic parameters can generally be factored linearly from the dynamic equations, the closed loop dynamics becomes

$$\ddot{\tilde{q}} + K_v\dot{\tilde{q}} + K_p\tilde{q} = M(q, \hat{\Phi})^{-1}Y(q, \dot{q}, \ddot{q})\tilde{p}, \quad (5.2.5)$$

where $Y \in \mathbb{R}^{n \times 10n}$ is the regressor function that satisfies

$$M(q, p)\ddot{q} + b(q, \dot{q}, p) = Y(q, \dot{q}, \ddot{q})p, \quad (5.2.6)$$

and $\tilde{p} = \hat{p} - p$ is the error vector of the parameter estimate. The state space formulation of (5.2.5) with augmented state vector defined as $e = [\tilde{q}^T, \dot{\tilde{q}}^T]^T$ is given by

$$\dot{e} = Ae + BM(q, \hat{p})^{-1}Y(q, \dot{q}, \ddot{q})\tilde{p}, \quad (5.2.7)$$

where $A = \begin{bmatrix} 0 & \mathbb{I}_n \\ -K_p & -K_v \end{bmatrix} \in \mathbb{R}^{2n \times 2n}$ is a Hurwitz matrix and $B = \begin{bmatrix} 0 & \mathbb{I}_n \end{bmatrix}^T \in \mathbb{R}^{2n \times n}$ (we denote \mathbb{I}_n the $n \times n$ identity matrix). One may then choose some constant symmetric $2n \times 2n$ positive-definite matrix Q , and let P be the unique symmetric positive-definite matrix satisfying the Lyapunov equation,

$$A^T P + PA = -Q. \quad (5.2.8)$$

The Lyapunov function candidate V is defined as follows:

$$V = e^T P e + \tilde{p}^T \Gamma \tilde{p},$$

where Γ is a constant symmetric positive-definite matrix. The time derivative of the Lyapunov function candidate \dot{V} is given by

$$\dot{V} = -e^T Q e + 2\tilde{p}^T \{Y(q, \dot{q}, \ddot{q})^T M(q, \hat{p})^{-1} B^T P e + \Gamma \dot{\tilde{p}}\},$$

which makes use of the fact that $\dot{\tilde{p}} = \dot{\hat{p}}$ since p is constant. Defining the parameter update law (or adaptation law) as

$$\dot{\hat{p}} = -\Gamma^{-1} Y(q, \dot{q}, \ddot{q})^T M(q, \hat{p})^{-1} B^T P e, \quad (5.2.9)$$

we have $\dot{V} = -e^T Q e \leq 0$. From the stability analysis provided in [71] it follows that the position tracking error e converges to zero asymptotically, and the parameter

estimate error $\tilde{p}^T \Gamma \tilde{p}$ remains bounded, provided that the estimated mass matrix $M(q, \hat{p})$ is uniformly bounded and invertible.

In fact, $M(q, \hat{p})$ must be invertible in order to implement the adaptation law (5.2.9), but the existence of $M(q, \hat{p})^{-1}$ is not rigorously justified. In [71], the authors provide a switching scheme to reset the parameters to always reside inside some feasible bound, defined by a set of linear inequality constraints on \hat{p} that lead to sufficient conditions guaranteeing the boundedness and invertibility (or equivalently, positive-definiteness) of $M(q, \hat{p})$. Another important issue in implementing the adaptation law (5.2.9) is the requirement of the joint acceleration feedback term \ddot{q} , which is difficult to obtain accurately from position and velocity measurements. The adaptive controller presented below proposed by Slotine and Li [72] removes both of these difficulties.

5.2.1.2 Passivity-based Adaptive Control

The control input for passivity-based adaptive control proposed by [72] is given as

$$\begin{aligned} u &= M(q, \hat{\Phi})a + C(q, \dot{q}, \hat{\Phi})v + g(q, \hat{\Phi}) \\ &\quad + k(v, \hat{\mu}) - Kr, \\ &= Y(q, \dot{q}, a, v)\hat{p} - Kr \end{aligned} \tag{5.2.10}$$

where the vectors $v, a, r \in \mathbb{R}^n$ are defined as

$$v = \dot{q}_d - \Lambda \tilde{q}, \tag{5.2.11}$$

$$a = \dot{v} = \ddot{q}_d - \Lambda \dot{\tilde{q}}, \tag{5.2.12}$$

$$r = \dot{q} - v = \dot{\tilde{q}} + \Lambda \tilde{q}, \tag{5.2.13}$$

and K and Λ are symmetric positive definite matrices of constant gains. The closed-loop dynamics is given by

$$M(q, \Phi)\dot{r} + C(q, \dot{q}, \Phi)r + Kr + [k(\dot{q}, \mu) - k(v, \mu)] = Y(q, \dot{q}, a, v)\tilde{p}.$$

The following Lyapunov function candidate is introduced:

$$V = \frac{1}{2}r^T M(q, \Phi)r + \frac{1}{2}\tilde{p}^T \Gamma \tilde{p}, \quad (5.2.14)$$

where Γ as before is set to be a constant symmetric positive-definite matrix. Using the fact that $\dot{M} - 2C$ as constructed is skew-symmetric [72], \dot{V} reduces to

$$\dot{V} \leq -r^T K r + \tilde{p}^T \{\Gamma \dot{\tilde{p}} + Y(q, \dot{q}, a, v)^T r\},$$

under the mild condition that the joint friction model $k_i(\dot{q}_i, \mu)$ is a monotonically increasing function of \dot{q}_i . Choosing the parameter update law as

$$\dot{\tilde{p}} = -\Gamma^{-1} Y(q, \dot{q}, a, v)^T r, \quad (5.2.15)$$

we have $\dot{V} = -r^T K r \leq 0$. It can also be shown that the position tracking error e converges to zero asymptotically from the convergence of r , and the parameter estimate error $\tilde{p}^T \Gamma \tilde{p}$ remains bounded. Note that neither the acceleration measurements \ddot{q} nor the inverse of the mass matrix $M(q, \hat{\Phi})$ is required in the control input (5.2.10) and the parameter adaptation law (5.2.15).

Remark 5.1. As noted in [46], a practical choice for the gain matrices K , Λ is given by $K = \lambda \cdot M(q, \hat{\Phi})$, and $\Lambda = \lambda \cdot \mathbb{I}_n$ for some scalar constant λ . The above choice implies that higher gain values are used for joints with higher inertias. It is further shown that even with such a time-varying choice of K , global convergence can be guaranteed with only a slight modification of a in the regressor function Y in the adaptation law (5.2.15), to $a - \lambda r$.

5.3 Barrier-Hessian Manifolds

In this Section, we try to extend the previous framework of characterizing geometric structure of the physically feasible set of inertial parameters (in Section 3.3) to general mechanical parameters defined on an arbitrary convex subset in \mathbb{R}^N . The contents described here are the key ingredients for the derivation of our main contribution: geometric parameter adaptation laws presented in Section 5.4.

To begin with, recall that evaluating numerical values of mechanical parameters require a specific coordinate choice and the coordinate choice includes unit system, and body-attached frames to evaluate spatial quantities, e.g., moments of inertia, center of mass (CoM), stiffness matrix, etc. One useful observation here is that the change of these coordinates can always be defined in the form of affine transformation of the parameters. Mathematically, such a space can be characterized as an affine manifold.

Definition 5.3.1. An **affine manifold** is a manifold which only admits affine coordinate systems; that is, coordinate change is always an affine function.

There is a particularly convenient way of imposing a Riemannian manifold structure to the affine manifold which introduces to the definition of a Hessian manifold.

Definition 5.3.2. Hessian manifold (\mathcal{M}, h) is an affine manifold \mathcal{M} with Riemannian metric g given by the Hessian of a strictly convex, twice differentiable function $h : \mathcal{M} \rightarrow \mathbb{R}$, i.e., in local affine coordinate $p = (p^1, \dots, p^N) \in \mathbb{R}^N$ the coordinate representation of the Riemannian metric $g(p) \in \mathbb{R}^{N \times N}$ is given by

$$[g(p)]_{i,j} \triangleq \frac{\partial^2 h(p)}{\partial p^i \partial p^j}, \quad \text{or} \quad (5.3.16)$$

$$g(p) \triangleq \nabla^2 h(p) \succ 0. \quad (5.3.17)$$

Remark 5.2. Note however that Hessian can not be generally well-defined as a valid

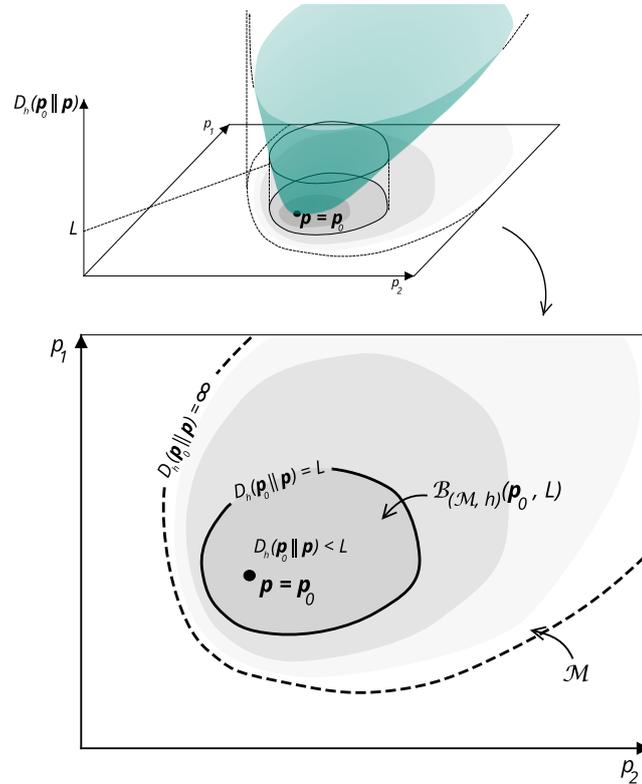


Figure 5.1: Example of a 2-dimensional barrier-Hessian manifold (\mathcal{M}, h) .

coordinate-invariant Riemannian metric for a manifold which admits nonlinear coordinate transformations.

Bregman divergence (2.2.2) associated with function h can then be naturally defined on a Hessian manifold (\mathcal{M}, h) . By the property (2.2.13), the following second-order approximation holds between the Bregman divergence and the Riemannian metric (5.3.17):

$$d_h(p||p + dp) = \frac{1}{2} dp^T g(p) dp + o(\|dp\|^2) \quad (5.3.18)$$

Lastly, we add a few technical conditions to the original definition of Hessian manifold to define what we would call a barrier-Hessian manifold (see Figure 1).

Definition 5.3.3. A **barrier-Hessian manifold** is a Hessian manifold (\mathcal{M}, h) that satisfies the following conditions:

- (i) \mathcal{M} can be identified as an N -dimensional convex open subset in \mathbb{R}^N
- (ii) [84, pp. 201] For every $p_0 \in \mathcal{M}$ and positive scalar L , the set

$$\mathcal{B}_{(\mathcal{M}, h)}(p_0, L) \triangleq \{p \in \mathcal{M} : d_h(p_0 \| p) \leq L\} \quad (5.3.19)$$

is compact, i.e., closed and bounded in \mathbb{R}^N . Or equivalently, $d_h(p_0 \| p)$ approaches to infinity as p approaches to the boundary of \mathcal{M} .

The conditions above essentially reduce the degree to which the choice of function h can be made, yet will allow us to construct parameter adaptation laws with desirable properties in Section 5.4. Below we characterize the space of rigid body inertial parameters and joint friction/stiffness parameters as explicit examples of affine manifolds with barrier-Hessian manifold structures.

5.3.1 Rigid Body Inertial Parameters

Recall that the manifold \mathcal{M}_I of physically consistent inertial parameters of a single rigid body is given by

$$\mathcal{M}_I \simeq \{\phi_b \in \mathbb{R}^{10} : P(\phi_b) \succ 0\}. \quad (5.3.20)$$

where the elements in \mathcal{M}_I can be identified with different choices of coordinates $\{b\}$.

We first claim that \mathcal{M}_I is an affine manifold. From the relation (3.3.24), the pseudo-inertia matrix representation of inertial parameters always admits the linear coordinate transformation rule $\Psi_{ab} : \mathcal{P}(4) \rightarrow \mathcal{P}(4)$ of the form,

$$P_a = \Psi_{ab}(P_b) \triangleq S_{ab} P_b S_{ab}^T, \quad (5.3.21)$$

for some constant nonsingular matrix $S_{ab} \in \mathbb{R}^{4 \times 4}$. By the linearity of the mappings P and Ψ_{ab} , the vector representation ϕ hence always admits the linear transformation rule $\psi_{ab} = P \circ \Psi_{ab} \circ P^{-1} : \mathbb{R}^{10} \rightarrow \mathbb{R}^{10}$, i.e.,

$$\phi_a = \psi_{ab}(\phi_b) = P^{-1}(\Psi_{ab}(P(\phi_b))). \quad (5.3.22)$$

Therefore, the manifold \mathcal{M}_I (5.3.20) is an affine manifold. Now, we propose that \mathcal{M}_I exhibits a particular barrier-Hessian manifold structure.

Proposition 5.1. *Affine manifold \mathcal{M}_I (5.3.20) of inertial parameters admits barrier-Hessian manifold structure (\mathcal{M}_I, h_I) with*

$$h_I(\phi) \triangleq -\log \det(P(\phi)). \quad (5.3.23)$$

Hessian of the strictly convex function h_I induces the Riemannian metric $g_I \in \mathbb{R}^{10 \times 10}$ as

$$[g_I(\phi)]_{i,j} \triangleq \text{tr}(P(\phi)^{-1}P(e_i)P(\phi)^{-1}P(e_j)) \quad (5.3.24)$$

($e_i \in \mathbb{R}^{10}$ is a vector having 1 at the i -th element and zero for the remaining elements) and the Bregman divergence of h_I is given by

$$\begin{aligned} d_{h_I}(\phi \parallel \psi) &= -\log \det(P(\psi)^{-1}P(\phi)) \\ &\quad + \text{tr}(P(\psi)^{-1}P(\phi)) - 4 \end{aligned} \quad (5.3.25)$$

Proof is given in the Appendix A.5.

Remark 5.3. *It is noted that the expression of $h_I : \mathcal{M}_I \rightarrow \mathbb{R}$ given in (5.3.23) is yet well-defined in terms of coordinate invariance; for arbitrary two coordinates $\{b\}$ and $\{a\}$ the values of h_I deviate by a constant, i.e.,*

$$\begin{aligned} h_I(\phi_b) &= -\log \det(S_{ab}P(\phi_a)S_{ab}^T) \\ &= -\log \det(P(\phi_a)) - 2 \log \det(S_{ab}) \\ &= h_I(\phi_a) + \text{constant}. \end{aligned}$$

Nonetheless, both the Riemannian metric (5.3.24) and the Bregman divergence D_{h_I} (5.3.25) admit well-defined coordinate-free expressions, as Hessian operator and Bregman divergence is invariant to the constant additions to the function.

It is noted that further useful convex constraints can be appended to the characterization of the inertial parameter set \mathcal{M}_I . For instance, bounding box constraint on CoM, min-max constraint on the mass, bounding ellipsoid constraint on the support of the underlying mass density (3.2.12) [28] can all be expressed as linear inequality constraints on ϕ . Let us denote the set \mathcal{M}_I^+ of physically consistent inertial parameters with additional m number of linear inequality constraints as

$$\mathcal{M}_I^+ \simeq \{\phi \in \mathcal{M}_I : a_k^T \phi - b_k > 0, \quad k = 1, \dots, m\}, \quad (5.3.26)$$

which is also an affine manifold by the same argument as above.

Proposition 5.2. *The barrier-Hessian manifold structure (\mathcal{M}_I^+, h_I^+) can be defined based on the previous (\mathcal{M}_I, h_I) as,*

$$h_I^+(\phi) \triangleq h_I(\phi) - \alpha \cdot \sum_{k=1}^m \log(a_k^T \phi - b_k) \quad (5.3.27)$$

for some positive scalar α ; Hessian of the function h_I^+ induces the Riemannian metric $g_I^+ \in \mathbb{R}^{10 \times 10}$ as

$$g_I^+(\phi) \triangleq g_I(\phi) + \alpha \cdot \sum_{k=1}^m \frac{1}{(a_k^T \phi - b_k)^2} a_k a_k^T \succ 0. \quad (5.3.28)$$

and the Bregman divergence of h_I^+ is given by

$$\begin{aligned} D_{h_I^+}(\phi \parallel \psi) &= D_{h_I}(\phi \parallel \psi) \\ &+ \alpha \cdot \sum_{i=1}^m \left[\log \frac{a_i^T \psi - b_i}{a_i^T \phi - b_i} + \frac{a_i^T \phi - b_i}{a_i^T \psi - b_i} - 1 \right]. \end{aligned} \quad (5.3.29)$$

Proof is given in the Appendix A.6.

Remark 5.4. *Similar to Remark 5.3, the Riemannian metric (5.3.28) and the Bregman divergence $D_{h_I^+}$ (5.3.29) are invariant to the coordinate choice and also arbitrary scaling of the specification of the linear inequality constraints, i.e., substitutions $a_k \leftarrow c \cdot a_k$ and $b_k \leftarrow c \cdot b_k$ for any scalar $c > 0$ yield the same form of Riemannian metric and Bregman divergence. This allows the choice of dimensionless value α to be free of scaling and coordinate choices.*

5.3.2 Joint Friction/Stiffness Parameters

One-dimensional joint friction or elastic force is typically modeled with a set of positive scalar-valued parameters. For instance, the joint Coulomb-viscous friction model (5.2.2) can be parametrized with two positive scalars $(\mu_c, \mu_v) \in \mathbb{R}_+ \times \mathbb{R}_+$, and simple linear (torsional) spring model often used to take into account the joint elasticity is also parametrized by a scalar spring constant $k \in \mathbb{R}_+$. Hence, the manifold \mathcal{M}_J of feasible joint friction/stiffness parameters can be given by

$$\mathcal{M}_J \simeq \{\mu \in \mathbb{R}^{n_J} : \mu_i > 0, i = 1, \dots, n_J\}. \quad (5.3.30)$$

\mathcal{M}_J is clearly an affine manifold whose elements can be identified with different scaling of each of the parameters.

Proposition 5.3. *Affine manifold \mathcal{M}_J (5.3.30) admits barrier-Hessian manifold structure (\mathcal{M}_J, h_J) with*

$$h_J(\mu) = - \sum_{i=1}^{n_J} \log \mu_i, \quad (5.3.31)$$

where the Hessian of the convex function h_J induces the Riemannian metric $g_J \in \mathbb{R}^{n_J \times n_J}$ as

$$g_J(\mu) = \text{diag} \left(\frac{1}{\mu_1^2}, \dots, \frac{1}{\mu_{n_J}^2} \right), \quad (5.3.32)$$

and the Bregman divergence of h_J is given by

$$D_{h_J}(\mu||\nu) = \sum_{i=1}^{n_J} \left[\log \frac{\nu_i}{\mu_i} + \frac{\mu_i}{\nu_i} - 1 \right] \quad (5.3.33)$$

Also, if each of the parameters is to be bounded on the finite interval, i.e., $\mu_i \in (l_i, r_i)$, denote

$$\mathcal{M}_J^+ \simeq \{\mu \in \mathbb{R}^{n_J} : 0 < l_i < \mu_i < r_i, i = 1, \dots, n_J\}. \quad (5.3.34)$$

Proposition 5.4. *The barrier-Hessian manifold structure (\mathcal{M}_J^+, h_J^+) can be defined with,*

$$h_J^+(\mu) = - \sum_{i=1}^{n_J} [\log(\mu_i - l_i) + \log(r_i - \mu_i)], \quad (5.3.35)$$

where the Hessian of the function h_J^+ induces the Riemannian metric $g_J^+ \in \mathbb{R}^{n_J \times n_J}$ as

$$g_J^+(\mu) = \text{diag} \left(\frac{1}{(\mu_1 - l_1)^2}, \dots, \frac{1}{(\mu_{n_J} - l_{n_J})^2} \right) + \text{diag} \left(\frac{1}{(\mu_1 - r_1)^2}, \dots, \frac{1}{(\mu_{n_J} - r_{n_J})^2} \right), \quad (5.3.36)$$

and the Bregman divergence of h_J^+ is given by

$$D_{h_J^+}(\mu||\nu) = \sum_{i=1}^{n_J} \left[\log \frac{\nu_i - l_i}{\mu_i - l_i} + \frac{\mu_i - l_i}{\nu_i - l_i} - 1 \right] + \sum_{i=1}^{n_J} \left[\log \frac{r_i - \nu_i}{r_i - \mu_i} + \frac{r_i - \mu_i}{r_i - \nu_i} - 1 \right]. \quad (5.3.37)$$

The proofs of Proposition 5.3 and 5.4 are similar to those of Proposition 5.1 and Proposition 5.2.

We close this section by showing how one can directly define a barrier-Hessian manifold structure on the product space of two affine manifolds already having barrier-Hessian manifold structures, e.g., (\mathcal{M}_I, h_I) and (\mathcal{M}_J, h_J) .

Proposition 5.5. *Given two barrier-Hessian manifolds (\mathcal{M}_x, h_x) and (\mathcal{M}_y, h_y) , the barrier-Hessian manifold structure (\mathcal{M}, h) on the product space $\mathcal{M} = \mathcal{M}_x \times \mathcal{M}_y$ can be define with*

$$h(p) = w_x \cdot h_x(x) + w_y \cdot h_y(y), \quad (5.3.38)$$

for some constant positive scalars w_x , w_y , and $p = (x, y) \in \mathcal{M} = \mathcal{M}_1 \times \mathcal{M}_2$. Then, the Riemannian metric g and the Bregman divergence D_h are given by

$$g(p) = \text{diag}(w_x \cdot g_x(\phi), w_y \cdot g_y(\mu)) \quad (5.3.39)$$

$$d_h(p_1 \| p_2) = w_x \cdot d_{h_x}(x_1 \| x_2) + w_y \cdot d_{h_y}(y_1 \| y_2). \quad (5.3.40)$$

Proof is given in the Appendix A.7. By recursion, barrier-Hessian manifold structure (\mathcal{M}, h) with Riemannian metric g on the product space of arbitrary n number of barrier-Hessian manifolds, (\mathcal{M}_i, h_i) , $i = 1, \dots, n$ can be given as follows:

$$\mathcal{M} = \mathcal{M}_1 \times \dots \times \mathcal{M}_n, \quad (5.3.41)$$

$$h(p) = \sum_{i=1}^n w_i \cdot h_i(x_i), \quad (5.3.42)$$

$$g(p) = \text{diag}(w_1 \cdot g_1(x_1), \dots, w_n \cdot g_n(x_n)), \quad (5.3.43)$$

$$d_h(p_1 \| p_2) = \sum_{i=1}^n w_i \cdot d_{h_i}(x_{i_1} \| x_{i_2}), \quad (5.3.44)$$

where x_i are the coordinate values of \mathcal{M}_i . Unless otherwise noted, we would normally set $w_i = 1$ for all i for the sake of brevity.

5.4 Geometric Parameter Update Laws

Lyapunov stability analysis provides a way of assessing the stability of a closed-loop system by a suitable choice of Lyapunov function. In some cases what at first appears

to be a natural Lyapunov function candidate can complicate the design of a stabilizing control law. Particularly for a large class of mechanical systems including robot manipulators, a branch of passivity-based control methods that exploits the skew-symmetry of $\dot{M} - 2C$ have been developed for a class of energy-like Lyapunov functions, e.g., a system's kinetic and elastic energy terms as reflected in (5.2.14). Such methods allow for the closed-loop system to inherit some of the intrinsic physical properties of the original system [85], rather than replacing entirely the original dynamics with that of a virtual spring-damper system via feedback linearization as done in computed-torque control. As described in the previous section, Slotine and Li have shown that such physically motivated control designs can also be successfully applied to adaptive manipulator control.

The general class of globally convergent adaptive robot control laws as discussed in Section 5.2.1 consider a Lyapunov function candidate V that is the sum of a tracking error term V_t and a parameter error term V_p :

$$V = \underbrace{V_t(q, \dot{q}, q_d, \dot{q}_d, p)}_{\text{(tracking error)}} + \underbrace{V_p(p, \hat{p})}_{\text{(parameter error)}} . \quad (5.4.45)$$

Here, as opposed to the choice of tracking error term V_t , the natural choice of parameter error term V_p is often overlooked; a quadratic parameter error of the form $\tilde{p}^T \Gamma \tilde{p}$ with constant positive-definite matrix Γ is often considered. As the inverse of Γ serves as an adaptation gain in (5.2.9), (5.2.15), selecting an appropriate choice of Γ can be a time-consuming and arduous process. Moreover, as discussed in Chapter 3, different choices of metric on the parameters, for example the standard Euclidean metric corresponding to $\Gamma = \gamma \mathbb{I}$, can lead to fatal scaling problems and physically inconsistent estimators, especially when the excitation trajectories fail to be persistently exciting. Such problems are pervasive in adaptive control, where only limited trajectory data is available for online adaptation purposes.

There have been several approaches or modifications toward more robust parameter adaptation under the framework of adaptive computed-torque control and passivity based adaptive control. For the case of adaptive computed-torque control, several projection or resetting based modifications of the adaptation law have been proposed to ensure the estimated inertial parameters are physically consistent [71, 77, 76, 78, 79], by sufficiently guaranteeing the uniform positive-definiteness of estimated mass matrix $M(q, \hat{\Phi})$. We believe that such switching-based schemes are more often than not ad hoc remedies to the issue of physical inconsistency: with a poor choice of Γ , the estimated parameters are highly prone to converge to the boundary of a pre-defined feasible set, which leads to a trivial adaptation of the parameters over a large portion of the operation time. On the other hand, Slotine and Li have suggested indirect [77] or composite versions [46] of their passivity-based method. These take into account additional filtered torque prediction errors in a least-squares sense, which essentially updates the matrix Γ to a more well-conditioned value over the time.

However, to the best of our knowledge, none of the adaptation laws so far respect the physical consistency of the estimated parameters in an intrinsic and smooth manner. In the following sections, we argue that these goals can in fact be achieved by consideration of the natural and coordinate-invariant choice of a smooth parameter error term V_p . This in turn significantly reduces the human burden of having to select an excessive number of adaptation gains compared to existing methods.

5.4.1 Geometric Adaptation Law

To begin with, observe that the time derivative of the tracking error term V_t is of the form

$$\dot{V}_t \leq \tilde{p}^T b, \quad (5.4.46)$$

where $b = Y^T \hat{M}^{-1} B^T P e$ for adaptive computed torque control, and $b = Y^T r$ for the passivity-based adaptive control. In order to ensure that time derivative of $V = V_t + V_p$ is negative, the time derivative of V_p must be of the form

$$\dot{V}_p(p, \hat{p}) = \tilde{p}^T w(\dot{\hat{p}}, \hat{p}), \quad (5.4.47)$$

with the adaptation law $w(\dot{\hat{p}}, \hat{p}) = -b$. This eventually enforces the negativity of \dot{V} , i.e., $\dot{V} = \dot{V}_t + \dot{V}_p \leq \tilde{p}^T (b + w(\dot{\hat{p}}, \hat{p})) = 0$. To illustrate, when V_p is set to be the conventional quadratic error, i.e., $V_p = \frac{1}{2} \tilde{p}^T \Gamma \tilde{p}$, then $w(\dot{\hat{p}}, \hat{p}) = \Gamma \dot{\hat{p}}$ which yields the conventional parameter adaptation laws of the form

$$\dot{\hat{p}} = -\Gamma^{-1} b. \quad (5.4.48)$$

It is important here to note that w should be a function of only the estimated parameters \hat{p} and their time derivatives $\dot{\hat{p}}$, and not include the true parameter p , since it is not known a priori, but rather a desired value we wish to estimate. We now claim that the Bregman divergence-based distance measure (2.2.2) allows for a valid Lyapunov function candidate for V_p whose time derivative is of the form (5.4.47).

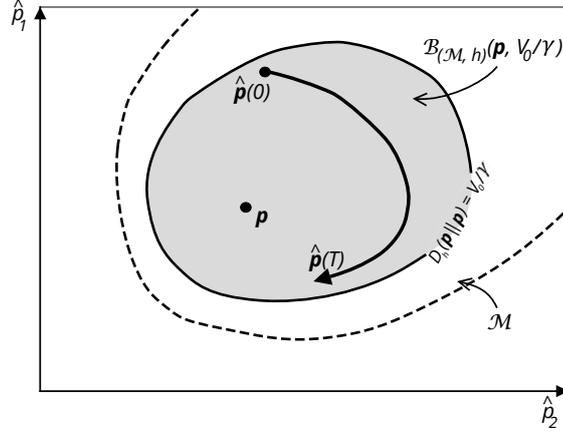
Proposition 5.6. *Given a barrier Hessian manifold structure (\mathcal{M}, h) on the parameter space, a function $V_p : \mathcal{M} \times \mathcal{M} \rightarrow \mathbb{R}$ defined with the Bregman divergence,*

$$V_p(p, \hat{p}) = \gamma d_h(p, \hat{p}), \quad (5.4.49)$$

for constant positive scalar gain γ is a valid Lyapunov function candidate. Its time derivative satisfies the form given in (5.4.47):

$$\dot{V}_p(p, \hat{p}) = \gamma \tilde{p}^T \underbrace{\nabla^2 h(\hat{p})}_{w(\dot{\hat{p}}, \hat{p})} \dot{\hat{p}}. \quad (5.4.50)$$

Proof is given in the Appendix A.8.

Figure 5.2: Geometric adaptation on (\mathcal{M}, h)

Remark 5.5. With V_p chosen as geodesic distance of the Riemannian metric, its time derivative cannot be generally factored in the form (5.4.47).

Based on the choice of V_p as in (5.4.49), we now propose a novel adaptation law, which we refer to as a *geometric adaptation law*, applicable to any Lyapunov-based adaptive control law in which the time derivative of the tracking error-related Lyapunov function candidate V_t is expressible in the form (5.4.46).

Proposition 5.7. (Geometric Adaptation Law) Let $g = \nabla^2 h$ be the Riemannian metric of a barrier Hessian manifold structure (\mathcal{M}, h) . Given a control law that results in the time derivative of V_t of the form (5.4.46), the following adaptation law

$$\dot{\hat{p}} = -\frac{1}{\gamma} [g(\hat{p})]^{-1} \cdot b \quad (5.4.51)$$

guarantees the asymptotic convergence of the tracking error to zero and bounded parameter error, i.e., for $V_0 = V|_{t=0}$

$$\hat{p}(t) \in \mathcal{B}_{(\mathcal{M}, h)}(p, V_0/\gamma) \subset \mathcal{M}, \quad \forall t \in [0, \infty) \quad (5.4.52)$$

in which the initial estimate $p(0)$ is chosen to be $p(0) \in \mathcal{M}$.

Proof is given in the Appendix A.9. Although the vectorized coordinate representation of the adaptation law (5.4.52) involves the matrix inverse of $g(\hat{p})$, we note that in practice this matrix inverse need not be implemented at all. To illustrate, suppose the barrier-Hessian manifold \mathcal{M} be the product space of inertial parameter manifold \mathcal{M}_I ($\dim(\mathcal{M}_I) = 10$) and joint friction parameter manifold \mathcal{M}_J ($\dim(\mathcal{M}_J) = n_J$). By Proposition 5.5, the Riemannian metric g is a block diagonal matrices of g_I and g_J . Therefore, the update rule (5.4.52) can be decomposed as

$$\dot{\hat{p}} = -\frac{1}{\gamma} \left([g_I(\hat{\phi})]^{-1} b_I + [g_J(\hat{\mu})]^{-1} b_J \right). \quad (5.4.53)$$

The metric $g_J(\cdot)$ (5.3.32) is a diagonal matrix for which the inverse is trivial to implement. For the first term $[g_I(\hat{\phi})]^{-1} b_I$, one can use the identity (B.1.10) in the Appendix B.1 to avoid the matrix inverse computation, i.e., $[g_I(\hat{\phi})]^{-1} b_I = 2 \cdot P^{-1}(P(\hat{\phi})X(b_I)P(\hat{\phi}))$. More details on the general implementation of $[g(\hat{p})]^{-1} b$ is given in the Appendix B.4.

As opposed to an arbitrary constant gain matrix Γ in conventional approaches (5.4.48), it can be observed that a state-dependent, time-varying choice $\Gamma(\hat{p}) = g(\hat{p})$ is admissible. As a by-product, one only has to determine a single tuning variable γ , which is essentially the adaptation speed.

From an optimization perspective, the existing parameter update rule (5.4.48) can be viewed as a steepest gradient descent-like update [2, page 414] in flat Euclidean space with constant metric Γ . On the other hand, our method is analogous to the natural gradient descent method, which generalizes the steepest descent method to general curved Riemannian manifolds. The natural gradient is not only coordinate-invariant, but also known to provide a more efficient direction of descent in terms of the metric predefined on the space [80]. Moreover, the resulting natural gradient-like realization of the

adaptation law actually turns out to guarantee physical consistency of the adapted parameters, which we proved using the Bregman divergence in a Lyapunov-based analysis rather than by using the exact Riemannian distance metric.

Remark 5.6. *In principle, tracking error in the passivity-based direct adaptive controllers [72] is guaranteed to asymptotically converge to zero with arbitrary constant choice of gain matrix Γ without regarding to the physical consistency of the estimated parameters. However, as will be shown in the following section, the transient performance of a direct adaptive controller can in fact be largely improved by the proposed natural choice of $\Gamma(\hat{p})$ that also ensures the physical consistency of the estimated parameters.*

5.4.2 Geometric Projection Law

In this Section, parameter projection laws for robust adaptive controllers [76] conventionally defined on the Euclidean space are generalized to the manifold \mathcal{M} with Hessian manifold structure.

Let the convex constraint be defined with $c(p) \leq 0$ for some convex function $c : \mathbb{R}^N \rightarrow \mathbb{R}$, i.e., define

$$\mathcal{C} = \{p \in \mathcal{M} : c(p) \leq 0\}. \quad (5.4.54)$$

Technical objective of robust adaptation is to update the parameters $\hat{p}(t)$ (on \mathcal{M}) while ensuring $\hat{p}(t) \in \mathcal{C}$ as well as the closed-loop system's stability for all t . This is achieved by switching the smooth adaptation law only when \hat{p} is about to cross the boundary of \mathcal{C} , i.e., $\hat{p}(t) \in \partial\mathcal{C}$. Specifically, the ensuing optimization problem formulation for projecting the instantaneous direction of update v to the tangent space of $\partial\mathcal{C}$ is given

by,

$$\min_{\hat{v}} \|\hat{v} - v\|_{\hat{p}}^2 \quad (5.4.55)$$

$$\text{s.t. } \nabla c(\hat{p})^T \hat{v} = 0 \quad (\hat{v} \in \mathcal{T}_{\hat{p}} \partial \mathcal{C}) \quad (5.4.56)$$

where $\|v\|_p \triangleq \sqrt{v^T g(p) v}$ and $\hat{p} \in \partial \mathcal{C}$. The closed-form solution to the above is given by, $\hat{v}^* = v - \frac{\nabla c^T v}{\nabla c^T w} \cdot w$ where $w = [g(\hat{p})]^{-1} \cdot \nabla c(\hat{p})$. Below we propose that this solution actually provides a valid parameter projection rule.

Proposition 5.8. (Geometric Projection Law on $\dot{\hat{p}}$) Define $v = -\frac{1}{\gamma} [g(\hat{p})]^{-1} \cdot b$ and $w = [g(\hat{p})]^{-1} \cdot \nabla c(\hat{p})$.

$$\dot{\hat{p}} = \begin{cases} v - \frac{\nabla c^T v}{\nabla c^T w} \cdot w & , \text{ if } \hat{p} \in \partial \mathcal{C} \text{ and } \nabla c^T v > 0 \\ v & , \text{ otherwise} \end{cases} \quad (5.4.57)$$

guarantees the asymptotic convergence of the tracking error to zero and bounded parameter error,

$$\hat{p}(t) \in \mathcal{B}_{(\mathcal{M}, h)}(p, V_0/\gamma) \cap \mathcal{C}, \quad \forall t \in [0, \infty) \quad (5.4.58)$$

in which the initial estimate $p(0)$ is chosen to be $p(0) \in \mathcal{M} \cap \mathcal{C}$.

Proof is given in the Appendix A.10. Meanwhile, if the boundary of \mathcal{C} is curved (i.e. if the function $c(p)$ is nonlinear), standard numerical integration of the projection law (5.4.57), e.g., Euler integration, can make \hat{p} to cross the boundary. To prevent this numerical problem, we propose an additional projection step that makes sure \hat{p} to always exactly evolve on \mathcal{C} while preserving the stability guarantees.

Proposition 5.9. (Geometric Projection Law on p) If it happens that $\hat{p}(t) \notin \mathcal{C}$, then the following instantaneous correction at t^+ ,

$$\hat{p}(t^+) = \text{Proj}_{\mathcal{C}}(\hat{p}(t)), \quad (5.4.59)$$

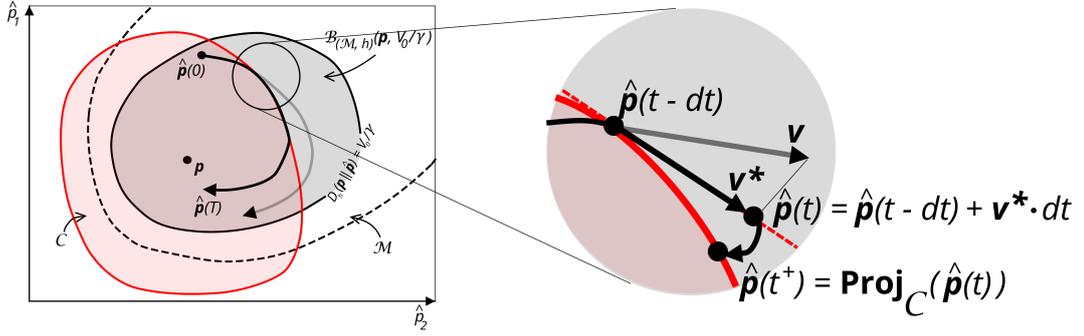


Figure 5.3: Geometric adaptation on (\mathcal{M}, h) with projection at the boundary of \mathcal{C}

where the projection operator $\text{Proj}_{\mathcal{C}}(\cdot)$ is defined as (2.2.12), decreases the Lyapunov function V by

$$V|_{t+} \leq V|_t - \underbrace{D_h(\hat{p}(t^+)||\hat{p}(t))}_{\geq 0}, \quad (5.4.60)$$

and hence guarantees stability.

Proof is given in the Appendix A.11

5.5 Simulation Study: Barret WAM7 Manipulator

In this section, we present extensive simulation experiments to assess the tracking performance of our geometric adaptive controller vis-a-vis existing adaptive controllers. The simulations are conducted on a torque-controllable Barret WAM seven-DOF manipulator arm model under the passivity-based adaptive control framework of [72]. For the dynamic model, only the rigid body dynamics is used with no joint frictions, and hence only the inertial parameters are considered for parameter adaptation. The modified recursive Newton-Euler algorithm is used to implement the adaptive controller based on [86]. The simulations are all implemented in MATLAB. Euler integration with a fixed

stepsize of 1e-3 seconds is used for the parameter adaptation implementation. For improved simulation fidelity, the forward dynamics is updated at a faster rate of 10kHz.

The four different adaptation laws to be compared are labelled as follows: ‘No-adaptation’, ‘Euclidean’, ‘Const. pullback’ and ‘Natural’. ‘No-adaptation’ indicates a pure passivity-based tracking controller with a fixed parameter estimate, i.e. $\hat{\Phi}(t) = \hat{\Phi}(0)$. ‘Euclidean’ and ‘Const. pullback’ are passivity-based adaptive controllers with the constant adaptation gain matrix respectively chosen as $\Gamma = \gamma \cdot \mathbb{I}$ and $\Gamma = \gamma \cdot g(\hat{\Phi}(0))$ (a constant pullback Riemannian metric evaluated on the initial estimate $\hat{\Phi}(0)$), for some constant scalar γ . When manually tuning the constant adaptation gain matrix, for reasonably close initial estimates, a non-identity or non-diagonal Γ for ‘Const. pullback’ is a reasonable choice that balances the update law with respect to the scales for each of the parameters. Finally, ‘Natural’ indicates a passivity-based adaptive controller using our geometric adaption law (5.4.52). The adaptation speed factor γ^{-1} is set at a scale as large as possible without accruing large numerical integration errors for the given stepsize; for ‘Euclidean’, $\gamma^{-1} = 5e-5$, while for ‘Const. pullback’ and ‘Natural’, $\gamma^{-1} = 1.0$. As noted in the Remark 5.1, the control gains are set to $K = \lambda \cdot M(q, \hat{\Phi})$, $\Lambda = \lambda \cdot \mathbb{I}$, and $\lambda = 10.0$.

The common task in our experiments is to track a periodic joint trajectory of the form

$$q_{d,i}(t) = A_i \left(\cos\left(\frac{2\pi}{T_i}t\right) - 1 \right) \quad , i = 1, \dots, n,$$

where $q_{d,i}$ is the desired joint angle of the i^{th} joint, A_i is the amplitude, and T_i is the period.

5.5.1 Full Adaptation

We adapt the entire set of inertial parameters for each of the links. First, to compare the robustness of the controllers with respect to various levels of model uncertainty, initial estimates of the inertial parameters are deliberately perturbed from the true parameters, i.e., $\hat{\phi}_i(0) = \phi_i \cdot (1 + \epsilon_i)$, where ϵ_i is drawn 100 times from a zero-mean Gaussian distribution with standard deviation σ evenly spaced at 0.05 intervals, from 0.05 to 0.40 (ϵ_i is truncated to be within the range $[-\sigma, \sigma]$ to prevent the initial estimate from being physically inconsistent). The desired trajectory parameters are fixed with amplitude $A_i = 0.8$ and period $T_i = 5$ sec.

As shown in Figure 5.4 our method ‘Natural’ and ‘Const. pullback’ outperform other methods in tracking performance at all noise levels. For the ‘Euclidean’ method, the estimated parameters were physically inconsistent for over 90 % of the operation time, which implies that using projection or resetting methods for ‘Euclidean’ merely enforce the physical consistency of the estimated parameters to be on the boundary of a user-specified (and ad hoc) feasible set. On the other hand, ‘Natural’ and ‘Const. pullback’ were shown to always provide physically consistent estimation parameters. Although the ‘Const. pullback’ method is not yet theoretically guaranteed to always ensure physical consistency, our results highlight the fact that a reasonable choice of metric or adaptation gain matrix Γ plays a more critical role in physically consistent estimation than only considering physical consistency on the boundary of the feasible region.

We now fix the model uncertainty to $\sigma = 0.40$ and consider trajectory tracking for a varied sequence of desired trajectories, each having different amplitudes A_i evenly spaced at 0.2 from 0.4 to 1.2. As can be observed in Figure 5.5, ‘Natural’ and ‘Const.

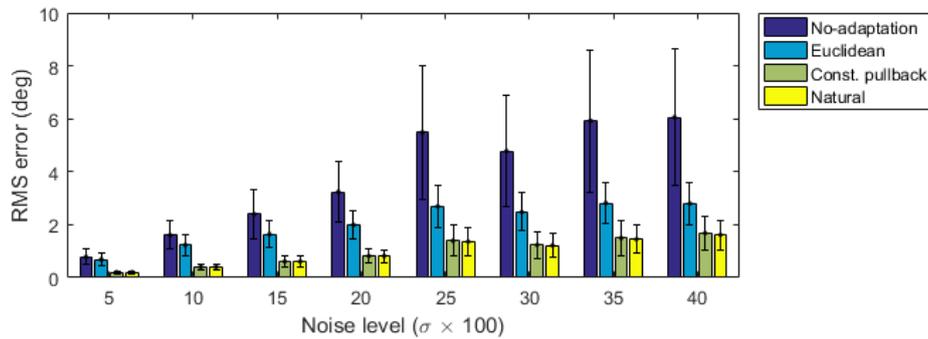


Figure 5.4: Plot of joint tracking RMS error when adapting the entire set of the inertial parameters with various levels of noise in the initial estimate

pullback’ considerably reduce the tracking error after the first round compared to ‘Euclidean’ and continues to maintain this level of tracking performance. The adapted parameters from ‘Natural’ and ‘Const. pullback’, while achieving physical consistency, are more generalizable to varied task situations. A lack of generalizability implies that the robot could behave unexpectedly in unexpected task situations. In this regard, ‘Natural’ and ‘Const. pullback’ methods show more promise with respect to robustness.

5.5.2 Unknown Payload Adaptation

We now consider adaptation of only the last end-effector link, to emulate situations where an unknown load or tool is loaded on the end-effector. The repeated sequence of constant amplitude 0.8 is used as the desired trajectory. The inertial parameters (except for the end-effector link) are fixed as true parameters. For the initial estimates, we assume a blind guess of the unknown loads assuming a 1kg sphere of radius 0.1m. The unknown loads considered for the experiment are of three types: ‘Sphere’, ‘Ellipsoid1’, and ‘Ellipsoid2’, respectively indicating a uniform sphere with radius 0.05m, a uniform

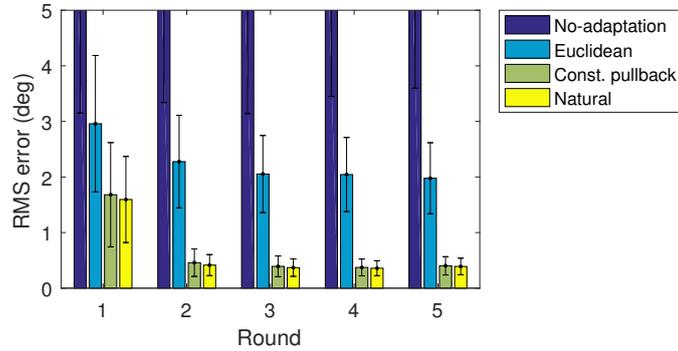


Figure 5.5: Plot of joint tracking RMS error when adapting the entire set of the inertial parameters on a varied sequence of desired trajectories, with a fixed noise level for the initial estimate; $\sigma = 0.4$. RMS errors for ‘No-adaptation’ all exceeded 7 deg

ellipsoid with aspect ratio [0.0323m: 0.3227m: 0.0323m], and a uniform ellipsoid with aspect ratio [0.1375m: 0.2165m: 0.1082m] and center of mass offset from the initial estimate by $p = (0.05\text{m}, 0.04\text{m}, -0.07\text{m})$, $\|p\| = 0.095\text{m}$. All the unknown loads weigh 3kg.

In Figure 5.6, our method, ‘Natural’, is shown to outperform other methods in terms of transient behavior and tracking error convergence. However, for the case of ‘Ellipsoid1’, our method did not show superior error convergence to other methods, compared to the cases of ‘Sphere’ and ‘Ellipsoid2’. In fact, the initial estimate, sphere, is highly biased from the pipe-shaped load, ‘Ellipsoid1’, which makes the adaptation most challenging out of the three cases. We believe that such limitations arising from bad initial estimates are an inherent feature of direct adaptive control methods, since the adaptation is done in a local gradient-like sense, directly compensating the model uncertainties with the instant *tracking error*. Indirect [77], [74] or modularized approaches [87], in which the accumulated past trajectory information is preserved in a more consistent way via

propagation of the covariance or observation matrix from, e.g., least-squares estimation, may perform better when there are significant errors in the initial estimate. However, they often may not satisfy the asymptotic convergence condition, e.g. uniform invertibility of the estimated mass matrix, or require additional torque and joint acceleration measurements. A composite approach [46] involving setting the initial covariance matrix to be the inverse of the pullback Riemannian metric on the initial estimate, could be a viable solution in this regard.

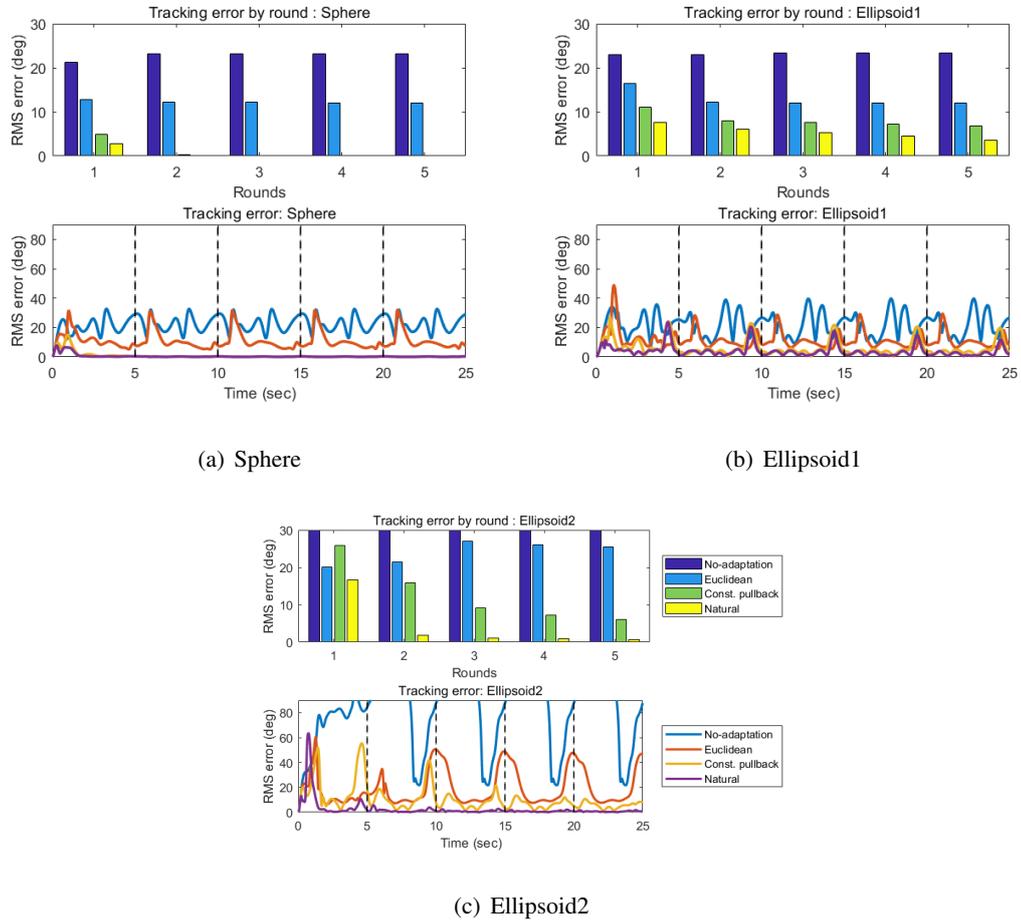


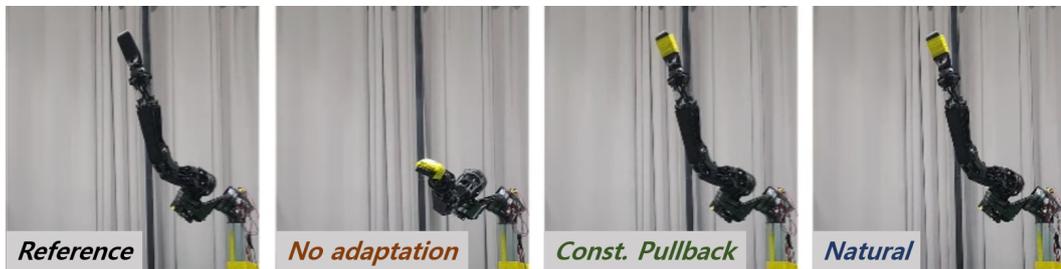
Figure 5.6: Joint tracking RMS error plots for adaptation of three kinds of unknown loads; 5.6(a) Sphere, 5.6(b) Ellipsoid1, and 5.6(c) Ellipsoid2. The plots above show mean RMS tracking error for repeated rounds of periodic desired trajectories, while the bottom plot shows a time plot of RMS tracking error. The color legends for all the plots are depicted in 5.6(c). Mean RMS tracking error for ‘No-adaptation’ in 5.6(c) all exceed 70 deg.

5.6 Experiment: AMBIDEX Robot Manipulator

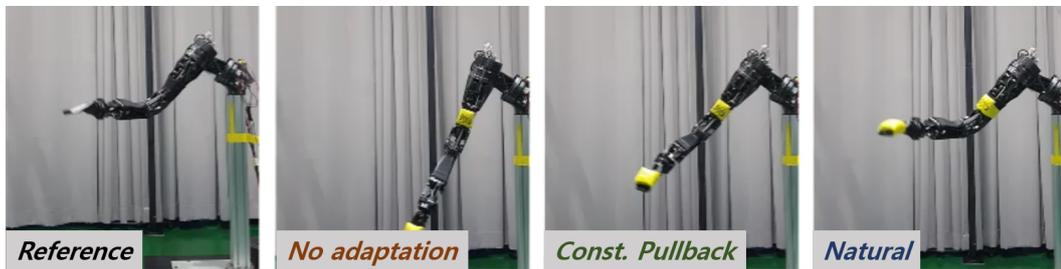
Our geometric robust adaptation algorithm is implemented on the AMBIDEX cable-driven robot manipulator. The modified recursive Newton-Euler algorithm is used to implement the passivity-based adaptive controller with a control frequency of 1kHz. Again referring to Remark 5.1, the control gains are set to $K = \lambda \cdot M(q, \hat{\Phi})$, $\Lambda = \lambda \cdot \mathbb{I}$, and $\lambda = 3.0$. Observe that the control gain λ is set to be very small as compared to the previous simulation experiment.

The task in this experiment is to sequentially track 14 sets of randomly chosen point-to-point motion trajectories; the duration of each point-to-point motion is 1.5 sec. Both the inertial parameters and Coulomb-viscous friction parameters are considered for parameter adaptation. The initial values for all the parameters are obtained from geometric dynamic identification as conducted in Section 3.6.1. To emulate situations where an unknown load or tool is loaded to the bodies, 500g sandbags are rigidly attached to the bodies. 3 different cases are considered: 500g sandbag on link 3, 500 sandbag each on link 3 and 10, and 500g sandbag each on link 3, 5 and 10. Friction parameter adaptation is mainly due to the relatively large model mismatch from using a simplified friction model. The scalar adaptation gain $\gamma = 1$ is used for all the experiments. We however set the adaptation speed for the friction parameters twice faster than the inertial parameters; the w_i are set to 0.5 for the friction parameter manifolds \mathcal{M}_J and 1 for the inertial parameter manifolds \mathcal{M}_I (see (5.3.43)).

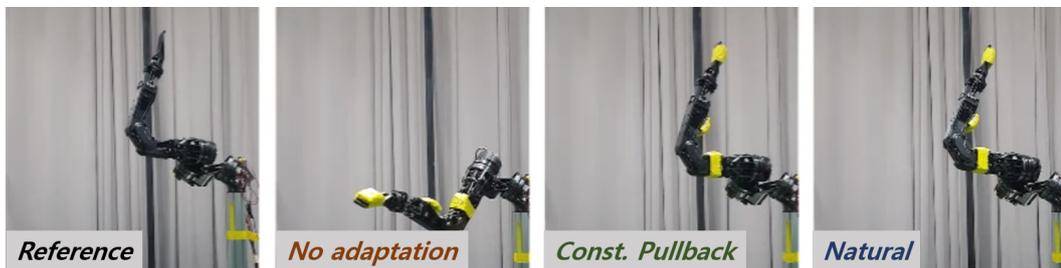
We note that the real experiment with the Euclidean metric as the constant adaptation gain matrix, i.e., $\Gamma = \gamma \cdot \mathbb{1}$, is not conducted for safety issue. In the simulation test, the parameter update with $\Gamma = \gamma \cdot \mathbb{1}$ was either numerically unstable (large γ) or too slow (small γ). This issue is shown to be more prominent than the previous experiment, perhaps for the inclusion of the friction parameters for adaptation (friction parameters



(a) Link 10 inertia + friction adaptation



(b) Link 3, 10 inertia + friction adaptation



(c) Link 3, 5, 10 inertia + friction adaptation

Figure 5.7: Snapshots of adaptation scenarios on the AMBIDEX robot manipulator. 500g (yellow colored) sandbags are placed on multiple body links.

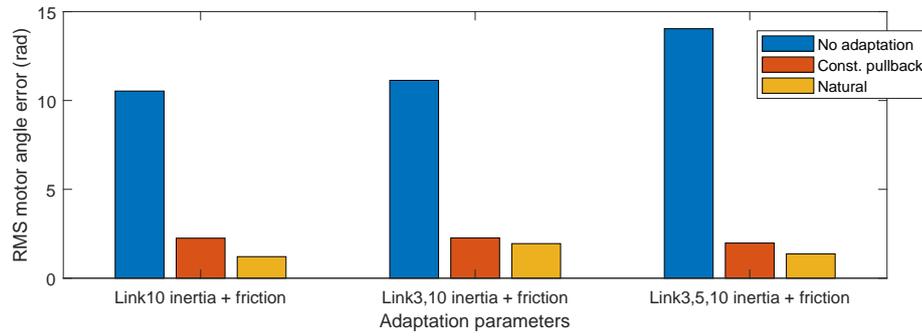


Figure 5.8: Joint tracking RMS errors during adaptive control of AMBIDEX robot manipulator subject to link inertia and friction uncertainties

are in a completely different scale as compared with the inertial parameters). Here we again highlight that our method does not require excessive tuning of the full adaptation gain matrix Γ .

Figure 5.8 shows the conspicuous reduction of the mean RMS tracking error (of the motor angle) from using our geometric adaptation methods. However, considering the reduction ratio from motor angle to joint angle due to the cable wiring, the mean RMS joint angle error from using our method is around the range of 1-3 deg. As can be observed in Figure 5.9, the steady state of the tracking error is also not clearly identifiable, although the parameter adaptation is nearly converged. We currently suspect that such performance degradation (as compared with the previous simulation experiment) mainly comes again from using a simplified joint friction model. In fact, we have observed that the friction model of AMBIDEX is highly complex and also configuration dependant, perhaps due to the unique kinematic design of AMBIDEX. We expect more precise modeling of the friction is a key to decrease the absolute tracking error (under the same condition where the control gain is set to be small, c.f., $\lambda = 3.0$).

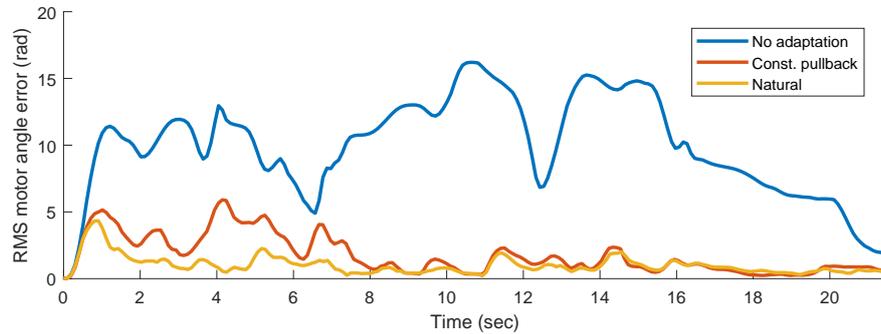


Figure 5.9: Joint tracking RMS error trajectory during adaptive control of AMBIDEX robot manipulator with 500g sandbag loaded at link 10.

Nonetheless, the present experiment still shows that for a lightweight robot like AMBIDEX loading moderately heavy object (0.5-1.5 kg) actually change the dynamic behavior much. In this regard, we have shown that the proposed geometric adaptive control framework allows to quickly adapt to the change in the mass-inertial dynamics while being robust to the model mismatch, and also can be employed in practice with a minimum tuning of the adaptation gain.

5.7 Conclusion

We have proposed a new set of geometric adaptation laws for direct adaptive control of robot manipulators. The robustness and practical utility of our algorithm, which has the added advantage of being implementable without excessive prior tuning, are both traceable in large part to exploiting the coordinate-invariant geometric structure of the physical parameter space. To highlight, our geometric approach does not impose significant additional computational burden compared to existing methods.

Adaptive control methods are currently being developed for more advanced robotic

tasks, e.g., interaction, constrained or hybrid position/force control [88], [89], or for complex structures, e.g., underactuated [44], flexible-joint [90], and spacecraft [91], with most existing works adopting the Lyapunov-based framework used in [71], [72]. Applying our natural adaptation laws to these more advanced tasks appears to be potentially profitable and direction for future work.

6

Conclusion

6.1 Summary

This thesis has shown that many of the challenges that arise in dynamic parameter estimation for multibody systems can be mitigated by appealing to coordinate-invariant, differential geometric methods. The key lies in the finding that physical parameters, in particular mass-inertial parameters, reside in a curved space with a natural Riemannian metric which captures the perturbations on the parameters in a coordinate-invariant, physically meaningful way. Taking this geometric perspective as our point of departure, we have devised geometric formulations and algorithms for robust dynamic identification, geometric criteria and algorithms for optimal excitation trajectory generation, and geometric adaptive control framework for robot manipulators. The main contributions of each chapter are now summarized.

- **Geometric Dynamic Identification of Multibody Systems**

A Riemannian manifold structure on a physically consistent rigid body inertial

parameter space is defined using the symmetric positive definite pseudo-inertia matrix representation of the inertial parameters. We then reformulate the classical linear least squares problem of inertial parameter identification to a nonlinear one by using a Riemannian geodesic distance. Further, convex second-order approximations of the Riemannian geodesic distance, entropic divergence and constant pullback distance, are proposed. These distances are used to relax the original non-convex geometric dynamic identification problem to a convex optimization problem. The proposed convex programming approach to the geometric dynamic identification leads to a globally optimal solution, reduced computations, faster and more reliable convergence, and easy inclusion of additional convex constraints as well as other mechanical parameters like friction parameters as optimization variables. Extensive experiments with the three-dof MIT Cheetah leg, the seven-dof AMBIDEX tendon-driven arm, and a sixteen-link articulated human model show markedly improved robustness and generalizability vis-à-vis existing vector space methods.

- **Geometric Criteria for Excitation Trajectory Optimization**

A new set of coordinate-invariant optimality criteria for generating excitation trajectories for dynamic identification of multibody mechanical systems has been presented. Practical benefit of using the proposed criterion for optimal excitation lies not only in allowing more accurate identification of the model parameters, but also in resulting optimal excitation trajectories, which are invariant to the choice of coordinate frames and linear reparametrizations of the observable parameters. The proposed geometric framework is also extended to devise a coordinate-invariant algorithm for assessing the effectively identifiable set of parameters given

a set of excitation trajectory samples. Reduced identification of only the effectively identifiable parameters is shown to significantly improve the accuracy of identification for high-dimensional humanoid robot.

- **Geometric Robust Adaptive Control of Robot Manipulators**

We have presented geometric adaptation laws for globally convergent direct adaptive control of robot manipulators. We first generalize the Riemannian geometric structure of the feasible inertial parameter space to general mechanical parameters by using the concept of Hessian manifold structure. Then, a coordinate-invariant Bregman divergence is shown to be naturally defined on the corresponding Hessian manifold in addition to the Riemannian metric. By using this Bregman divergence as a Lyapunov function candidate, the natural gradient parameter update law is derived with which physical feasibility of the parameters can be guaranteed in a completely smooth manner. Perhaps most importantly, our method considerably reduces the degree to which engineering choices must be made in the constant adaptation gain as compared to existing methods; only the choice of a scalar constant gain is required for adjusting the speed of adaptation, without regard to the number of parameters for adaptation. The method is further extended to define robust adaptation laws in which arbitrary convex constraint on the parameters can be satisfied using geometric projections at the boundary of the constraint. Validation through adaptive trajectory tracking control tasks on both simulation and real hardware experiment shows markedly improved tracking error convergence, again to highlight, without having to excessively tune the full adaptation gain matrix.

6.2 Future Work

There are a number of potential directions in which our current geometric line of research can be extended. While some of them are directly referred in the previous Chapters, here we raise some major directions and open problems of future research, which we believe merit further investigations.

- Throughout the thesis, we have mainly focused our attention on parameter estimation for rigid robots. We believe that the geometric approach adopted in this thesis can be extended to the case of soft robots. Parameters that are associated with elastic or stiffness forces should also be, like mass-inertia, positive in nature, since any elastic component should physically define a corresponding positive elastic energy. For instance, spatial stiffness matrices for modeling, e.g., compliant spatial mechanisms, reside in positive definite matrix manifold which like mass-inertia tensors exhibits a change of coordinate system in a form of affine group transformation. Adopting a geometric approach to more general types of compliant mechanical systems seems to be an interesting direction for future research.
- Geometric approaches for improving robustness in model-based methods proposed in this thesis are likely to share similar perspective with the ones toward developing efficient and stable learning algorithms for robot control. A class of natural gradient-based reinforcement learning algorithms [92, 93, 94, 95], in analogy to our geometric adaptive control algorithm, have been shown to allow much more stable and efficient learning process in comparison to non-geometric, coordinate-dependant approaches. However, these model-free methods typically require impractically long duration of training time to be used in real systems. The main

reason we believe is in making too little assumptions about the system or environment in an attempt to derive a completely general-purpose algorithm. Making a bridge or balance between model-based estimation/control methods and learning-based approaches from a unified geometric perspective thus seem to be a meaningful direction of research toward making an intelligent system that can learn in a robust, safe and efficient manner.

- Another interesting direction for robot learning control that has recently gained more attention is a simulation-to-real learning approach [10, 11, 96]; a control law is learned entirely in the simulation environment and then the learned control law is directly deployed in the real world. While, of course, accurate simulator or equivalently accurate dynamic model used in the simulator is in general desirable, one simple and effective method that is commonly used to circumvent the problem of modeling errors, often referred to as dynamic randomization technique [96], is to deliberately perturb or randomize the model used in the simulator during the learning process, i.e., a robot tries to learn a single control law in such a way that it is able to accomplish the given task for multiple perturbed simulation environments. Here we note that *robust control* methods in the control systems field essentially share the same philosophy or objective in designing the controller. In robust control, it is mostly the case that there exists a trade-off between the performance of the controller and allowable range of uncertainty of the model. In this regard, it would be interesting to investigate to what level the randomization can be applied for learning a robust control law without sacrificing too much on the performance. Here, the level of randomization applied to the dynamic model should be quantified in a coordinate invariant way. Lastly, as adaptive control is often regarded as an appealing alternative to robust control in

the control domain, it would be interesting to explore similar arguments or formal comparisons in terms of learning-based control, e.g., in what situations online model learning/adaptation approach would be more necessary than a single robust control law learned offline.

6.3 Concluding Remark

Dynamic model-based control methods in robotics are becoming mainstream for ever more complex systems like humanoids and multi-legged robots. However, rich and complete measurements for accurate dynamic parameter estimation can often be impractical to achieve due to many practical limitations. We would argue that our contribution is made more meaningful to the practitioner by offering methods that work robustly despite many of the practical difficulties. It is hoped that the set of geometric methods proposed in this thesis will apply to a more wide range of mechanical systems and also further be extended in a way to allow robots to perform more dynamic and more versatile set of tasks than ever before.

A

Proofs and Supplemental Derivations

A.1 Supplemental Derivations

Proposition A.1 (Derivation of (3.3.20)). *The differential metric on \mathcal{M} satisfies*

$$\begin{aligned} ds^2 &= \frac{1}{2} \text{tr}(P^{-1} dP P^{-1} dP) \\ &= \frac{1}{2} \left(\frac{dm}{m} \right)^2 + \frac{1}{2} \text{tr} \left(\left((\Sigma^C)^{-1} d\Sigma^C \right)^2 \right) \\ &\quad + m \cdot dp^T (\Sigma^C)^{-1} dp \end{aligned}$$

Proof. Using (3.3.17), it can be verified that a pseudo inertia P is given in terms of (m, p, Σ^C) by

$$P = \begin{bmatrix} \mathbb{I}_3 & p \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Sigma^C & 0 \\ 0 & m \end{bmatrix} \begin{bmatrix} \mathbb{I}_3 & p \\ 0 & 1 \end{bmatrix}^T, \quad (\text{A.1.1})$$

The affine-invariant metric, in particular, is invariant under translations. So, without loss

of generality, we consider the case when $p = 0$. Then, using the form of P from (A.1.1):

$$P = \begin{bmatrix} \Sigma^C & 0 \\ 0 & m \end{bmatrix}, \quad dP = \begin{bmatrix} d\Sigma^C & mdp \\ mdp^T & dm \end{bmatrix} \quad (\text{A.1.2})$$

$$P^{-1} dP = \begin{bmatrix} (\Sigma^C)^{-1} d\Sigma^C & m(\Sigma^C)^{-1} dp \\ dp^T & dm/m \end{bmatrix} \quad (\text{A.1.3})$$

The desired result then follows via matrix multiplication. \square

Proposition A.2 (Derivation of (3.3.29)). *In the coordinates (m, p, Σ^C) , the entropic divergence is given by*

$$\begin{aligned} d_{\mathcal{M}} ({}^1\phi \| {}^2\phi) &= d_h ({}^1m \| {}^2m) + d_h ({}^1\Sigma^C \| {}^2\Sigma^C) \\ &\quad + {}^1m ({}^1p - {}^2p)^T ({}^2\Sigma^C)^{-1} ({}^1p - {}^2p). \end{aligned} \quad (\text{A.1.4})$$

Proof. To show the result, we use the definition of the entropic divergence from (3.3.28) and (2.2.14), which is recalled as

$$d_{\mathcal{M}} ({}^1\phi \| {}^2\phi) = -\log \frac{|{}^1P|}{|{}^2P|} + \text{tr} ({}^2P^{-1} \cdot {}^1P) - 4$$

From (A.1.1) it can be deduced that

$$\log(|P|) = \log(m) + \log(|\Sigma^C|).$$

This leads to

$$\log \frac{|{}^2P|}{|{}^1P|} = \log \left(\frac{{}^2m}{{}^1m} \right) + \log \frac{|{}^2\Sigma^C|}{|{}^1\Sigma^C|}.$$

Noting that ${}^2P^{-1}$ can be written as

$${}^2P^{-1} = \begin{bmatrix} \mathbb{I}_3 & 0 \\ -{}^2p^T & 1 \end{bmatrix} \begin{bmatrix} ({}^2\Sigma^C)^{-1} & 0 \\ 0 & 1/{}^2m \end{bmatrix} \begin{bmatrix} \mathbb{I}_3 & -{}^2p \\ 0 & 1 \end{bmatrix}$$

it then follows via straightforward matrix multiplication that

$$\begin{aligned} \text{tr}({}^2P^{-1} \cdot {}^1P) &= \frac{{}^1m}{{}^2m} + \text{tr}(({}^2\Sigma^C)^{-1} \cdot {}^1\Sigma^C) \\ &\quad + {}^1m({}^1p - {}^2p)^T ({}^2\Sigma^C)^{-1} ({}^1p - {}^2p) \end{aligned}$$

The final form follows from the definition of the Bregman log determinant divergence (2.2.14). \square

Proposition A.3. *When mass is fixed to 1, the entropic divergence (3.3.28) is double the KL divergence between two multivariate Gaussian densities ${}^i\rho = \mathcal{N}({}^ip, {}^i\Sigma^C)$, $i = 1, 2$:*

$$\begin{aligned} D_{KL}({}^1\rho \| {}^2\rho) &\triangleq \int_{\Omega} {}^1\rho \log \left(\frac{{}^1\rho}{{}^2\rho} \right) dV \\ &= \frac{1}{2} \left[\log \frac{|{}^2\Sigma^C|}{|{}^1\Sigma^C|} + \text{tr}(({}^2\Sigma^C)^{-1} \cdot {}^1\Sigma^C) - 3 \right. \\ &\quad \left. + ({}^1p - {}^2p)^T ({}^2\Sigma^C)^{-1} ({}^1p - {}^2p) \right] \end{aligned} \quad (\text{A.1.5})$$

Proof. Directly substituting ${}^1m = {}^2m = 1$ into (A.1.4) yields,

$$\begin{aligned} d_{\mathcal{M}}({}^1\phi \| {}^2\phi) \Big|_{{}^1m, {}^2m=1} &= d_F({}^1\Sigma^C \| {}^2\Sigma^C) \\ &\quad + ({}^1p - {}^2p)^T ({}^2\Sigma^C)^{-1} ({}^1p - {}^2p) \\ &= 2 \cdot D_{KL}({}^1\rho \| {}^2\rho). \end{aligned}$$

\square

A.2 Proof of Proposition 3.2

The Lagrangian for $\hat{\Phi} \sim \{\hat{P}_j\}_{j=1}^n$ can be defined as

$$\mathcal{L}(\{\hat{P}_j\}_{j=1}^n, \lambda) = \sum_{j=1}^n \frac{1}{2} d_{\mathcal{P}(4)}(P_j, \hat{P}_j)^2 + \lambda \left(b - \sum_{j=1}^n \text{tr}(\hat{P}_j X_j) \right) \quad (\text{A.2.6})$$

where λ is the Lagrange multiplier. Then for all admissible curves $c_j : \mathbb{R} \rightarrow \mathcal{P}(4)$, $j = 1, \dots, n$ such that $c_j(0) = \hat{P}_j$ and $\dot{c}_j(0) = V_j \in \mathcal{S}(4)$,

$$\left. \frac{d}{dt} \right|_{t=0} \mathcal{L}(\{c_j(t)\}_{j=1}^n, \lambda) = 0 \quad (\text{A.2.7})$$

should hold as the first-order necessary condition. Since

$$\left. \frac{d}{dt} \right|_{t=0} d_{\mathcal{P}(4)}(P_j, c_j(t))^2 = \langle \text{Log}_{\hat{P}_j}(P_j), V_j \rangle_{\hat{P}_j} \quad (\text{A.2.8})$$

holds [36], we have

$$\sum_{j=1}^n \langle \text{Log}_{\hat{P}_j}(P_j) - \lambda \cdot \hat{P}_j X_j \hat{P}_j, V_j \rangle_{\hat{P}_j} = 0,$$

which should hold for all variations of V_j . Hence, the first-order necessary condition reduces to

$$\text{Log}_{\hat{P}_j}(P_j) = \lambda \cdot \hat{P}_j X_j \hat{P}_j, \quad j = 1, \dots, n. \quad (\text{A.2.9})$$

Define $Q_j \triangleq \log(P_j^{-1/2} \hat{P}_j P_j^{-1/2}) \in \mathcal{S}(4)$, which also implies $\hat{P}_j = P_j^{1/2} e^{Q_j} P_j^{1/2}$. Then the first-order necessary condition further simplifies to

$$(-Q_j) e^{-Q_j} = \lambda \bar{X}_j, \quad j = 1, \dots, n. \quad (\text{A.2.10})$$

Let the eigen decomposition of $-Q_j \in \mathcal{S}(4)$ be

$$-Q_j = R_j \Gamma_j R_j^T, \quad \text{where } \Gamma_j = \text{diag}(\gamma_j^1, \dots, \gamma_j^4). \quad (\text{A.2.11})$$

Then Equation (A.2.10) becomes

$$\Gamma_j e^{\Gamma_j} = \text{diag}(\gamma_j^1 e^{\gamma_j^1}, \dots, \gamma_j^4 e^{\gamma_j^4}) = \lambda R_j^T \bar{X}_j R_j, \quad (\text{A.2.12})$$

which implies that the matrix on the right-hand side is diagonal, and leads to the eigen-decomposition of \bar{X}_j to be of the form

$$\bar{X}_j = R_j \Sigma_j R_j^T, \quad \text{where } \Sigma_j = \text{diag}(\sigma_j^1, \dots, \sigma_j^4). \quad (\text{A.2.13})$$

Then γ_j^k is the solution of

$$\gamma_j^k e^{\gamma_j^k} = \lambda \sigma_j^k,$$

which is given by $\gamma_j^k = W(\lambda \sigma_j^k)$ from the definition of W , also referred to as the Lambert W function. Note that there exists two solutions γ_j^k when $\lambda \sigma_j^k \in (-1/e, 0)$, one greater than -1 , $W(\lambda \sigma_j^k)$, and the smaller $W_0(\lambda \sigma_j^k)$. However, since

$$\gamma_j^k = -\log(\lambda_k (P_j^{-1} \hat{P}_j)) \geq -\log(e) = -1 \quad (\text{A.2.14})$$

from the restriction that $\hat{\Phi} \in \mathcal{B}_\Phi$, a monotonically increasing inverse function W defined on the interval $[-1/e, \infty)$ is the only one of interest. For the solution γ_j^k to exist, $\lambda \sigma_j^k \geq -1/e$ should hold for all $j = 1, \dots, n$ and $k = 1, \dots, 4$. This condition can equivalently be stated as $\lambda \in [-g(\sigma_{max}), g(-\sigma_{min})]$.

Now the original constraint $\sum_{j=1}^n \text{tr}(\hat{P}_j X_j) - b = 0$ can be equivalently rewritten as

$$\begin{aligned} \sum_{j=1}^n \text{tr}(X_j \hat{P}_j) - b &= \sum_{j=1}^n \text{tr}(\bar{X}_j e^{Q_j}) - b \\ &= \sum_{j=1}^n \text{tr}(\Sigma_j e^{-\Gamma_j}) - b \\ &= \sum_{j=1}^n \sum_{k=1}^4 \sigma_j^k e^{-\gamma_j^k} - b \\ &= \sum_{j=1}^n \sum_{k=1}^4 \sigma_j^k e^{-W(\lambda \sigma_j^k)} - b = 0 \end{aligned} \quad (\text{A.2.15})$$

, where the last line (A.2.15) implies $\mathcal{C}(\lambda) = 0$ and it is clear that $\mathcal{C}(\cdot)$ is a monotonically decreasing function from the monotonicity of $W(\cdot)$.

In summary, the necessary condition for the optimal solution $\hat{\Phi} \in \mathcal{B}_\Phi$ to exist can be transformed into the existence of a unique solution $\hat{\lambda} \in [-g(\sigma_{max}), g(-\sigma_{min})]$ to $\mathcal{C}(\lambda) = 0$. If such a solution exists, $\hat{\phi}_j$ is given by (3.4.41); the solution is in fact a

local minima by the given convex invariant objective function [97] and linear constraint with respect to $P_j^{-1/2} \hat{P}_j P_j^{-1/2}$, whose eigenvalues are bounded on the interval $(0, e]$ from the restriction $\hat{\Phi} \in \mathcal{B}_{\Phi}$.

A.3 Supplemental Propositions for Section 3.5.1

We present two main propositions that rigorously justify the use of our method provided in Section 3.5.1 for provably comparing the generalizability of using different types of convex regularizers for dynamic parameter identification.

The optimization variable x can generally be decomposed as $x = [\Phi^T, f_{\text{fric}}^T]^T$ or simply $x = \Phi$, where Φ is a vector of inertial parameters and f_{fric} is a vector of friction parameters. The residual error loss $f(x) = \|Ax - b\|_{\Sigma^{-1}}^2$ is set as a least square error, and regularizer $g(x) = d(\Phi, {}^0\Phi)$ is given by an arbitrary convex distance function from the prior inertial parameter ${}^0\Phi$. A convex constraint set \mathcal{C} can be arbitrarily characterized by convex constraints that might include physical consistency conditions (including the \mathcal{E} -density reliability condition, CoM bounds, etc.) on the inertial parameters Φ , non-negativity constraints on the friction parameters f_{fric} , or additional linear equality constraints on the parameters, e.g., to enforce a constant total mass.

We first propose that the original regularized least square error formulation, e.g., (3.2.16), can be equivalently reformulated as a bounded least square error formulation, e.g., (3.5.52), in which the identical optimal solution can be instead obtained by projecting the prior parameters to the set of bounded least square error, or, when the convex constraint $x \in \mathcal{C}$ is considered, intersection of the set of bounded least square error and \mathcal{C} .

Proposition A.4. *Let \mathcal{C} be a convex set, and f (residual error loss) and g (regularizer) be convex differentiable functions, with the requirement that regularized objective function*

$J_{\gamma,g}(x) = f(x) + \gamma g(x)$ be strictly convex for all $\gamma > 0$. The unique solution $x_{\gamma,g}^*$ to the regularized objective $\min_{x \in \mathcal{C}} J_{\gamma,g}(x)$ is also a global minimum solution to the constrained formulation,

$$\min_{x \in \mathcal{C}} g(x) \quad \text{s.t.} \quad f(x) \leq f(x_{\gamma,g}^*) \quad (\text{A.3.16})$$

Proof. Let y be a solution to the convex problem (A.3.16). Then, $g(y) \leq g(x_{\gamma,g}^*)$ and $f(y) \leq f(x_{\gamma,g}^*)$ holds, which leads to $J_{\gamma,g}(y) \leq J_{\gamma,g}(x_{\gamma,g}^*)$. Since $x_{\gamma,g}^*$ is a unique minimizer to $J_{\gamma,g}(\cdot)$ over \mathcal{C} , $y = x_{\gamma,g}^*$. \square

Now, we claim that by only exchanging the regularization function g in the constrained formulation (A.3.16) to any other convex function h , the optimal solution preserves to lie at the intersection of the boundary of the set of bounded least square error and \mathcal{C}

Proposition A.5. *Let $h \neq g$ be another convex differentiable regularizer such that (1) $\arg \min_{x \in \mathcal{C}} h(x) = \arg \min_{x \in \mathcal{C}} g(x)$ and (2) $f(x) + \alpha h(x)$ strictly convex for all $\alpha > 0$. Then, the constrained optimization formulation:*

$$\min_{x \in \mathcal{C}} h(x) \quad \text{s.t.} \quad f(x) \leq f(x_{\gamma,g}^*) \quad (\text{A.3.17})$$

has a unique solution corresponding to the regularized solution $x_{\alpha,h}^$ for some $\alpha > 0$ satisfying $f(x_{\alpha,h}^*) = f(x_{\gamma,g}^*)$ (We exclude the trivial case when the regularized solution happens to exactly match the prior value, i.e., $x_{\gamma,g}^* \in \arg \min_{x \in \mathcal{C}} g(x)$, which can actually be realized with a probability of measure zero)*

Proof. Let y be a solution to the convex problem (A.3.17).

If $y \in \arg \min_{x \in \mathcal{C}} h(x) = \arg \min_{x \in \mathcal{C}} g(x)$, then $f(y) \leq f(x_{\gamma,g}^*)$ and $g(y) \leq g(x_{\gamma,g}^*)$ holds, which leads to $J_{\gamma,g}(y) \leq J_{\gamma,g}(x_{\gamma,g}^*)$. Since $x_{\gamma,g}^*$ is a unique minimizer to $J_{\gamma,g}(\cdot)$ over \mathcal{C} , $y = x_{\gamma,g}^*$. This corresponds to the trivial case where $x_{\gamma,g}^* = y \in \arg \min_{x \in \mathcal{C}} g(x)$.

Now, consider $y \notin \arg \min_{x \in \mathcal{C}} h(x)$. Then, the inequality constraint is an active constraint; the corresponding optimality conditions [98, pp. 5-7] can be abbreviated as,

$$y \in \arg \min_{x \in \mathcal{C}} h(x) + \lambda f(x) \quad (\text{A.3.18})$$

$$f(y) = f(x_{\gamma,g}^*) \quad (\text{A.3.19})$$

$$\lambda > 0 \quad (\text{A.3.20})$$

The positive scalar λ is a Lagrange multiplier or optimal multiplier to the constraint (A.3.19). Let $\alpha = 1/\lambda > 0$, then from (A.3.18) and (A.3.19), the optimal solutions are of the regularized solution form, $x_{\alpha,h}^*$ satisfying $f(x_{\alpha,h}^*) = f(x_{\gamma,g}^*)$. Suppose there exists more than a single optimal solution, say $x_{\alpha_1,h}^* \neq x_{\alpha_2,h}^*$ with $\alpha_1 \neq \alpha_2$. From the optimality conditions, $h(x_{\alpha_1,h}^*) = h(x_{\alpha_2,h}^*)$ and $f(x_{\alpha_1,h}^*) = f(x_{\alpha_2,h}^*)$ should hold. Then, $f(x_{\alpha_1,h}^*) + \alpha_1 h(x_{\alpha_1,h}^*) = f(x_{\alpha_2,h}^*) + \alpha_1 h(x_{\alpha_2,h}^*)$ holds which contradicts the fact that $x_{\alpha_1,h}^*$ is the unique global solution of $J_{\alpha_1,h}$. Therefore the constrained convex problem (A.3.17) has a unique global solution of the form $x_{\alpha,h}^*$, satisfying $f(x_{\alpha,h}^*) = f(x_{\gamma,g}^*)$. \square

A.4 Proof of Proposition 4.1

We first prove that (4.3.39) is invariant to body-fixed reference frames and physical units; observe that the expression itself is already regardless of the choice of B for base parameter representations. Denote Φ_{0a} , A_a , Σ_a the nominal inertial parameter, regressor matrix and observation covariance matrix evaluated at some set of body-fixed reference frames and under some physical units; superscript a represents the particular choice. Then the coordinate transformation rules can be given by the linear mappings: $\Phi_{0a} = P_{ab}\Phi_{0b}$, $A_a = Q_{ab}^{-1}A_bP_{ab}^{-1}$, $\Sigma_a = Q_{ab}\Sigma_bQ_{ab}^T$, for some nonsingular constant matrices P_{ab}

and Q_{ab} [81]. This results in,

$$A_a^T \Sigma_a^{-1} A_a = P_{ab}^{-T} (A_b^T \Sigma_b A_b) P_{ab}^{-1} \quad (\text{A.4.21})$$

From the coordinate invariance of the metric, i.e.,

$$ds^2 = d\Phi_{0a}^T G_{0a} d\Phi_{0a} = d\Phi_{0b}^T P_{ab}^T G_{0a} P_{ab} d\Phi_{0b}, \quad (\text{A.4.22})$$

it follows that,

$$G_{0a} = P_{ab}^{-T} G_{0b} P_{ab}^{-1}. \quad (\text{A.4.23})$$

From (A.4.21) and (A.4.23), transformation rule on $D = A^T \Sigma^{-1} A G_0^{-1}$ is finally given by

$$D_a = P_{ab}^{-T} D_b P_{ab}^T. \quad (\text{A.4.24})$$

Since the set of eigenvalues of a matrix X is always preserved under the transformation $X \mapsto G^{-1} X G$ for any nonsingular matrix G , we conclude that $f(\lambda_{\neq 0}(D))$ is invariant under the desired coordinate transformations.

Now, we claim that the nonzero eigenvalues of $A^T \Sigma^{-1} A G_0^{-1}$ is identical to the eigenvalues of $A_B^T \Sigma^{-1} A_B H_0^{-1}$, which suffices to complete the proof. The characteristic equation for $A^T \Sigma^{-1} A G_0^{-1}$ is given by,

$$\begin{aligned} p(s) &= \det(s\mathbb{I} - A^T \Sigma^{-1} A G_0^{-1}) \\ &= \det(s\mathbb{I} - B^T A_B^T \Sigma^{-1} A_B B G_0^{-1}) \\ &= s^{10n-n_B} \cdot \det(s\mathbb{I} - A_B^T \Sigma^{-1} A_B B G_0^{-1} B^T) \\ &= s^{10n-n_B} \cdot \det(s\mathbb{I} - A_B^T \Sigma^{-1} A_B H_0^{-1}), \end{aligned} \quad (\text{A.4.25})$$

where the last equality holds by the Sylvester's determinant identity, i.e., $\det(\mathbb{I} + XY) = \det(\mathbb{I} + YX)$ for square matrices XY and YX . Therefore, n_B number of nonzero eigenvalues of $A^T \Sigma^{-1} A G_0^{-1}$ exactly matches to the positive eigenvalues of $A_B^T \Sigma^{-1} A_B H_0^{-1}$.

A.5 Proof of Proposition 5.1

We first show that Hessian of h_I defined by (5.3.23) induces the Riemannian metric g_I given as (5.3.24). Gradient of the function h_I is given by,

$$\begin{aligned} [\nabla h_I(\phi)]_i &= \left. \frac{\partial}{\partial t} h_I(\phi + t \cdot e_i) \right|_{t=0} \\ &= - \operatorname{tr} (f(\phi + t \cdot e_i)^{-1} f(e_i)) \Big|_{t=0} \\ &= - \operatorname{tr} (f(\phi)^{-1} f(e_i)). \end{aligned} \quad (\text{A.5.26})$$

Hessian of h_I can also be derived in the same manner as

$$\begin{aligned} [\nabla^2 h_I(\phi)]_{i,j} &= \left. \frac{\partial}{\partial t} [\nabla h_I(\phi + t \cdot e_j)]_i \right|_{t=0} \\ &= - \left. \frac{\partial}{\partial t} \operatorname{tr} (f(\phi + t \cdot e_j)^{-1} f(e_i)) \right|_{t=0} \\ &= \operatorname{tr} (f(\phi)^{-1} f(e_i) f(\phi)^{-1} f(e_j)). \end{aligned}$$

Therefore, $g_I = \nabla^2 h_I$.

Now, we claim that Hessian manifold (\mathcal{M}_I, h_I) satisfies the conditions i and ii in the Definition 5.3.3 of barrier-Hessian manifold.

Condition i: Linear matrix inequality constraint, $f(\phi_b) \succ 0$, imposes an open convex cone in \mathbb{R}^{10} .

Condition ii: Bregman divergence of h_I can be calculated (from A.5.26) as

$$\begin{aligned} D_{h_I}(\phi \|\psi) &= -\log \det (f(\psi)^{-1} f(\phi)) + \operatorname{tr} (f(\psi)^{-1} f(\phi)) - 4 \\ &= \sum_{i=1}^4 \left[\log \lambda_i + \frac{1}{\lambda_i} - 1 \right], \end{aligned} \quad (\text{A.5.27})$$

where λ_i are the eigenvalues of $f(\phi)^{-1} f(\psi)$. Note that $\lambda_i \in (0, \infty)$ defines the set \mathcal{M}_I . As $\lambda_i \rightarrow 0^+$ or $\lambda_i \rightarrow \infty$ then $D_{h_I}(\phi \|\psi) \rightarrow \infty$. Therefore, upper-bounded D_{h_I} , i.e.,

$D_{h_I}(\phi\|\psi) \leq L$, characterizes a compact set on the eigenvalues λ_i and hence a compact set on $f(\phi)^{-1}f(\psi)$. For fixed ϕ , this implies compact set on ψ , as $\psi \mapsto f(\phi)^{-1}f(\psi)$ is a linear one-to-one mapping. Therefore, the set

$$\mathcal{B}_{(\mathcal{M}_I, h_I)}(\phi, L) \triangleq \{\psi \in \mathcal{M}_I : D_{h_I}(\phi\|\psi) \leq L\} \quad (\text{A.5.28})$$

is compact for every $\phi \in \mathcal{M}_I$ and L .

A.6 Proof of Proposition 5.2

We claim that Hessian manifold (\mathcal{M}_I^+, h_I^+) satisfies the conditions i and ii in the Definition 5.3.3 of barrier-Hessian manifold.

Condition i: Linear matrix inequality constraint, $f(\phi_b) \succ 0$, and linear inequalities, $a_k^T \phi - b_k > 0$, $k = 1, \dots, m$, imposes an open convex subset in \mathbb{R}^{10} .

Condition ii: By the linearity of Bregman divergence (iii),

$$D_{h_I^+}(\phi\|\psi) = D_{h_I}(\phi\|\psi) + \alpha \cdot \sum_{k=1}^m \left[\log \nu_k + \frac{1}{\nu_k} - 1 \right], \quad (\text{A.6.29})$$

where $\nu_k = \frac{a_k^T \psi - b_k}{a_k^T \phi - b_k}$ and $D_{h_I}(\phi\|\psi) = \sum_{i=1}^4 [\log \lambda_i + 1/\lambda_i - 1]$. Note that $\lambda_i \in (0, \infty)$ and $\nu_k \in (0, \infty)$ defines the set \mathcal{M}_I^+ . As $\lambda_i \rightarrow 0^+$ or $\lambda_i \rightarrow \infty$ or $\nu_k \rightarrow 0^+$ or $\nu_k \rightarrow \infty$ then $D_{h_I^+}(\phi\|\psi) \rightarrow \infty$. Therefore, upper-bounded $D_{h_I^+}$, i.e., $D_{h_I^+}(\phi\|\psi) \leq L$, characterizes a compact set on λ_i and ν_k , and hence (for fixed $\phi \in \mathcal{M}_I^+$) a compact set on ψ by a similar argument as above. Therefore, the set

$$\mathcal{B}_{(\mathcal{M}_I^+, h_I^+)}(\phi, L) \triangleq \{\psi \in \mathcal{M}_I^+ : D_{h_I^+}(\phi\|\psi) \leq L\} \quad (\text{A.6.30})$$

is compact for every $\phi \in \mathcal{M}_I^+$ and L .

A.7 Proof of Proposition 5.5

First of all, $\mathcal{M} = \mathcal{M}_x \times \mathcal{M}_y$ is clearly an affine manifold which can be identified as $N = \dim(\mathcal{M}_x) + \dim(\mathcal{M}_y)$ dimensional open convex subset in \mathbb{R}^N . Function $h = w_x h_x + w_y h_y$ is strictly convex as h_x and h_y are strictly convex functions. By the linearity of the Hessian operator and the Bregman divergence, the Riemannian metric and Bregman divergence can be straightforwardly derived as (5.3.39) and (5.3.40) respectively. Therefore (\mathcal{M}, h) is a Hessian manifold.

Now, it suffices to prove that the set

$$\mathcal{B}_{(\mathcal{M}, h)}(p_0, L) = \{(x, y) \in \mathcal{M} : D_{h_x}(x_0 \| x) + D_{h_y}(y_0 \| y) \leq L\}$$

is compact for every $p_0 = (x_0, y_0) \in \mathcal{M}$ and $L > 0$. For any $(x, y) \in \mathcal{B}_{(\mathcal{M}, h)}(p_0, L)$, $D_{h_x}(x_0 \| x) \leq D_{h_x}(x_0 \| x) + D_{h_y}(y_0 \| y) \leq L$ and hence $x \in \mathcal{B}_{(\mathcal{M}_x, h_x)}(x_0, L)$. Likewise, $y \in \mathcal{B}_{(\mathcal{M}_y, h_y)}(y_0, L)$. Therefore,

$$\mathcal{B}_{(\mathcal{M}, h)}(p_0, L) \subset \mathcal{B}_{(\mathcal{M}_x, h_x)}(x_0, L) \times \mathcal{B}_{(\mathcal{M}_y, h_y)}(y_0, L), \quad (\text{A.7.31})$$

which leads the set $\mathcal{B}_{(\mathcal{M}, h)}(p_0, L)$ is bounded. As $D_h(p_0 \| p) = D_{h_x}(x_0 \| x) + D_{h_y}(y_0 \| y)$ is a continuous function of p , it follows that its sublevel set $\mathcal{B}_{(\mathcal{M}, h)}(p_0, L)$ is closed. Therefore, $\mathcal{B}_{(\mathcal{M}, h)}(p_0, L)$ is a compact set, which completes the proof.

A.8 Proof of Proposition 5.6

By the non-negativity property (i) of the Bregman divergence, $V_p(p, \hat{p})$ is a valid Lyapunov function candidate. The time derivative of $V_p(p, \hat{p})$ can be derived as,

$$\begin{aligned} \frac{d}{dt} V_p(p, \hat{p}) &= \gamma \left[-\nabla h(\hat{p}) \dot{\hat{p}} - \dot{\hat{p}}^T \nabla^2 (p - \hat{p}) + \nabla h(\hat{p}) \dot{\hat{p}} \right] \\ &= \gamma \cdot \dot{\hat{p}}^T \nabla^2 h(\hat{p}) \dot{\hat{p}} \end{aligned}$$

where $\dot{p} = 0$ and $h(p) = 0$ are used since p is a constant true parameter vector.

A.9 Proof of Proposition 5.7

Consider the valid Lyapunov function candidate

$$V = V_t + V_p,$$

where V_p is defined by (5.4.49). Then,

$$\begin{aligned} \dot{V} &= \dot{V}_t + \dot{V}_p \\ &\leq \tilde{p}^T b + \gamma \cdot \tilde{p}^T \nabla^2 h(\hat{p}) \dot{\hat{p}} \\ &= \tilde{p}^T \left[\gamma \cdot \nabla^2 h(\hat{p}) \dot{\hat{p}} + b \right] \end{aligned} \tag{A.9.32}$$

holds from (5.4.46) and (5.4.50). From the adaptation rule defined by (5.4.51), $\dot{V} \leq 0$ holds, and asymptotic convergence of the tracking error can be shown in the same manner as before. Given that the initial parameter estimate is set to be physically consistent, i.e. $\hat{p}(0) \in \mathcal{M}$, $V_p|_{t=0}$ is a finite positive scalar. Hence, $V|_{t=0}$ is also a finite positive scalar. Moreover, since $\dot{V} \leq 0$, for all $T \in [0, \infty)$,

$$V|_{t=0} \geq V|_{t=T} \tag{A.9.33}$$

$$\geq V_p(p, \hat{p}(T)) \tag{A.9.34}$$

$$= \gamma \cdot d_h(p || \hat{p}(T)), \tag{A.9.35}$$

which leads to $d_h(p || \hat{p}(T)) \leq V_0/\gamma$. Therefore, the parameter estimate \hat{p} is always bounded in a compact set $\mathcal{B}_{(\mathcal{M}, h)}(p, V_0/\gamma) \subset \mathcal{M}$.

A.10 Proof of Proposition 5.8

Plugging adaptation law (5.4.57) into (A.9.32)

$$\begin{aligned}
 \dot{V} &\leq -\gamma \frac{\nabla c^T v}{\nabla c^T w} \tilde{p}^T \nabla c \\
 &= \alpha \cdot (p - \hat{p})^T \nabla c(\hat{p}) \\
 &\leq \alpha \cdot (c(p) - c(\hat{p})) \\
 &= \alpha \cdot c(p) < 0,
 \end{aligned}$$

where $\alpha = \gamma \frac{\nabla c^T v}{\nabla c^T w} > 0$ and $(p - \hat{p})^T \nabla c(\hat{p}) \leq c(p) - c(\hat{p})$ holds from the convexity of the function $c(\cdot)$.

A.11 Proof of Proposition 5.9

$$\begin{aligned}
 V|_{t^+} - V|_t &= V_p|_{t^+} - V_p|_t \\
 &= \gamma \cdot (d_h(p|\hat{p}(t^+)) - d_h(p|\hat{p}(t))) \\
 &= \gamma \cdot (d_h(p|\text{Proj}_{\mathcal{C}}(\hat{p}(t))) - d_h(p|\hat{p}(t))) \\
 &\leq -\gamma \cdot d_h(\text{Proj}_{\mathcal{C}}(\hat{p}(t))|\hat{p}(t)) \\
 &= -\gamma \cdot d_h(\hat{p}(t^+)|\hat{p}(t)),
 \end{aligned}$$

where the last inequality holds by the Pythagorean theorem (2.2.12) of the Bregman divergence.

B

Algorithms and Implementation Details

B.1 Mappings Associated with Inertial Parameters

- Given the rotational inertia tensor

$$I = \begin{bmatrix} I^{xx} & I^{xy} & I^{xz} \\ I^{yx} & I^{yy} & I^{yz} \\ I^{zx} & I^{zy} & I^{zz} \end{bmatrix} \in \mathcal{S}(3), \quad (\text{B.1.1})$$

define its vectorized representation as

$$\text{vec}(I) = [I^{xx}, I^{yy}, I^{zz}, I^{xy}, I^{yz}, I^{zx}] \in \mathbb{R}^6. \quad (\text{B.1.2})$$

Then, the vector representation of the full inertial parameters (3.2.1) can be given by,

$$\phi = [m, h^T, \text{vec}(I)^T]^T \in \mathbb{R}^{10}. \quad (\text{B.1.3})$$

- The mapping P from $\phi = [m, h^T, \text{vec}(I)^T]^T$ to the pseudo-inertia matrix (3.2.5) is given by

$$P(\phi) = \begin{bmatrix} \frac{1}{2}\text{tr}(I) \cdot \mathbb{I}_3 - I & h \\ h^T & m \end{bmatrix} \in \mathbb{R}^{4 \times 4}. \quad (\text{B.1.4})$$

- The inverse mapping P^{-1} from the pseudo-inertia matrix

$$Q = \begin{bmatrix} \Sigma & h \\ h^T & m \end{bmatrix} \in \mathbb{R}^{4 \times 4} \quad (\text{B.1.5})$$

to the vectorized inertial parameters is given by

$$P^{-1}(Q) = [m, h^T, \text{vec}(\text{tr}(\Sigma) \cdot \mathbb{I}_3 - \Sigma)^T]^T \in \mathbb{R}^{10}. \quad (\text{B.1.6})$$

- The pullback form of the affine-invariant Riemannian metric $g(\cdot) \in \mathbb{R}^{10 \times 10}$ on $\phi \in \mathbb{R}^{10}$ can be straightforwardly constructed from the definition (3.3.19) as,

$$[g(\phi)]_{i,j} = \frac{1}{2} \text{tr} (P(\phi)^{-1} P(e_i) P(\phi)^{-1} P(e_j)), \quad (\text{B.1.7})$$

for all $1 \leq i, j \leq 10$, where $e_k \in \mathbb{R}^{10}$ a vector having 1 at the k -th element and zero for the remaining elements.

- The standard inner product defined between the vectorized inertial parameters $\phi \in \mathbb{R}^{10}$ and some vector $a = [a_1, \dots, a_{10}] \in \mathbb{R}^{10}$ can be uniquely restated as the trace product between pseudo inertia matrix and some 4×4 symmetric matrix as,

$$\phi^T a = \text{tr}(P(\phi)X(a)), \quad (\text{B.1.8})$$

where the linear mapping $X : \mathbb{R}^{10} \rightarrow \mathcal{S}(4)$ is defined as,

$$X(a) = \begin{bmatrix} a_6 + a_7 & -\frac{1}{2}a_8 & -\frac{1}{2}a_{10} & \frac{1}{2}a_2 \\ -\frac{1}{2}a_8 & a_7 + a_5 & -\frac{1}{2}a_9 & \frac{1}{2}a_3 \\ -\frac{1}{2}a_{10} & -\frac{1}{2}a_9 & a_5 + a_6 & \frac{1}{2}a_4 \\ \frac{1}{2}a_2 & \frac{1}{2}a_3 & \frac{1}{2}a_4 & a_1 \end{bmatrix}. \quad (\text{B.1.9})$$

- For vectorized inertial parameters $\phi \in \mathbb{R}^{10}$ and some vector $b \in \mathbb{R}^{10}$, the following identity holds:

$$[g(\phi)]^{-1} \cdot b = 2 \cdot P^{-1}(P(\phi)X(b)P(\phi)). \quad (\text{B.1.10})$$

B.2 Inertial Parameter Perturbation Strategy

The following steps are followed for each body. Given a support region $\mathcal{E} \subset \mathbb{R}^3$, mass m , and pseudo-inertia matrix $P = P(\phi)$:

- (i) Designate the number of point masses N_m and mass uncertainty factor $\delta \in (0, 1)$. Uniformly sample the mass deviation Δ_m from the interval $[-m\delta, m\delta]$
- (ii) Uniformly sample positive scalars η_i such that $\sum_{i=1}^{N_m} \eta_i = 1$ and position vectors $x_i \in \mathcal{E} \subset \mathbb{R}^3$ for $i = 1, \dots, N_m$.
- (iii) Generate a “perturbed” pseudo-inertia matrix \tilde{P} by perturbing the given pseudo-inertia P according to:

$$\tilde{P} = P + \sum_{i=1}^{N_m} (\Delta_m \eta_i) \cdot \begin{bmatrix} x_i \\ 1 \end{bmatrix} \begin{bmatrix} x_i \\ 1 \end{bmatrix}^T. \quad (\text{B.2.11})$$

B.3 Recursive Regressor Gradient for Multibody Systems

We provide recursive algorithms for computing the regressor matrix and its gradient for a fixed-base tree-structured multibody system. Regressor matrices for both joint torques and ground reaction force are considered which suffices to cover the cases made in Chapter 4. We first present the original Newton-Euler inverse dynamics algorithm for a fixed-base tree-structured multibody system. Then, we demonstrate that computation

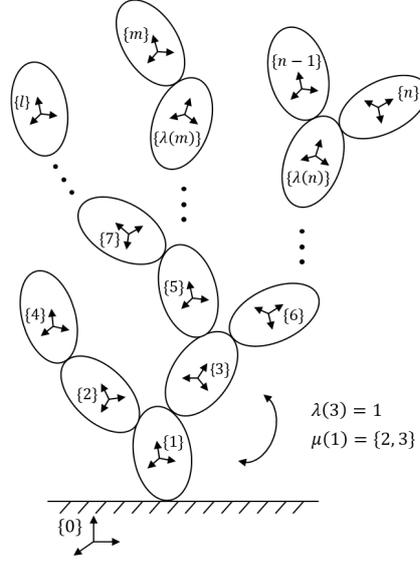


Figure B.1: Fixed-base tree-structured multibody system

of regressor matrices and their gradients can be induced naturally by following similar recursive computations. In particular, a coordinate-invariant form of the algorithm is derived using notations and operations associated with the Lie group of rigid body motions introduced in Section 2.3.

For n link fixed-base tree-structured multibody system (See Figure B.1), denote $\{i\}$ the body-fixed reference frame of link i , $1 \leq i \leq n$, and $\{0\}$ the stationary frame attached to the, e.g., ground. We use the following convention and definitions for link indices. First, the links and joints of the tree topology system are numbered in such a way that the numbers increase outward from the base. Define $\lambda(i)$ (called the connectivity index) to be the number of the link directly precedes link i . Then, let the displacement from $\{\lambda(i)\}$ to $\{i\}$, denoted $T_{\lambda(i),i} \in \text{SE}(3)$, be expressed as,

$$T_{\lambda(i),i} = M_i e^{[A_i]q_i}, \quad (\text{B.3.12})$$

where q_i is the angle of i -th joint connecting link $\lambda(i)$ and i , and $A_i \in \mathfrak{se}(3)$ is a screw representation of joint axis i in terms of frame $\{i\}$, and $M_i \in \text{SE}(3)$ denotes the location of frame $\{i\}$ relative to frame $\{\lambda(i)\}$ in the zero position. Also define

$$\mu(i) = \begin{cases} \{k \mid \lambda(k) = i\} & \text{if link } i \text{ is not a tip link of a branch} \\ \emptyset & \text{if link } i \text{ is a tip link of a branch} \end{cases} \quad (\text{B.3.13})$$

The set $\mu(i)$ is simply a list of the links that are direct successors of link i . Denote $V_i \in \mathfrak{se}(3)$ the body velocity of link i , $F_i \in \mathfrak{se}^*(3)$ the wrench applied by link $\lambda(i)$ to link i , all expressed in frame $\{i\}$. Also denote $\tau_i \in \mathbb{R}$ as the torque at joint i and $\tau = [\tau_1, \dots, \tau_n]^T \in \mathbb{R}^n$ as the joint torque vector. Finally, the recursive Newton-Euler inverse dynamics algorithm can be given as below.

Algorithm 1 Recursive Newton-Euler inverse dynamics algorithm for a floating-base tree-structured system

- 1: Initialize $V_0 = 0$, $\dot{V}_0 = -\dot{V}_s^{\text{grav}}$ (gravity acceleration)
 - 2: **for** $i = 1 : n$ **do**
 - 3: $T_{\lambda(i),i} = M_i e^{[A_i]q_i}$
 - 4: $T_{s,i} = T_{s,\lambda(i)} T_{\lambda(i),i}$
 - 5: $V_i = A_i \dot{q}_i + [\text{Ad}_{T_{\lambda(i),i}^{-1}}] V_{\lambda(i)}$
 - 6: $\dot{V}_i = A_i \ddot{q}_i + [\text{Ad}_{T_{\lambda(i),i}^{-1}}] \dot{V}_{\lambda(i)} + [\text{ad}_{V_i}] A_i \dot{q}_i$
 - 7: **end for**
 - 8: **for** $i = n : 1$ **do**
 - 9: $F_i = \sum_{k \in \mu(i)} [\text{Ad}_{T_{i,k}^{-1}}]^T F_k + G_i \dot{V}_i - [\text{ad}_{V_i}]^T G_i V_i$
 - 10: $\tau_i = A_i^T F_i$ (Joint torques)
 - 11: **end for**
 - 12: $F_0 = [\text{Ad}_{T_{0,1}^{-1}}]^T F_1$ (Ground reaction wrench)
-

Directly differentiating the line 5-6 in Algorithm 1 with respect to the motion parameter $p \in P$, i.e., $q(t) = q(t, \mathbf{P})$, yields the forward recursive algorithm below to be used later to compute the regressor gradient.

Algorithm 2 Recursive algorithm for kinematic derivatives

- 1: Initialize $\frac{\partial V_0}{\partial p} = 0$, $\frac{\partial \dot{V}_0}{\partial p} = 0$, $\forall p \in P$
 - 2: **for** $i = 1 : n$ **do**
 - 3: $\frac{\partial V_i}{\partial p} = A_i \frac{\partial \dot{q}_i}{\partial p} + [\text{Ad}_{T_{\lambda(i),i}^{-1}}] \frac{\partial V_{\lambda(i)}}{\partial p} + [\text{ad}_{A_i}] [\text{Ad}_{T_{\lambda(i),i}^{-1}}] V_{\lambda(i)} \frac{\partial q_i}{\partial p}$
 - 4: $\frac{\partial \dot{V}_i}{\partial p} = A_i \frac{\partial \ddot{q}_i}{\partial p} + [\text{Ad}_{T_{\lambda(i),i}^{-1}}] \frac{\partial \dot{V}_{\lambda(i)}}{\partial p} + [\text{ad}_{V_i}] A_i \frac{\partial \dot{q}_i}{\partial p}$
 $\quad - [\text{ad}_{A_i}] [\text{Ad}_{T_{\lambda(i),i}^{-1}}] \dot{V}_{\lambda(i)} \frac{\partial q_i}{\partial p} - [\text{ad}_{A_i}] \frac{\partial V_i}{\partial p} \dot{q}_i$
 - 5: **end for**
-

B.3.1 Joint Torque Regressor

Observe that the joint torque vector would eventually be given in the form,

$$\tau = Y\Phi, \quad (\text{B.3.14})$$

where $Y \in \mathbb{R}^{n \times 10n}$ is the joint torque regressor matrix. Now, the relation in line 9 of Algorithm 1 can be viewed as a set of linear systems of equations in terms of $\mathcal{F} = [F_1^T, \dots, F_n^T]^T \in \mathbb{R}^{6n}$. The solutions are then given in the linear form as

$$\mathcal{F} = W\Phi, \quad (\text{B.3.15})$$

$$W = \begin{bmatrix} W_{1,1} & W_{1,2} & \cdots & W_{1,n} \\ W_{2,1} & W_{2,2} & \cdots & W_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ W_{n,1} & W_{n,2} & \cdots & W_{n,n} \end{bmatrix} \in \mathbb{R}^{6n \times 10n},$$

where a sub-regressor matrix $W_{i,j} \in \mathbb{R}^{6 \times 10}$ relates the contribution of inertial parameter $\phi_i \in \mathbb{R}^{10}$ of link j to F_i . Replacing F_i and F_k in line 10 of Algorithm 1 with the representation (B.3.15) and comparing the left- and right-hand sides yield the following identities:

$$W_{ij} = \begin{cases} \Gamma(\dot{V}_i) - [\text{ad}_{V_i}]^T \Gamma(V_i) & \text{if } i = j \\ \sum_{k \in \mu(i)} [\text{ad}_{T_{i,k}^{-1}}]^{T} W_{kj} & \text{if } i < j \\ 0_{6 \times 10} & \text{otherwise} \end{cases}, \quad (\text{B.3.16})$$

where a linear mapping $\Gamma : \text{se}(3) \rightarrow \mathbb{R}^{6 \times 10}$ is explicitly given by

$$\Gamma(V) = \begin{bmatrix} & & & \omega_2 & 0 & \omega_3 \\ 0_{3 \times 1} & -[v] & \text{diag}(\omega) & \omega_1 & \omega_3 & 0 \\ & & & 0 & \omega_2 & \omega_1 \\ v & [\omega] & 0_{3 \times 3} & 0_{3 \times 3} & & \end{bmatrix}, \quad (\text{B.3.17})$$

for $V = (\omega, v) \in \text{se}(3)$.

Finally, from the relation $\tau = \mathcal{A}^T \mathcal{F}$ (where $\mathcal{A} \in \mathbb{R}^{n \times 6n}$ is a block diagonal matrix of A_i s), the joint torque regressor mapping can be given by $Y = \mathcal{A}^T W$. Differentiating the relations (B.3.16) with respect to the motion parameter $p \in P$ further yields the recursive set of equations for computing the gradient of the regressor Y . The algorithm is shown below.

Algorithm 3 Recursive algorithm for evaluating and computing gradient of the joint torque regressor matrix

- 1: Run line 1-7 of Algorithm 1
 - 2: Run Algorithm 2
 - 3: Initialize $W = 0$ and $\frac{\partial W}{\partial p} = 0, \forall p \in P$
 - 4: **for** $i = n : 1$ **do**
 - 5: **for** $j = 1 : n$ **do**
 - 6: **if** $i = j$ **then**
 - 7: $W_{i,j} = \Gamma(\dot{V}_i) - [\text{ad}_{V_i}]^T \Gamma(V_i)$
 - 8:
 - 9: $\frac{\partial W_{i,j}}{\partial p} = \Gamma(\frac{\partial \dot{V}_i}{\partial p}) - [\text{ad}_{\frac{\partial V_i}{\partial p}}]^T \Gamma(V_i) - [\text{ad}_{V_i}]^T \Gamma(\frac{\partial V_i}{\partial p})$
 - 10: **else if** $i < j$ **then**
 - 11: $W_{i,j} = \sum_{k \in \mu(i)} [\text{ad}_{T_{i,k}^{-1}}]^T W_{k,j}$
 - 12: $\frac{\partial W_{i,j}}{\partial p} = \sum_{k \in \mu(i)} \left\{ [\text{ad}_{T_{i,k}^{-1}}]^T \frac{\partial W_{k,j}}{\partial p} - [\text{ad}_{T_{i,k}^{-1}}]^T [\text{ad}_{A_k}]^T W_{k,j} \frac{\partial q_k}{\partial p} \right\}$
 - 13: **end if**
 - 14: **end for**
 - 15: **end for**
 - 16: $Y = \mathcal{A}^T W$ (Joint torque regressor)
 - 17: $\frac{\partial Y}{\partial p} = \mathcal{A}^T \frac{\partial W}{\partial p}$ (Joint torque regressor gradient)
-

B.3.2 Ground Reaction Wrench Regressor

Obtaining ground reaction wrench regressor $Y_0 \in \mathbb{R}^{6 \times 10n}$ and its gradient is straightforward from the relation,

$$\begin{aligned} F_0 &= [\text{Ad}_{T_{0,1}^{-1}}]^T F_1 \\ &= \underbrace{[\text{Ad}_{T_{0,1}^{-1}}]^T \begin{bmatrix} W_{1,1}, & \dots, & W_{1,n} \end{bmatrix}}_{\triangleq Y_0} \Phi. \end{aligned} \quad (\text{B.3.18})$$

The algorithm is given below.

Algorithm 4 Recursive algorithm for evaluating and computing gradient of the joint torque regressor matrix

- 1: Given $T_{0,1}$
 - 2: Run line 1-15 of Algorithm 3
 - 3: $Y_0 = [\text{Ad}_{T_{0,1}^{-1}}]^T \begin{bmatrix} W_{1,1}, & \dots, & W_{1,n} \end{bmatrix}$
(Ground reaction wrench regressor)
 - 4: $\frac{\partial Y_0}{\partial p} = [\text{Ad}_{T_{0,1}^{-1}}]^T \begin{bmatrix} \frac{\partial W_{1,1}}{\partial p}, & \dots, & \frac{\partial W_{1,n}}{\partial p} \end{bmatrix}, \forall p \in P$
(Ground reaction wrench regressor gradient)
-

B.4 Matrix Inverse-free Computation of Geometric Adaptation Laws

In this Section, we provide explicit formulas for efficient implementation of our geometric adaptation law (5.4.52), i.e., how to compute $[g(p)]^{-1}b$ without having to compute the inverse of $[g(p)]$. We assume that the barrier-Hessian manifold (\mathcal{M}, h) of the full set of parameters is defined by using the construction (5.3.41)-(5.3.44). Then $[g(p)]^{-1}b$

can be decomposed as,

$$[g(p)]^{-1}b = \sum_{i=1}^n \frac{1}{w_i} [g_i(x_i)]^{-1}b_i, \quad (\text{B.4.19})$$

where $b = [b_1^T, \dots, b_n^T]^T$. Therefore, it suffices to show how to compute the form $[g_i(x_i)]^{-1}b_i$ for each barrier-Hessian manifold (\mathcal{M}_i, h_i) that constitutes the full barrier-Hessian manifold (\mathcal{M}, h) . Below we recall the definitions of the barrier-Hessian manifolds presented in Section 5.3.1 and 5.3.2, and provide explicit formulas for implementing the geometric adaptation law.

- $\mathcal{M}_I = \{\phi \in \mathbb{R}^{10} : P(\phi) \succ 0\}$, $h_I(\phi) = -\log \det(P(\phi))$:

$$[g_I(\phi)]^{-1}b_I = 2 \cdot P^{-1}(P(\phi)X(b_I)P(\phi)) \quad (\text{B.4.20})$$

- $\mathcal{M}_I^+ = \{\phi \in \mathcal{M}_I : a_k^T \phi - b_k > 0, k = 1, \dots, m\}$, $h_I^+(\phi) = h_I(\phi) - \alpha \cdot \sum_{k=1}^m \log(a_k^T \phi - b_k)$:

$$[g_I^+(\phi)]^{-1}b_I = 2 \cdot P^{-1}(P(\phi)X(b_I)P(\phi)) - \bar{A} \left(\frac{1}{\alpha} \mathbb{I}_m + A^T \bar{A} \right)^{-1} \bar{A}^T b_I, \quad (\text{B.4.21})$$

where the k -th column of $A \in \mathbb{R}^{10 \times m}$ is given by $A_k = \frac{1}{(a_k^T \phi - b_k)} a_k$, and the k -th column of $\bar{A} \in \mathbb{R}^{10 \times m}$ is given by $\bar{A}_k = [g_I(\phi)]^{-1}A_k = 2 \cdot P^{-1}(P(\phi)X(A_k)P(\phi))$.

- $\mathcal{M}_J = \{\mu \in \mathbb{R}^{n_J} : \mu_i > 0, i = 1, \dots, n_J\}$, $h_J(\mu) = -\sum_{i=1}^{n_J} \log \mu_i$:

$$[g_J(\mu)]^{-1}b_J = \sum_{i=1}^{n_J} b_{Ji} \cdot \mu_i^2 \quad (\text{B.4.22})$$

- $\mathcal{M}_J^+ = \{\mu \in \mathbb{R}^{n_J} : l_i < \mu_i < r_i, i = 1, \dots, n_J\}$, $h_J^+(\mu) = -\sum_{i=1}^{n_J} [\log(\mu_i - l_i) + \log(r_i - \mu_i)]$:

$$[g_J^+(\mu)]^{-1}b_J = \sum_{i=1}^{n_J} \frac{b_{Ji}}{1/(\mu_i - l_i)^2 + 1/(\mu_i - r_i)^2} \quad (\text{B.4.23})$$

Bibliography

- [1] David E Orin, Ambarish Goswami, and Sung-Hee Lee. Centroidal dynamics of a humanoid robot. *Autonomous Robots*, 35(2-3):161–176, 2013.
- [2] Marco Hutter, Hannes Sommer, Christian Gehring, Mark Hoepflinger, Michael Bloesch, and Roland Siegwart. Quadrupedal locomotion using hierarchical operational space control. *The International Journal of Robotics Research*, 33(8):1047–1062, 2014.
- [3] Jonas Koenemann, Andrea Del Prete, Yuval Tassa, Emanuel Todorov, Olivier Stasse, Maren Bennewitz, and Nicolas Mansard. Whole-body model-predictive control applied to the hrp-2 humanoid. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2015)*, page 8p, 2015.
- [4] Scott Kuindersma, Robin Deits, Maurice Fallon, Andrés Valenzuela, Hongkai Dai, Frank Permenter, Twan Koolen, Pat Marion, and Russ Tedrake. Optimization-based locomotion planning, estimation, and control design for the atlas humanoid robot. *Autonomous Robots*, 40(3):429–455, 2016.
- [5] Alexander Herzog, Nicholas Rotella, Sean Mason, Felix Grimmering, Stefan Schaal, and Ludovic Righetti. Momentum control with hierarchical inverse dynamics on a torque-controlled humanoid. *Autonomous Robots*, 40(3):473–491, 2016.
- [6] G. Bledt, P. M. Wensing, S. Ingersoll, and S. Kim. Contact model fusion for event-based locomotion in unstructured terrains. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1–8, May 2018.
- [7] Roy Featherstone. *Rigid body dynamics algorithms*. Springer, 2014.

- [8] Abhinandan Jain. *Robot and multibody dynamics: analysis and algorithms*. Springer Science & Business Media, 2010.
- [9] Frank C. Park, Beobkyoon Kim, Cheongjae Jang, and Jisoo Hong. Geometric algorithms for robot dynamics: A tutorial review. *Applied Mechanics Reviews*, 70(1):010803, feb 2018.
- [10] Jemin Hwangbo, Joonho Lee, Alexey Dosovitskiy, Dario Bellicoso, Vassilios Tsounis, Vladlen Koltun, and Marco Hutter. Learning agile and dynamic motor skills for legged robots. *Science Robotics*, 4(26):eaau5872, 2019.
- [11] Jie Tan, Tingnan Zhang, Erwin Coumans, Atil Iscen, Yunfei Bai, Danijar Hafner, Steven Bohez, and Vincent Vanhoucke. Sim-to-real: Learning agile locomotion for quadruped robots. *arXiv preprint arXiv:1804.10332*, 2018.
- [12] Wenwei Yu, Chuanjian Liu, and Greg Turk. Preparing for the unknown: Learning a universal policy with online system identification. *ArXiv*, abs/1702.02453, 2017.
- [13] Wenhao Yu, Visak C. V. Kumar, Greg Turk, and C. Karen Liu. Sim-to-real transfer for biped locomotion. *CoRR*, abs/1903.01390, 2019.
- [14] Guillaume Rao, David Amarantini, Eric Berton, and Daniel Favier. Influence of body segments' parameters estimation models on inverse dynamics solutions during gait. *Journal of Biomechanics*, 39(8):1531–1536, 2006.
- [15] Christopher G Atkeson, Chae H An, and John M Hollerbach. Estimation of inertial parameters of manipulator loads and links. *The International Journal of Robotics Research*, 5(3):101–119, 1986.

- [16] Maher Moakher and Philipp G Batchelor. Symmetric positive-definite matrices: From geometry to applications and visualization. In *Visualization and Processing of Tensor Fields*, pages 285–298. Springer, 2006.
- [17] William James and Charles Stein. Estimation with quadratic loss. In *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 361–379, 1961.
- [18] Brian Kulis, Mátyás Sustik, and Inderjit Dhillon. Learning low-rank kernel matrices. In *Proceedings of the 23rd international conference on Machine learning*, pages 505–512. ACM, 2006.
- [19] Andrzej Cichocki, Sergio Cruces, and Shun-ichi Amari. Log-determinant divergences revisited: Alpha-beta and gamma log-det divergences. *Entropy*, 17(5):2988–3034, 2015.
- [20] Pradeep K Khosla and Takeo Kanade. Parameter identification of robot dynamics. In *1985 24th IEEE Conference on Decision and Control*, pages 1754–1760. IEEE, 1985.
- [21] Maxime Gautier and Wisama Khalil. On the identification of the inertial parameters of robots. In *Proceedings of the 27th IEEE Conference on Decision and Control*, pages 2264–2269. IEEE, 1988.
- [22] Michael Mistry, Stefan Schaal, and Katsu Yamane. Inertial parameter estimation of floating base humanoid systems using partial force sensing. In *2009 9th IEEE-RAS International Conference on Humanoid Robots*, pages 492–497. IEEE, 2009.
- [23] Ko Ayusawa, Gentiane Venture, and Yoshihiko Nakamura. Identifiability and identification of inertial parameters using the underactuated base-link dynamics

- for legged multibody systems. *The International Journal of Robotics Research*, 33(3):446–468, 2014.
- [24] Jan Swevers, Chris Ganseman, Dilek Bilgin, Joris De Schutter, and Hendrik Van Brussel. Optimal robot excitation and identification. *IEEE transactions on robotics and automation*, 13(5):730–740, 1997.
- [25] Jovana Jovic, Adrien Escande, Ko Ayusawa, Eiichi Yoshida, Abderrahmane Kheddar, and Gentiane Venture. Humanoid and human inertia parameter identification using hierarchical optimization. *IEEE Transactions on Robotics*, 32(3):726–735, 2016.
- [26] Vincent Bonnet, Philippe Fraisse, André Crosnier, Maxime Gautier, Alejandro González, and Gentiane Venture. Optimal exciting dance for identifying inertial parameters of an anthropomorphic structure. *IEEE Transactions on Robotics*, 32(4):823–836, 2016.
- [27] Cristóvão D Sousa and Rui Cortesão. Physical feasibility of robot base inertial parameter identification: A linear matrix inequality approach. *The International Journal of Robotics Research*, 33(6):931–944, 2014.
- [28] Patrick M Wensing, Sangbae Kim, and Jean-Jacques E Slotine. Linear matrix inequalities for physically consistent inertial parameter identification: A statistical perspective on the mass distribution. *IEEE Robotics and Automation Letters*, 3(1):60–67, 2018.
- [29] Silvio Traversaro, Stanislas Brossette, Adrien Escande, and Francesco Nori. Identification of fully physical consistent inertial parameters using optimization on manifolds. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5446–5451. IEEE, 2016.

- [30] Benoît Raucent, Guy Campion, Georges Bastin, Jean-Claude Samin, and Pierre Y Willems. Identification of the barycentric parameters of robot manipulators from external measurements. *Automatica*, 28(5):1011–1016, 1992.
- [31] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.
- [32] Miquel Calvo and Josep M Oller. A distance between multivariate normal distributions based in an embedding into the siegel group. *Journal of multivariate analysis*, 35(2):223–242, 1990.
- [33] Minyeon Han and Frank C Park. Dti segmentation and fiber tracking using metrics on multivariate normal distributions. *Journal of mathematical imaging and vision*, 49(2):317–334, 2014.
- [34] S. Kullback and R. A. Leibler. On information and sufficiency. *Ann. Math. Statist.*, 22(1):79–86, 1951.
- [35] Huaiyu Zhu and Richard Rohwer. Information geometric measurements of generalisation. *Technical Report*, 1995.
- [36] Hermann Karcher. Riemannian center of mass and mollifier smoothing. *Communications on pure and applied mathematics*, 30(5):509–541, 1977.
- [37] Nicolas Boumal, Bamdev Mishra, P-A Absil, and Rodolphe Sepulchre. Manopt, a matlab toolbox for optimization on manifolds. *The Journal of Machine Learning Research*, 15(1):1455–1459, 2014.
- [38] K. Ayusawa and Y. Nakamura. Identification of standard inertial parameters for large-dof robots considering physical consistency. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 6194–6201, Oct 2010.

- [39] Y. Kim. Anthropomorphic low-inertia high-stiffness manipulator for high-speed safe interaction. *IEEE Transactions on Robotics*, 33(6):1358–1374, Dec 2017.
- [40] Erling D Andersen, Cornelis Roos, and Tamas Terlaky. On implementing a primal-dual interior-point method for conic quadratic optimization. *Mathematical Programming*, 95(2):249–277, 2003.
- [41] Michael Grant, Stephen Boyd, and Yinyu Ye. *Cvx: Matlab software for disciplined convex programming*, 2008.
- [42] Guido Tournois, Michele Focchi, Andrea Del Prete, Romeo Orsolino, Darwin G Caldwell, and Claudio Semini. Online payload identification for quadruped robots. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4889–4896. IEEE, 2017.
- [43] Taeyoon Lee and Frank C. Park. A natural adaptive control law for robot manipulators. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018.
- [44] Daniele Pucci, Francesco Romano, and Francesco Nori. Collocated adaptive control of underactuated mechanical systems. *IEEE Transactions on Robotics*, 31(6):1527–1536, 2015.
- [45] Jean-Jacques E Slotine, Weiping Li, et al. *Applied nonlinear control*. Prentice hall Englewood Cliffs, NJ, 1991.
- [46] Jean-Jacques E Slotine and Weiping Li. Composite adaptive control of robot manipulators. *Automatica*, 25(4):509–519, 1989.
- [47] Miomir Vukobratovic and Branislav Borovac. Zero-moment point - thirty five years of its life. *I. J. Humanoid Robotics*, 1:157–173, 03 2004.

- [48] 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 *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018.
- [49] Patrick M Wensing, Albert Wang, Sangok Seok, David Otten, Jeffrey Lang, and Sangbae Kim. Proprioceptive actuator design in the mit cheetah: Impact mitigation and high-bandwidth physical interaction for dynamic legged robots. *IEEE Transactions on Robotics*, 33(3):509–522, 2017.
- [50] Patrick M Wensing, Günter Niemeyer, and Jean-Jacques E Slotine. Observability in inertial parameter identification. *arXiv preprint arXiv:1711.03896*, 2017.
- [51] J. Di Carlo, P. M. Wensing, B. Katz, G. Bledt, and S. Kim. Dynamic locomotion in the mit cheetah 3 through convex model-predictive control. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018.
- [52] Y. Ogawa, G. Venture, and C. Ott. Dynamic parameters identification of a humanoid robot using joint torque sensors and/or contact forces. In *2014 IEEE-RAS International Conference on Humanoid Robots*, pages 457–462, Nov 2014.
- [53] B. J. Stephens. State estimation for force-controlled humanoid balance using simple models in the presence of modeling error. In *2011 IEEE International Conference on Robotics and Automation*, pages 3994–3999, May 2011.
- [54] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, Princeton, NJ, 2008.
- [55] Dinesh K Pai. Muscle mass in musculoskeletal models. *Journal of biomechanics*, 43(11):2093–2098, 2010.

- [56] Brian Armstrong. On finding exciting trajectories for identification experiments involving systems with nonlinear dynamics. *The International Journal of Robotics Research*, 8(6):28–48, 1989.
- [57] Ko Ayusawa, Antoine Rioux, Eiichi Yoshida, Gentiane Venture, and Maxime Gautier. Generating persistently exciting trajectory based on condition number optimization. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6518–6524. IEEE, 2017.
- [58] Friedrich Pukelsheim. *Optimal Design of Experiments (Classics in Applied Mathematics) (Classics in Applied Mathematics, 50)*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2006.
- [59] A. N. Donev, A. C. Atkinson, and R. D. Tobias. *Optimum Experimental Designs, with SAS*. Oxford Statistical Science Series. Oxford University Press, United Kingdom, 2007.
- [60] Holger Dette. Designing experiments with respect to ‘standardized’ optimality criteria. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 59(1):97–110, 1997.
- [61] J López-Fidalgo, MJ Rivas-López, and B Fernández-Garzón. A-optimality standardized through the coefficient of variation. *Communications in Statistics?Theory and Methods*, 36(4):781–792, 2007.
- [62] John P. Morgan and Xiaowei Wang. Weighted optimality in designed experimentation. *Journal of the American Statistical Association*, 105(492):1566–1580, 2010.
- [63] Jonathan W Stallings and JP Morgan. General weighted optimality of designed experiments. *Biometrika*, 102(4):925–935, 2015.

- [64] C Presse and Maxime Gautier. New criteria of exciting trajectories for robot identification. In *Robotics and Automation, 1993. Proceedings., 1993 IEEE International Conference on*, pages 907–912. IEEE, 1993.
- [65] Maxime Gautier and Wisama Khalil. Exciting trajectories for the identification of base inertial parameters of robots. *The International Journal of Robotics Research*, 11(4):362–375, 1992.
- [66] Andrew D Wilson, Jarvis A Schultz, and Todd D Murphey. Trajectory synthesis for fisher information maximization. *IEEE Transactions on Robotics*, 30(6):1358–1370, 2014.
- [67] Bryan J. Martin and James E. Bobrow. Minimum-effort motions for open-chain manipulators with task-dependent end-effector constraints. *The International Journal of Robotics Research*, 18(2):213–224, feb 1999.
- [68] Sung-Hee Lee, Junggon Kim, F.C. Park, Munsang Kim, and J.E. Bobrow. Newton-type algorithms for dynamics-based robot movement optimization. *IEEE Transactions on Robotics*, 21(4):657–667, aug 2005.
- [69] Ko Ayusawa and Eiichi Yoshida. Comprehensive theory of differential kinematics and dynamics towards extensive motion optimization framework. *The International Journal of Robotics Research*, 37(13-14):1554–1572, 2018.
- [70] John J Craig. *Introduction to robotics: mechanics and control*, volume 3. Pearson Prentice Hall Upper Saddle River, 2005.
- [71] John J Craig, Ping Hsu, and S Shankar Sastry. Adaptive control of mechanical manipulators. *The International Journal of Robotics Research*, 6(2):16–28, 1987.

- [72] Jean-Jacques E Slotine and Weiping Li. On the adaptive control of robot manipulators. *The international journal of robotics research*, 6(3):49–59, 1987.
- [73] G Feng and M Palaniswami. Adaptive control of robot manipulators in task space. *IEEE Transactions on Automatic Control*, 38(1):100–104, 1993.
- [74] RH Middleton and GC Goodwin. Adaptive computed torque control for rigid link manipulators. In *1986 25th IEEE Conference on Decision and Control*, pages 68–73. IEEE, 1986.
- [75] Mark W Spong and Romeo Ortega. On adaptive inverse dynamics control of rigid robots. *IEEE Transactions on Automatic Control*, 35(1):92–95, 1990.
- [76] Petros A Ioannou and Jing Sun. *Robust adaptive control*, volume 1. PTR Prentice-Hall Upper Saddle River, NJ, 1996.
- [77] Weiping Li and J-JE Slotine. Indirect adaptive robot control. In *Proceedings. 1988 IEEE International Conference on Robotics and Automation*, pages 704–709. IEEE, 1988.
- [78] Hanlei Wang and Yongchun Xie. On the uniform positive definiteness of the estimated inertia for robot manipulators. In *Proceedings of the 18th IFAC World Congress*, pages 4089–4094. Milano, Italy, 2011.
- [79] Graham C Goodwin and David Q Mayne. A parameter estimation perspective of continuous time model reference adaptive control. *Automatica*, 23(1):57–70, 1987.
- [80] Shun-Ichi Amari. Natural gradient works efficiently in learning. *Neural computation*, 10(2):251–276, 1998.

- [81] Tayoon Lee and Frank C. Park. A geometric algorithm for robust multibody inertial parameter identification. *IEEE Robotics and Automation Letters*, 3(3):2455–2462, July 2018.
- [82] Taeyoon Lee, Patrick M. Wensing, and Frank C. Park. Geometric robot dynamic identification: A convex programming approach. *IEEE Transactions on Robotics*, 2019, To appear.
- [83] Hirohiko Shima and Katsumi Yagi. Geometry of hessian manifolds. *Differential Geometry and its Applications*, 7, 09 1997.
- [84] L.M. Bregman. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *USSR Computational Mathematics and Mathematical Physics*, 7(3):200 – 217, 1967.
- [85] J-JE Slotine. Putting physics in control—the example of robotics. *IEEE Control Systems Magazine*, 8(6):12–18, 1988.
- [86] Gunter Niemeyer and J-JE Slotine. Performance in adaptive manipulator control. In *Proceedings of the 27th IEEE Conference on Decision and Control*, pages 1585–1591. IEEE, 1988.
- [87] Marcio S de Queiroz, Darren M Dawson, and M Agarwal. Adaptive control of robot manipulators with controller/update law modularity. *Automatica*, 35(8):1379–1390, 1999.
- [88] Miomir Vukobratovic and Atanasko Tuneski. Adaptive control of single rigid robotic manipulators interacting with dynamic environment—an overview. *Journal of Intelligent and Robotic Systems*, 17(1):1–30, 1996.

- [89] J-J Slotine and Weiping Li. Adaptive strategies in constrained manipulation. In *Proceedings. 1987 IEEE International Conference on Robotics and Automation*, volume 4, pages 595–601. IEEE, 1987.
- [90] Christian Ott, Alin Albu-Schaffer, and Gerd Hirzinger. Comparison of adaptive and nonadaptive tracking control laws for a flexible joint manipulator. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 2, pages 2018–2024. IEEE, 2002.
- [91] Hanlei Wang and Yongchun Xie. Passivity based adaptive jacobian tracking for free-floating space manipulators without using spacecraft acceleration. *Automatica*, 45(6):1510–1517, 2009.
- [92] Sham M Kakade. A natural policy gradient. In *Advances in neural information processing systems*, pages 1531–1538, 2002.
- [93] Jan Peters and Stefan Schaal. Natural actor-critic. *Neurocomputing*, 71(7-9):1180–1190, 2008.
- [94] John Schulman, Sergey Levine, Pieter Abbeel, Michael Jordan, and Philipp Moritz. Trust region policy optimization. In *International Conference on Machine Learning*, pages 1889–1897, 2015.
- [95] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.
- [96] Xue Bin Peng, Marcin Andrychowicz, Wojciech Zaremba, and Pieter Abbeel. Sim-to-real transfer of robotic control with dynamics randomization. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1–8. IEEE, 2018.

- [97] Chandler Davis. All convex invariant functions of hermitian matrices. *Archiv der Mathematik*, 8(4):276–278, 1957.
- [98] Arthur M Geoffrion. Duality in nonlinear programming: a simplified applications-oriented development. *SIAM review*, 13(1):1–37, 1971.

국문초록

다물체 기계시스템은 모델 기반 계획과 제어를 필요로 하는 다수의 입력-출력 시스템들을 이룬다. 최근 들어, 휴머노이드나 4족 보행 로봇과 같은 복잡하고 고차원의 로봇들이 매우 동적인 임무와 동작들을 수행하기 시작하면서 로봇틱스 분야에서 모델 기반 계획과 제어 기법들이 더욱 주목을 받고 있다. 이런 모델 기반 방법들의 성능이 모델 추정치의 정확도와 직결되는 반면, 복잡하고 고차원의 로봇들에 대한 모델 추정은 일반적으로 수많은 어려움을 야기한다.

본 논문에서는 이런 어려움이 좌표계에 불변한 미분 기하학적인 접근으로 해결될 수 있음을 입증한다. 주 요점은 질량 관성 파라미터와 같은 물리적 값을 지닌 파라미터들이 휘어진 리만 공간상의 요소로 확인됨에 따라 물리적으로 자연스러운 거리 측량이 가능해짐에 있다. 이런 기하학적인 접근에 기반을 두어, 다양한 단계와 상황들에 있어서 강건한 모델 파라미터 추정을 가능하게 하는 좌표계에 불변한 기하학적인 알고리즘들을 제안하였다.

먼저, 비 실시간 파라미터 추정을 위한 새로운 기하학적 문제 정의와 그에 따른 효율적인 알고리즘들을 제안하였다. 특히, 리만 거리 함수의 이차 근사 함수들을 이용하여 기하학적 파라미터 추정 문제를 볼록 최적화 문제로 재정의 하였다. 이로써 빠르고 유일해로의 수렴성이 보장되는 볼록 최적화 기법들의 사용이 가능해짐뿐만 아니라, 좌표계 불변의 성질이 보존되며, 추가적인 볼록 구속 조건의 도입이 용이해지게 된다. 로봇팔 부터 보행로봇 그리고 인체 모델에 이르는 제한된 센서 측정 값을 가지는 고차원의 다물체 시스템들에 대해 제안된 방법들을 심도 있게 검증하였으며, 기존의 벡터 공간상의 방법들에 비하여 추정치들의 강건성과 보편성이 크게 증가함을 확인할 수 있었다.

다음으로, 파라미터 추정에 필요한 최적 궤적 데이터 생성에 대한 문제를 다루었다. 궤적 데이터들의 정보량을 좌표계 불변한 방식으로 측정할 수 있는 새로운 표준

함수들을 정의하였다. 따라서 제안된 표준함수로 최적화된 궤적 데이터들은 항상 좌표계에 불변하다. 또한, 최적 궤적 생성을 위한 효율적이고 강건한 구배 기반 수치적 최적화 알고리즘을 제시하였다. 또한, 제안된 기하학적 방법을 기반으로, 주어진 궤적데이터로부터 효과적으로 추정 가능한 파라미터들을 식별하는 좌표계 불변한 알고리즘을 고안하였다. 이는 매우 제한된 궤적의 구동만 가능한 휴머노이드와 같은 고차원 로봇의 파라미터 추정문제에 특별히 용이하다. 산업용 로봇과 휴머노이드 로봇에 대한 수치적 실험을 통해 제안된 방법을 검증한 결과 파라미터 추정의 강건성과 정확도가 크게 향상됨을 보였다.

마지막으로, 실시간 파라미터 추정이 수반되며 페루프 제어 시스템의 안전성까지 고려되어야 하는 로봇 적응 제어 기법을 위한 기하학적 알고리즘을 제안하였다. 먼저, 질량-관성 파라미터의 리만 다양체 구조를 더 일반적인 물리적 파라미터 공간으로 확장 하였다. 그 다음으로, 좌표계 불변하고 기하학적으로 자연스러운 리아푸노프 함수를 설정함에 따라 자연스럽게 안정성이 보장되는 파라미터 적응법을 유도되었으며, 실제로 리만 공간상의 자연스러운 구배 흐름 방정식과 같은 꼴로 나타남을 확인할 수 있었다. 특히, 제안된 접근법은 계인 매트릭스를 직접 시행착오를 통해 과도하게 조율해야하는 기존 방법들의 한계점을 해결하였다. 나아가 임의의 볼록 구속 조건을 만족시키는 강건 적응 제어 기법으로 본 방법을 확장할 수 있었다. 7 자유도 로봇팔의 궤적 추적 적응제어 작업에 대한 수치적 실험과 실제 실험을 통해 제안된 방법을 검증하였으며, 계인 수치의 과도한 조율 없이도 추적 에러가 크게 감소하는 것을 확인할 수 있었다.

주요어: 시스템 파라미터 추정, 적응 제어, 최적 실험 설계, 최적 궤적 생성, 볼록 최적화, 행렬 다양체, 리만 기하학

학번: 2015-20749