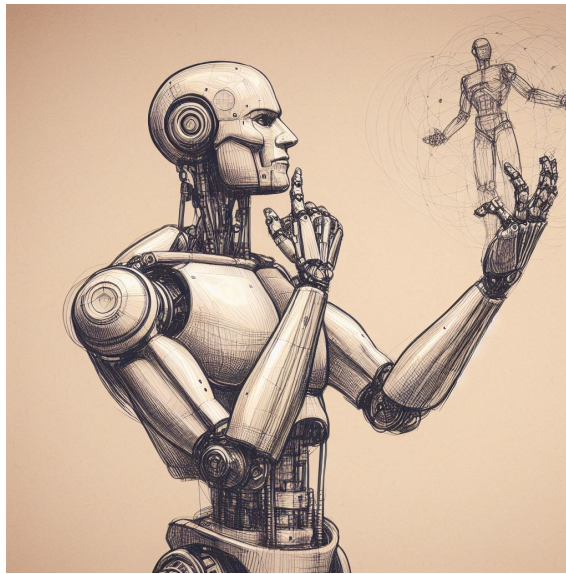


# Thesis structure: Fernando Díaz Ledezma

**Title:** Learning The-Self: Leveraging Proprioception to Guide the Autonomous Discovery of the Robot Body Schema

**Titel:** **TODO**



## Open TODOs

- ☒ Title
  - ☒ English
  - ☐ Deutsch
- ☐ Table of contents
- ☒ Abstract
- ☒ Summary BIB
  - ☒ English
  - ☐ Deutsch

- ☐ Nomenclature
- ☒ Introduction
  - ☒ Motivation
  - ☒ Problem statement
  - ☒ State of the art
  - ☒ Research questions
  - ☒ Contributions
  - ☒ Impact
- ☐ Chapter Introduction/Conclusion
- ☐ Conclusion
  - ☐ Contribution
  - ☐ Impact
  - ☐ Future work
- ☐ Feedback rounds with Sami for the chapters
- ☐ Final check

## Table of Contents

## Abstract

### Vision

- For future robots, the seamless integration of the body schema stands as a foundational pillar that fosters learning, motor control, coordination, and advanced spatial awareness that improves their versatility and seamless interaction with their surroundings.
- As robots develop and steadily permeate many aspects of human life, they need actively engage in the exploration and development of models for their own bodies, i.e. autonomous self-discovery of their body schema.
- Inspired by humans, future robots should be able to skillfully employ their body schema for advanced locomotion and motion planning, precise grasping, intricate object manipulation, and to anticipate and adapt the interaction with other agents.
- Constant self-monitoring of the sensorimotor state and the internal body models becomes the norm for instantaneous error detection and correction. These

models can adapt steadily to different situations developing a spatial awareness of the physical self that enables the rapid planning and deployment of contingent motion strategies providing advanced interaction capabilities with the environment.

- Robots will be self-sufficient to perform monitoring, calibration, and adaptation of their body representation relying only on onboard sensing capabilities. Fundamental modalities will include somatosensation (proprioception and touch) and vision.
- Understanding their own body structure enables robots to interact more effectively with other robots and with humans by adjusting movements for safety. Additionally, robots can optimize their energy consumption by adapting their motions based on physical properties, contributing to energy-aware robotics.

## Challenges

1. **Reliance on External Measurements:** Calibration and identification heavily depend on off-robot measurement devices, such as vision and motion-capturing systems, to discern kinematic structure properties. Despite various sensor signals from modern robots, determining the minimum set for constructing a body model based on robot sensing only remains unresolved.
2. **Limitations of Current Robot Learning Approaches:** Many current local and global machine learning frameworks for physical systems exclude structural knowledge and suffer from limited generalization capabilities and low sample efficiency. Additionally, learning a robot’s physical attributes parts from the assumption of a known mechanical topology and is often exclusive to calibration and offline parametric identification routines performed in controlled spaces (laboratories).
3. **Challenges in Learning Methods:** Many alternative learning methods, like neural networks, lack information about the body structure and require substantial data. Designing neural networks presents challenges in determining topology, and most data-based methods suffer from generalization limitations, confining learning to specific input-output regions.
4. **Research Gaps and Unifying Scheme:** There are significant gaps in research, including unclear understanding of how object handling extends the robotic body schema and limited exploration of the mechanical arrangement of joints and links (mechanical topology). Additionally, there is a lack of a unifying scheme to integrate all learning stages for a fully characterized robotic body schema solely from knowledge about sensorimotor signals.

## Contribution

This work:

- Consolidates the necessary and sufficient proprioceptive signal quantities (afferent and efferent sensory inputs and commands) that enable robots to autonomously acquire, monitor, and adapt knowledge about their body structure and decouple them from the need for exteroceptive off-body sensors.
- Reformulates robot kinematic calibration and parametric robot system identification as a computational graph whose topology reflects a modular structure amenable to machine learning. The architecture of this graph is abstracted into a pipeline consisting of a sequence of online learning phases where streams of proprioceptive signals are merged with first-order principles, imposed by the system’s embodiment, to enable the extraction of fundamental features of the robot body schema.
- Characterizes essential morphological properties of the broad class of tree-like floating base structures by studying the relationships among fundamental proprioceptive signals. The mechanical topology, i.e., the arrangement of links and joints, is initially inferred using model-free information-theoretic measures. Consequently, this topology is concurrently validated and employed to instantiate the kinematic description of the robot’s body independent of exteroceptive off-robot calibration devices.
- Complements the description of the robot body schema by instantiating the fundamental inertial properties of the links composing the inferred morphology. Given that these properties lie on the Riemannian manifold of symmetric positive definite matrices, a method is introduced to learn them online while ensuring physical feasibility at all times.

## Impact

- While acknowledging the undeniable versatility and representational power of current end-to-end learning approaches, this work incites to reconsider their naive application to physical systems and promote the assessment of their limitations when they deliberately exclude principled knowledge. In contrast, the arguments and findings presented here reveal avenues for machine learning frameworks for embodied systems. This research exposes the untapped potential arising from the synergistic integration of existing structural knowledge with data-driven method.
- The outlined concepts and methods demonstrate that crucial aspects of a robot’s body schema can be deduced through a fundamental set of proprioceptive signals. As future mobile robots are anticipated to feature a diverse, enhanced, and reliable array of on-board sensing modalities, extending beyond proprioception, the findings discussed in this thesis serve as a catalyst for research into the integration of these modalities. This integration, coupled with the online learning of body morphological and dynamic properties, holds the promise of refining and adapting body models, ultimately empowering robots with heightened levels of autonomy.

- This study contributes to an emerging research area that underscores building and maintaining a body schema as a crucial capability for embodied systems. Such a capability pertains robots characterized by conventional, immutable structures and a novel category of mechanical systems exhibiting dynamic morphologies and diverse multimodal sensory modalities. These systems will evolve their sense of self, recognizing the affordances inherent in their bodies..

## Abstract (text version)

As robots become increasingly integral to human life, the imperative emerges for them to autonomously explore and construct models of their bodies. Robots should take cues from human capabilities, aspiring to build and utilize their body schema for advanced locomotion, finer manipulation, and adaptive interactions. Thus, a crucial foundation lies in seamlessly integrating the body schema to elevate learning, motor control, coordination, and spatial awareness. Furthermore, future robots should become self-sufficient entities that conduct monitoring, calibration, and adaptation exclusively through onboard sensing modalities. Standardizing constant self-monitoring nurtures spatial awareness and facilitates rapid error detection and correction. A profound understanding of their body structure will undoubtedly lead to enhanced, safe, and energy-aware interactions. However, current robot learning approaches encounter limitations, such as suboptimal generalization and sample efficiency, exhibiting a need for more structural knowledge. Versatile methods, like neural networks, confront challenges related to data and topology, confining learning to specific regions. On the other hand, learning robot physical attributes still rely on a presumed knowledge of the mechanical topology, often involving calibration and offline identification in controlled environments with a persistent reliance on external measurements, such as vision and motion-capturing systems. The research landscape generally reveals the lack of a unified framework that enables robots to build representations of their body schema to achieve improved body awareness and interaction capabilities. This study addresses these challenges by consolidating necessary and sufficient proprioceptive signal quantities, enabling robots to autonomously acquire knowledge about their body structure without relying on exteroceptive off-body sensors. It introduces an approach that reformulates robot kinematic calibration and system identification as a modular computational graph amenable to machine learning. This abstracted architecture, applied in online learning phases, seamlessly merges proprioceptive signals with first-order principles, extracting fundamental features of the robot body schema. Characterizing morphological properties of tree-like structures, the study infers mechanical topology through information-theoretic measures, validating and applying it independently of off-robot calibration. The research extends its scope by complementing the robot body schema by instantiating inertial properties, ensuring online learning and physical feasibility. Ultimately, this work challenges the uncritical application of end-to-end learning in physical systems, urging a reevaluation of its limitations when excluding principled knowledge. It underscores opportunities for machine learning frameworks in embodied systems, emphasizing the untapped potential of synergizing structural knowledge with data-driven methods. This study catalyzes future research in an incipient field

that underscores building and maintaining a body schema by demonstrating that fundamental properties of a robot’s morphology can be deduced from proprioceptive signals. Its implications are far-reaching, addressing the needs of conventional and dynamic robotic structures with diverse sensory modalities that require a more profound sense of self.

## Summary for BIB (English)

This thesis explores the potential for enhanced robot autonomy through a self-discovery-oriented body schema, proposing a unified online learning framework exclusively reliant on proprioception and leveraging structural knowledge. It infers the robot morphology and associated inertial description. The work urges reconsidering end-to-end learning for physical systems, emphasizing the need for a synergistic integration of principled knowledge and sensorimotor data.

## Kurzzusammenfassung für BIB (Deutsch)

**TODO**

---

# Introduction

## Motivation

### 1. Empowering Robots

- Autonomous self-discovery is imperative for robots integrating into human life.
- Awareness of the physical self through the body schema is foundational.
- It enables the integration of sensory information and motor control.
- The evolving body schema serves as a dynamic map for interactions.
- Enhances robot motor control, precision, and coordination.
- Facilitates efficient learning, adapting to diverse environments.

### 2. Learning and the Body Schema

- The body schema is indispensable for multifaceted robot capabilities.
- Learning contributes to body schema development, forming a dual relationship.
- Detects structure in sensorimotor signals, aiding body schema construction.
- Incorporating body schema into learning refines skills and assimilates knowledge.
- Enhances motor control through adaptive internal body representations.
- Empowers robots to learn diverse tasks, providing versatility in dynamic settings.

### 3. Enhancing Locomotion, Manipulation, and Adaptability

- A well-integrated body schema improves adaptability and interaction.
- Enables precise and coordinated movements, advanced locomotion, and motion planning.
- Enhances manipulation capabilities with human-like dexterity and precision.
- Coordination with other agents, both robots and humans, becomes more refined.
- Anticipatory and adaptive capabilities are fundamental for safe and effective interactions.

### 4. Constant Self-Monitoring for Autonomy

- Continuous self-monitoring is fundamental for future robotic systems.
- Achieved through internal models and uninterrupted sensorimotor signals.

- Enables dynamic, real-time understanding of the robot's state.
- Successive error detection and correction phases enhance reliability.
- Rapid formulation and execution of contingency motion strategies in dynamic environments.

#### 5. Onboard Sensing for Self-Sufficiency

- True autonomy requires robots to rely exclusively on onboard sensing.
- Somatosensation (proprioception and touch) and vision are fundamental modalities.
- Liberates robots from external dependencies, enhancing self-sufficiency.
- Enables dynamic responses to changes in surroundings in real-time.
- Enhances autonomy and adaptability, previously unseen with off-board sensing.

#### 6. Safety- and Energy-Awareness

- The body schema serves as a predictive tool, fostering safety in interactions.
- Facilitates dynamic adjustments in movements to prioritize safety.
- Enables seamless coordination with other robots and humans, averting collisions.
- Comprehension of body structure optimizes energy consumption.
- Dual capability enhances safety and contributes to energy-aware robotics, fostering efficiency and collaboration.



## Problem Statement

- Prevalent challenges in robotics model learning include overlooking crucial structural knowledge, focusing primarily on global or local forward and inverse models without nuanced integration of essential structural information.
- Challenges faced by machine learning frameworks involve limitations in global techniques like Gaussian process regression and sensitivity in local methods like
- Locally Weighted Projection Regression, emphasizing the need for a balanced approach between data-driven and knowledge-driven methods.
- The dissertation aims to infer essential morphological properties of tree-like floating base structures, emulating the development of a body schema. Efforts in cognitive robotics stress the role of internal body models, enhancing spatial awareness, motor control, and adaptability, but lack consensus on defining the robot’s body schema.
- Classical robotics methods assume a known mechanical topology for identifying physical attributes, relying on calibration routines and offline system identification. However, these methods face challenges in floating base robots, and their integration into online learning frameworks is not well-established, limiting insights into the mechanical topology.
- Research gaps include the absence of a clear definition and construction method for the robotic schema, challenges in determining a fundamental set of signals for body schema construction, limited exploration of the mechanical topology, and a lack of synergy between engineering approaches and online learning for realizing a first robot body schema.
- The intricate connections between sensorimotor regularities and body structural knowledge in robotics remain poorly understood, requiring methods with plasticity to adapt to different motion policies and accurately reflect these effects. Addressing these gaps is crucial for advancing robotics and developing more sophisticated and adaptable embodied systems

## Assembled version

In robotics model learning, a pressing challenge is the prevalent oversight of incorporating crucial structural knowledge. Current efforts primarily concentrate on developing forward and inverse models, adopting either a global or local emphasis, often neglecting the nuanced integration of essential structural information [25]. Challenges faced by prevalent machine learning frameworks include the limitations of global techniques like Gaussian process regression, influenced by the curse of dimensionality, and the sensitivity of local methods such as Locally Weighted Projection Regression to hyperparameters. Despite the growing computational power and data availability favoring end-to-end learning methods, deep learning in particular, the

pervasive neglect of prior principled knowledge poses essential challenges. Determining dedicated neural network architectures remains elusive [2], [6]. Similarly, deep learning approaches, like other learning techniques, grapple with issues of low sample efficiency, extensive training times, and limited generalization, emphasizing the need for a balance between data-driven and knowledge-driven approaches within the field [26], [29].

Parting from the need for structure in robot model learning, **this dissertation addresses the problem of inferring essential morphological properties of the broad class of tree-like floating base structures that emulate the development of a body schema.** Similar efforts have been discussed in cognitive robotics, emphasizing the pivotal role of developing and maintaining internal body models for embodied systems. Indeed, enhanced understanding of the physical self through developing and integrating a body schema improves spatial awareness, motor control, and adaptability [12], [24]. **Yet, consensus is missing on what constitutes the body schema for a robot. Several existing endeavors in learning a body schema in robotics center exclusively in learning the particulars of the kinematic structure and predominantly rely on (off-body) vision [5], [10], [11], [19], [22], [28].** Alternative efforts lay their focus on exploring the sensorimotor associations between proprioceptive, tactile and visual sensory modalities [7], [15], [21], [23], [27] but **provide limited insights into the robot’s embodiment, in particular its physical structure.**

Classical robotics provides conventional methods for identifying the physical attributes of robots, typically assuming a known mechanical topology. Established techniques, such as conventional calibration routines [13] and offline system identification methods [16], [30], have proven effective for known kinematic structures in controlled environments. However, these methods encounter challenges when applied to floating base robots, where standardization in identification procedures is lacking [1], [17]. Importantly, **these conventional methods were not originally designed for integration into online learning frameworks.** While in addition to kinematic calibration classic robotics tackles forward and inverse kinematics problems, **it provides limited insights into a comprehensive understanding of the mechanical arrangement of joints and links, referred to as the mechanical topology.** Circling back to cognitive robotics, only a limited number of studies have approached the latter problem as a means to enhance self-modeling and monitoring, yet again depending on exteroceptive vision [3], [4]. Last, be it the black-box machine learning, cognitive, or classical robotics paradigms, there is a strong dependence on external devices like cameras and motion-capturing systems, neglecting on-robot sensing. Despite modern robots having various sensor signals, determining the minimum set for constructing a body model solely from on-robot sensing remains unresolved.

In summary, research in model learning for robotics reveals critical gaps hindering the construction and understanding of a comprehensive robot body schema. There needs to be a clear definition for the robotic schema and a method for its construction, coupled with the challenge of determining a fundamental set of proprioceptive and (on-body) exteroceptive signals essential for constructing a body schema, underscores the current limitations. Integrating cutting-edge machine learning techniques with prior information and first-order principles from mechanics offers a promising avenue to address challenges in data requirements, enhance generalization capabili-

ties, and gain deeper insights into body structure and properties. However, limited exploration into the mechanical topology, a lack of synergy between engineering approaches and online learning for realizing a first robot body schema, and the absence of a unifying scheme defining relevant learning stages remain notable gaps. Furthermore, the intricate connections between sensorimotor regularities and body structural knowledge in robotics remain poorly understood, necessitating methods with plasticity to adapt to different motion policies and accurately reflect these effects. Addressing these gaps is crucial for advancing the field of robotics, enabling the development of more sophisticated and adaptable embodied systems.

---

In the realm of robotics model learning, a pressing challenge is the prevalent oversight of incorporating crucial structural knowledge. Current efforts primarily concentrate on developing forward and inverse models, adopting either a global or local emphasis, often neglecting the nuanced integration of essential structural information [25]. Challenges faced by prevalent machine learning frameworks include the limitations of global techniques like Gaussian process regression, influenced by the curse of dimensionality, and the sensitivity of local methods such as Locally Weighted Projection Regression to hyperparameters. Despite the growing computational power and data availability favoring end-to-end learning methods, the pervasive neglect of prior principled knowledge poses recurrent challenges. Determining dedicated neural network architectures remains elusive [2], [6]. Similarly, deep learning approaches, like other learning techniques, grapple with issues of low sample efficiency, extensive training times, and limited generalization, emphasizing the need for a balance between data-driven and knowledge-driven approaches within the field [26], [29].

Additionally, the learning of a robot’s physical attributes, fundamental for effective model learning, often assumes a known mechanical topology. Conventional calibration routines and offline system identification methods, well-established for known kinematic structures in controlled environments, face challenges when applied to floating base robots, where identification procedures lack standardization [1], [13], [16], [17], [30]. Classical robotics research, while addressing kinematic calibration and standard inverse kinematics problems, offers limited insights into comprehensively learning the mechanical arrangement of joints and links, known as the mechanical topology.

Simultaneously, a burgeoning research area emphasizes the pivotal role of developing and maintaining internal body models for embodied systems in robotics. Learning the physical self and cultivating a body schema enhance spatial awareness, motor control, and adaptability [12], [24]. Existing achievements in learning a body schema predominantly rely on off-body vision, with limited exploration of other modalities such as tactile inputs. While some studies integrate vision and tactile input or consider various sensory modalities together, a notable scarcity of research delves into the need to learn the mechanical topology of the robot, hindering a comprehensive understanding of the robot’s physical structure beyond sensory data [3]–[5], [7], [8], [10], [11], [15], [18], [19], [21]–[23], [27], [31].

This research aims to address identified challenges and gaps in existing machine learning frameworks for model learning in robotics, emphasizing the nuanced integration of structural knowledge. Furthermore, it seeks to advance the understanding of robot physical attributes, specifically focusing on the mechanical topology, proposing innovative methodologies beyond conventional calibration and system identification. The investigation extends into the realm of learning the physical self and the body schema, exploring modalities beyond vision and tactile inputs. By doing so, this research endeavors to bridge the identified gaps, fostering the development of more robust and adaptive robotic systems in diverse environments.

In contemporary robotics research, the field of model learning grapples with critical challenges, notably the oversight of incorporating structural knowledge. Current efforts predominantly focus on the development of forward and inverse models, exhibiting either a global or local emphasis, neglecting the nuanced integration of es-

sential structural information [25]. The limitations of global techniques like Gaussian process regression, stemming from the curse of dimensionality, and the sensitivity of local methods such as Locally Weighted Projection Regression to hyperparameters underscore the complexities faced by prevalent machine learning frameworks. Despite the increasing computational power and data availability favoring end-to-end learning methods, the pervasive neglect of prior principled knowledge poses recurrent challenges. Determining dedicated neural network architectures remains elusive [2], [6], and the extensive training times and limited generalization of deep learning approaches highlight the imperative need for a nuanced balance between data-driven and knowledge-driven approaches in robotics [26], [29].

Moreover, the learning of a robot’s physical attributes, a fundamental aspect for effective model learning, often assumes a known mechanical topology. Conventional calibration routines and offline system identification methods, although well-established for known kinematic structures in controlled environments, face challenges when applied to floating base robots, where identification procedures lack standardization [1], [13], [16], [17], [30]. Classical robotics research, while addressing kinematic calibration and standard inverse kinematics problems, offers limited insights into comprehensively learning the mechanical arrangement of joints and links, known as the mechanical topology.

In parallel, a burgeoning research area emphasizes the development and maintenance of internal body models as a pivotal capability for embodied systems in robotics. Learning the physical self and cultivating a body schema enhance spatial awareness, motor control, and adaptability [12], [24]. However, existing achievements in learning a body schema predominantly rely on off-body vision, with limited exploration of other modalities such as tactile inputs. While some studies integrate vision and tactile input or consider various sensory modalities together, a notable scarcity of research delves into the need to learn the mechanical topology of the robot, hindering a comprehensive understanding of the robot’s physical structure beyond sensory data [3]–[5], [7], [8], [10], [11], [15], [18], [19], [21]–[23], [27], [31].

Identifying these challenges and gaps, the present research seeks to address the deficiencies in existing machine learning frameworks for model learning in robotics, focusing on the nuanced integration of structural knowledge. Additionally, it aims to advance the understanding of robot physical attributes, with a specific emphasis on the mechanical topology, proposing innovative methodologies that go beyond conventional calibration and system identification. The investigation extends into the realm of learning the physical self and the body schema, exploring modalities beyond vision and tactile inputs. By doing so, this research endeavors to bridge the identified gaps, fostering the development of more robust and adaptive robotic systems in diverse environments.

---

Recent endeavors in model learning within robotics often neglect the incorporation of structural knowledge, instead focusing on developing forward and inverse models with either a global or local emphasis [25]. While standard global machine learning techniques like Gaussian process regression face challenges related to the curse of dimensionality, rendering them computationally intractable, local methods such as Locally Weighted Projection Regression and support vector regression exhibit sensitivity to hyperparameters and struggle with generalization issues. The advent of

increased computational power and abundant data has elevated end-to-end learning methods, particularly with deep learning as the flagship. Despite their remarkable results, the widespread neglect of prior principled knowledge poses clear challenges. A recurring challenge is the difficulty of determining dedicated neural network architectures [2], [6]. Similarly, deep learning approaches, like other learning techniques, grapple with issues of low sample efficiency, extensive training times, and limited generalization, emphasizing the need for a balance between data-driven and knowledge-driven approaches within the field [26], [29]. In general, most data-driven learning methods face limitations in generalization, restricting learning to specific input-output regions.

The learning of a robot’s physical attributes often assumes a known mechanical topology, primarily addressed through calibration routines for known kinematic structures [13] and conventional offline system identification methods conducted in controlled environments [16], [30]. However, for floating base robots, unlike their fixed-base counterparts, identification procedures are less standardized [1], [17]. Classical robotics research, beyond addressing kinematic calibration and standard inverse kinematics problems, provides limited insights into learning the mechanical arrangement of joints and links that define the kinematic chain, referred to as the mechanical topology.

Calibration and identification processes in robotics rely extensively on off-robot measurement devices, including vision, laser metrology, and motion-capturing systems, to discern properties related to the kinematic structure. This strong dependence on external measurements underscores the complexity of achieving accurate calibration and identification through on-robot sensing alone. Moreover, despite the availability of various sensor signals from modern robots, the challenge of determining the minimum set required for constructing a body model based solely on robot sensing remains unresolved.

In the realm of robotics, there is a burgeoning research area emphasizing the development and maintenance of internal body models as a crucial capability for embodied systems, running parallel to current model learning and conventional system identification approaches. Learning the physical self and cultivating a body schema are pivotal for robotics, enhancing spatial awareness, motor control, and adaptability [24]. The body schema, functioning as a dynamic map, enables precise movements and facilitates efficient learning, allowing robots to navigate diverse environments effectively [12]. While recent achievements in learning a body schema for robots have been remarkable, they predominantly rely on (off-body) vision [5], [10], [11], [19], [22], with some works incorporating tactile inputs [8], [18], [31]. Few efforts explore combinations of modalities, such as vision and tactile input [7], somatosensation (proprioception and touch) [21], and consideration of all these modalities together [15], [23], [27]. However, very few works have delved into the need to learn the mechanical topology of the robot and the potential advantages it might bring [3], [4].

In the evolving landscape of robotics, a critical research gap exists, centering around the development and maintenance of internal body models. This area is gaining prominence alongside current model learning and traditional system identification

approaches. Learning the physical self and cultivating a body schema are recognized as pivotal for robotics, enhancing spatial awareness, motor control, and adaptability. Despite recent remarkable achievements in learning a body schema for robots, these efforts primarily lean on (off-body) vision, with limited exploration of other modalities such as tactile inputs. While some works integrate vision and tactile input or consider the combined use of various sensory modalities, there is a notable scarcity of studies delving into the need to learn the mechanical topology of robots and the potential advantages it might offer. Understanding the intricacies of the robot's physical structure, beyond sensory data, remains an underexplored frontier in the pursuit of advancing embodied systems. Bridging this research gap is crucial for the holistic development of robots with enhanced spatial intelligence and adaptive capabilities in diverse environments.

Recent endeavors in model learning within robotics research often overlook the incorporation of structural knowledge, instead focusing on the development of forward and inverse models with either a global or local emphasis [25]. While standard global machine learning techniques like Gaussian process regression face challenges related to the curse of dimensionality, becoming computationally intractable, local methods such as Locally Weighted Projection Regression and support vector regression exhibit sensitivity to hyperparameters and struggle with generalization issues. The advent of increased computational power and abundant data has elevated end-to-end learning methods, particularly with deep learning as the flagship. Despite their remarkable results, the widespread neglect of prior principled knowledge poses clear challenges. A recurrent one is the difficulty to determine dedicated neural networks architectures [2], [6]. Similarly, deep learning approaches, like other learning techniques, grapple with issues of low sample efficiency, extensive training times, and limited generalization, emphasizing the need for a balance between data-driven and knowledge-driven approaches within the field [26], [29]. In general, most data-driven learning methods face limitations in generalization, restricting learning to specific input-output regions.

The learning of a robot’s physical attributes often assumes a known mechanical topology, primarily addressed through calibration routines for known kinematic structures [13] and conventional offline system identification methods conducted in controlled environments [16], [30]. However, for floating base robots, unlike their fixed-base counterparts, identification procedures are less standardized [1], [17]. Classical robotics research, beyond addressing kinematic calibration and standard inverse kinematics problems, provides limited insights into learning the mechanical arrangement of joints and links that define the kinematic chain, referred to as the mechanical topology.

Calibration and identification processes in robotics rely extensively on off-robot measurement devices, including vision, laser metrology, and motion-capturing systems, to discern properties related to the kinematic structure. This strong dependence on external measurements underscores the complexity of achieving accurate calibration and identification through on-robot sensing alone. Moreover, despite the availability of various sensor signals from modern robots, the challenge of determining the minimum set required for constructing a body model based solely on robot sensing remains unresolved.

In the realm of robotics, there is a burgeoning research area emphasizing the development and maintenance of internal body models as a crucial capability for embodied systems, running parallel to current model learning and conventional system identification approaches. Learning the physical self and cultivating a body schema are pivotal for robotics, enhancing spatial awareness, motor control, and adaptability [24]. The body schema, functioning as a dynamic map, enables precise movements and facilitates efficient learning, allowing robots to navigate diverse environments effectively [12]. While recent achievements in learning a body schema for robots have been remarkable, they predominantly rely on (off-body) vision [5], [10], [11], [19], [22], with some works incorporating tactile inputs [8], [18], [31]. Few efforts explore combinations of modalities, such as vision and tactile input [7], somatosensation (proprioception and touch) [21], and consideration of all these modalities together



[15], [23], [27]. However, very few works have delved into the need to learn the mechanical topology of the robot and the potential advantages it might bring [3], [4].

In the evolving landscape of robotics, there exists a critical research gap that centers around the development and maintenance of internal body models, an area gaining prominence alongside current model learning and traditional system identification approaches. Learning the physical self and cultivating a body schema are recognized as pivotal for robotics, enhancing spatial awareness, motor control, and adaptability. Despite recent remarkable achievements in learning a body schema for robots, these efforts primarily lean on (off-body) vision, with limited exploration of other modalities such as tactile inputs. While some works integrate vision and tactile input or consider the combined use of various sensory modalities, there is a notable scarcity of studies delving into the need to learn the mechanical topology of robots and the potential advantages it might offer. Understanding the intricacies of the robot's physical structure, beyond sensory data, remains an underexplored frontier in the pursuit of advancing embodied systems. Bridging this research gap is crucial for the holistic development of robots with enhanced spatial intelligence and adaptive capabilities in diverse environments.

- **Challenges and limitations of current learning Approaches:**

1. Most of recent model learning endeavors in the robotics research has, unfortunately, excluded consideration of structural knowledge and rather centered around learning forward and inverse models with global or local focus [25].
2. Standard global machine learning techniques like Gaussian process regression suffer from the curse of dimensionality and rapidly become computationally intractable, alternative local methods such as Locally Weighted Projection Regression (and support vector regression) exhibit high sensitivity to hyperparameters and problematic generalization.
3. The rise in computational power and the availability of data has brought end-to-end learning methods to prominence. With deep learning as the flagship, these global methods have led to remarkable results. Yet, they have had the unwanted effect of making deliberately neglecting prior principled knowledge more widespread.
4. In spite the prowess and potentials of deep learning approaches, such exclusion of available prior knowledge makes it difficult to determine dedicated neural networks architectures (number of nodes and layers, connectivity, and activation functions) [2], [6] but, just like other learning techniques, deep learning approaches they still face problems related to the high demands of data to train (low sample efficiency), extensive training times, and performance that is tied to the training data, i.e. challenged generalization [26], [29].
5. In general, most data-driven learning methods suffer from generalization limitations, confining learning to specific input-output regions.

- **Limitations of conventional system identification:**

1. The learning of a robot’s physical attributes parts from the assumption of a known mechanical topology and is often exclusive to calibration routines for known kinematic structures [13] and conventional offline system identification methods performed in controlled spaces (laboratories). [16], [30].
2. In contrast to the conventional identification processes for fixed-base robots, the procedures for floating base robots are not as standardized [1], [17].
3. Besides kinematic calibration and standard inverse kinematics problems, classical robotics research offers limited insights on learning the mechanical arrangement of joints and links that define the essence of the kinematic chain, known as the mechanical topology.

- **Reliance on External Measurements:**

1. Calibration and identification heavily depend on off-robot measurement devices, such as vision and motion-capturing systems, to discern kinematic structure properties. Despite various sensor signals from modern robots, determining the minimum set for constructing a body model based on robot sensing only remains unresolved.
2. There is a strong dependence on off-robot measurement devices for calibration and identification, commonly involving exteroceptive measurements like vision, laser metrology, and motion-capturing systems, to determine the properties of the kinematic structure.

- **Inference of body knowledge:**

- In parallel to the current model learning and conventional system identification approaches there is a relatively young research area that underscores building and maintaining internal body models as a crucial capability for embodied systems
- Learning the physical self and developing a body schema are pivotal for robotics, enhancing spatial awareness, motor control, and adaptability [24]. The body schema serves as a dynamic map, enabling precise movements and fostering efficient learning. As robots encounter diverse environments, their adaptive body schema allows them to navigate real-world scenarios effectively [12].
- Recent remarkable achievements towards learning a body schema for robots have been made, yet they are strongly dependent but depend on (off-body) vision [5], [10], [11], [19], [22], [28]
- Some other works require tactile inputs [8], [18], [31]
- Few efforts explore combinations of these modalities, like the combination of vision and tactile input [7], somatosensation (proprioception and touch) [21], and the consideration of all these modalities [15], [23], [27]

- As stressed in current works [5], and in spite of the very much desired distributed and multimodal properties of a body schema, in robotics it is certainly relevant to capture the body morphology.
- Very few works have even contemplated the need to learn the mechanical topology of the robot and the advantages it might bring [3], [4]

• **Research Gaps:**

1. No clear definition of what the robotic schema is and how to construct it
2. Among the various proprioceptive and exteroceptive signals provided by modern robots' sensor suites, the determination of a fundamental set necessary to construct a body schema is yet to be established.
3. Identifying crucial research challenges in robotics involves delving into the integration of cutting-edge machine learning techniques with existing prior information and well-established first-order principles from mechanics, aiming to develop more effective and efficient learning algorithms for embodied systems. This area presents numerous opportunities for development, offering potential solutions to address data requirements, improve generalization capabilities, and gain deeper insights into body structure and properties. Recent contributions, such as the insightful work by Geist et al. [9] and the comprehensive study by Lutter et al. [20], have started to address these challenges in depth, shedding light on promising avenues for advancement in robotics research.
4. Unclear understanding of how object handling extends the robotic body schema
5. Limited exploration of the mechanical arrangement of joints and links (mechanical topology).
6. Leveraging engineering approaches for system identification and modern online learning techniques to provide a first realized robot body schema has not been extensively explored.
7. There is a lack of a unifying scheme that narrows the gap to define the relevant learning stages and their corresponding integration required to produce a robot body schema from knowledge about the sensorimotor signals and first-order principles that captures essential properties of the robot body (physical self).
8. Despite the general consensus that embodiment plays a crucial role in shaping the relationships among sensorimotor signals, the connections between sensorimotor regularities and body structural knowledge in robotics remain poorly understood [14]. A desired method addressing this gap should not only acknowledge the variability in statistical properties of signals but also demonstrate plasticity to adapt to different motion policies and accurately reflect these effects.

# Research Questions and Contributions

## Research questions

Overall the research questions addressed in this thesis pertain the learning of the robotic body schema, at least from the engineering perspective. In particular: -

**Q 1** *Which measurements are required to fully automate robot kinematics and inverse dynamics learning based on knowing only the adjacency graph along with kinematic and dynamic first-order principles?*

**Q 2** *How to transform robot system identification or end-to-end learning with meta parameter guessing into an automated learning scheme that determines both the structure and dynamical properties of the robot with minimal information?*

**Q 3** *How to leverage the inherent structure of the robot's sensorimotor system to gradually develop an understanding of the body structure despite being initially oblivious to its physical characteristics?*

## Contribution 1: Robot body structure as a learning problem

This thesis

1. Determines the type of proprioceptive signals and corresponding sensor requirements to enable robots to learn their body schema
2. Reformulates the classical kinematic calibration and parametric system identification of fixed base robots as an online learning problem that relies solely on the robot's proprioception and first-order principles from kinematic and dynamics
3. Discusses how embodiment and first-order principles (FOP) define network topologies of parameterized operators that model input-output mappings in robotic systems

## Contribution 2: Inferring the robot morphology

1. An application of classical gradient descent to learn three different representations of the robot kinematics; namely, modified Denavit-Hartenberg parameters, Euler angles, and angle axis representation
2. A demonstration that the given certain number of sensors with appropriate modalities the mechanical topology a tree-structure robot can be extracted by studying the mutual information among the signals

3. A method to infer the robot morphology, that is, the mechanical topology and the location and orientation of the robot's joint axes based only on the proprioceptive signals

### **Contribution 3: Online learning of physically feasible inertial parameters**

1. An offline learning (optimization) with constraints is presented to show that learning physically feasible inertial parameters of a manipulator can be done from joint data; i.e., joint position, velocity, acceleration, and torque
2. An online learning driven by state-of-the-art gradient descent method to facilitate the online learning of feasible inertial parameters applied to floating base robots
3. Introduce the Riemannian AMS gradient descent method, an optimization method for online learning on the manifold of symmetric positive definite matrices to guarantee the physical feasibility of the parameters at all times during the learning process

## **Overview of the Content**

The thesis discusses four main topics:

1. **Model learning and body schema.** Introduction of the fundamentals of robotic calibration and system identification and their relation to the concept of body schema. The chapter elaborates on the different meanings of the body schema and the definition applying within the context of this thesis is provided. Finally the learning stages to characterize the robotic body schema from an engineering perspective are introduced and discussed.
2. **Inferring the mechanical topology.** In this section, the concept of embodiment is presented and its significance to finding the robot structure is discussed. The fundamental idea that analyzing the relationships among the proprioceptive signals of a robot can convey information about the body structure as a result of embodiment is presented. Mutual information is pushed forward as a tool to unveil the mechanical topology of a robot given the right proprioceptive signals
3. **Characterizing the kinematic structure.** This chapter extends the classical exteroception-based kinematic calibration methods with proprioception-based online learning. Departing from the conventional assumption that the mechanical topology is known, it is discussed how the combination of mutual information basic differential kinematic laws can be used to characterize the location and orientation of the robot joint axes.
4. **Learning the inertial properties.** This section delves into the well established methods for robot inertial parameter identification and presents gradient-based online learning methods to produce valid sets of parameters.

In particular the fundamental property that the inertial parameters lie on the manifold of symmetric positive definite matrices is exploited to present a Riemannian gradient descent method that operates on this manifold.

## State of the art

## Impact

See abstract

---

# Ch. Introduction

## General

- The concept of body schema
- The body schema in robotics
- The body schema learning problem.
- Related work (State of the Art)

## Motivation

- Objectives of the research.
- Research questions or hypotheses.
- Significance of the study.

## Research questions and contribution

**TODO**

## The body schema learning problem

- The body schema
- The body learning problem
- Related work
- Different approaches to learn body properties
- Open research problems

- Contribution

## Conclusion

## Ch. Theoretical Framework

- The body schema in neuroscience
- The body schema in robotics
- Sensorimotor learning in robotics
  - Fundamentals
  - Taxonomy
    - Artificial neural networks
  - Statistical learning
    - Probabilistic learning
  - Decomposition
- Discussion
- The robot proprioceptive signals

## Introduction

- Robot kinematics
- Robot dynamics
- A modular view on learning

## Conclusion

## Ch. A Learning Perspective on the Inertial Parameters

## Introduction

- Classical system identification approach
- The advantages of online learning
- Relation to adaptive control
- The power of gradient descent

- Differential geometry
- Learning the inertial parameters the right way

## Conclusion

## Learning the kinematic description

## Introduction

## Conclusion

## The robot body topology

## Introduction

## Conclusion

## Conclusion

- One potential application area: Self-discovery in robots is crucial for applications in prosthetics and wearable robotics, allowing devices to align with the user’s body for natural and comfortable support.

## References

- [1] K. Ayusawa, G. Venture, and Y. Nakamura, “Identifiability and identification of inertial parameters using the underactuated base-link dynamics for legged multibody systems,” *The International Journal of Robotics Research*, vol. 33, no. 3, pp. 446–468, 2014.
- [2] B. Baker, O. Gupta, N. Naik, and R. Raskar, “Designing neural network architectures using reinforcement learning,” in *5th International Conference on Learning Representations*, 2017.
- [3] J. Bongard, V. Zykov, and H. Lipson, “Automated synthesis of body schema using multiple sensor modalities,” in *Proc. of the Int. Conf. on the Simulation and Synthesis of Living Systems (ALIFEX)*, Citeseer, 2006.
- [4] J. Bongard, V. Zykov, and H. Lipson, “Resilient machines through continuous self-modeling,” *Science*, vol. 314, no. 5802, pp. 1118–1121, 2006.
- [5] B. Chen, R. Kwiatkowski, C. Vondrick, and H. Lipson, “Fully body visual self-modeling of robot morphologies,” *Sci. Robotics*, vol. 7, no. 68, 2022. DOI: 10.1126/SCIROBOTICS.ABN1944.



- [6] T. Elsken, J. H. Metzen, and F. Hutter, “Neural architecture search: A survey,” *The Journal of Machine Learning Research*, vol. 20, no. 1, pp. 1997–2017, 2019.
- [7] S. Fuke, M. Ogino, and M. Asada, “Body image constructed from motor and tactile images with visual information,” *Int. J. Humanoid Robotics*, vol. 4, no. 2, pp. 347–364, 2007. DOI: 10.1142/S0219843607001096.
- [8] F. Gama, M. Shcherban, M. Rolf, and M. Hoffmann, “Goal-directed tactile exploration for body model learning through self-touch on a humanoid robot,” *IEEE Transactions on Cognitive and Developmental Systems*, 2021.
- [9] A. R. Geist and S. Trimpe, “Structured learning of rigid-body dynamics: A survey and unified view from a robotics perspective,” *GAMM-Mitteilungen*, vol. 44, no. 2, e202100009, 2021. DOI: <https://doi.org/10.1002/gamm.202100009>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/gamm.202100009>. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/gamm.202100009>.
- [10] J. W. Hart and B. Scassellati, “A robotic model of the ecological self,” in *Humanoid Robots (Humanoids), 2011 11th IEEE-RAS International Conference on*, IEEE, 2011, pp. 682–688.
- [11] M. Hersch, E. Sauser, and A. Billard, “Online learning of the body schema,” *International Journal of Humanoid Robotics*, vol. 5, no. 02, pp. 161–181, 2008.
- [12] M. Hoffmann, H. Marques, A. Arieta, H. Sumioka, M. Lungarella, and R. Pfeifer, “Body schema in robotics: A review,” *IEEE Transactions on Autonomous Mental Development*, vol. 2, no. 4, pp. 304–324, 2010.
- [13] J. M. Hollerbach and C. W. Wampler, “The calibration index and taxonomy for robot kinematic calibration methods,” *The International Journal of Robotics Research*, vol. 15, no. 6, pp. 573–591, 1996, ISSN: 0278-3649. DOI: 10.1177/027836499601500604.
- [14] L. Jacques, G. Baldassarre, V. G. Santucci, and K. O’Regan, “Sensorimotor contingencies as a key drive of development: From babies to robots,” *Frontiers in Neurorobotics*, vol. 13, p. 98, 2019.
- [15] P. Lanillos, E. Dean-Leon, and G. Cheng, “Yielding self-perception in robots through sensorimotor contingencies,” *IEEE Transactions on Cognitive and Developmental Systems*, vol. 9, no. 2, pp. 100–112, 2016.
- [16] Q. Leboutet, J. Roux, A. Janot, J. R. Guadarrama-Olvera, and G. Cheng, “Inertial parameter identification in robotics: A survey,” *Applied Sciences*, vol. 11, no. 9, p. 4303, ISSN: 2076-3417. DOI: 10.3390/app11094303.
- [17] A. C.-H. Lee, H.-K. Hsu, and H.-P. Huang, “Optimized system identification of humanoid robots with physical consistency constraints and floating-based exciting motions,” *International Journal of Humanoid Robotics*, vol. 19, no. 05, p. 2250015, 2022.
- [18] Q. Li, R. Haschke, and H. Ritter, “Towards body schema learning using training data acquired by continuous self-touch,” Seoul, Korea (South): IEEE, 2015, pp. 1109–1114, ISBN: 978-1-4799-6884-8. DOI: 10.1109/HUMANOIDS.2015.7363491.

- [19] H. Lipson, “Task-agnostic self-modeling machines,” *Sci. Robotics*, vol. 4, no. 26, 2019. DOI: 10.1126/SCIROBOTICS.AAU9354.
- [20] M. Lutter and J. Peters, “Combining physics and deep learning to learn continuous-time dynamics models,” *The International Journal of Robotics Research*, vol. 42, no. 3, pp. 83–107, 2023, ISSN: 0278-3649. DOI: 10.1177/02783649231169492.
- [21] K. Malinovska, I. Farkas, J. Harvanova, and M. Hoffmann, *A connectionist model of associating proprioceptive and tactile modalities in a humanoid robot*, 2022. DOI: 10.1109/icdl153763.2022.9962195.
- [22] R. Martinez-Cantin, M. Lopes, and L. Montesano, “Body schema acquisition through active learning,” in *2010 IEEE international conference on robotics and automation*, IEEE, 2010, pp. 1860–1866.
- [23] P. D. H. Nguyen, M. Hoffmann, U. Pattacini, and G. Metta, *Reaching development through visuo-proprioceptive-tactile integration on a humanoid robot - a deep learning approach*, 2019. DOI: 10.1109/devlrm.2019.8850681.
- [24] P. D. Nguyen, Y. K. Georgie, E. Kayhan, M. Eppe, V. V. Hafner, and S. Wermter, “Sensorimotor representation learning for an “active self” in robots: A model survey,” *KI-Künstliche Intelligenz*, pp. 1–27, 2021.
- [25] D. Nguyen-Tuong and J. Peters, “Model learning for robot control: A survey,” *Cognitive processing*, vol. 12, no. 4, pp. 319–340, 2011.
- [26] H. A. Pierson and M. S. Gashler, “Deep learning in robotics: A review of recent research,” *Adv. Robotics*, vol. 31, no. 16, pp. 821–835, 2017. DOI: 10.1080/01691864.2017.1365009.
- [27] G. Pugach, A. Pitti, O. Tolochko, and P. Gaussier, “Brain-inspired coding of robot body schema through visuo-motor integration of touched events,” *Frontiers Neurorobotics*, vol. 13, p. 5, 2019. DOI: 10.3389/FNBOT.2019.00005.
- [28] J. Sturm, C. Plagemann, and W. Burgard, “Body schema learning for robotic manipulators from visual self-perception,” 2009. DOI: 10.1016/j.jphysparis.2009.08.005. [Online]. Available: <https://www.semanticscholar.org/paper/c8982c68d7f7cace3375a875cb895a307d5b89dd>.
- [29] N. Sünderhauf, O. Brock, W. J. Scheirer, R. Hadsell, D. Fox, J. Leitner, B. Upcroft, P. Abbeel, W. Burgard, M. Milford, and P. Corke, “The limits and potentials of deep learning for robotics,” *Int. J. Robotics Res.*, vol. 37, no. 4-5, pp. 405–420, 2018. DOI: 10.1177/0278364918770733.
- [30] J. Swevers, W. Verdonck, and J. De Schutter, “Dynamic model identification for industrial robots,” *IEEE Control Systems*, vol. 27, no. 5, pp. 58–71, 2007.
- [31] R. Zenha, P. Vicente, L. Jamone, and A. Bernardino, *Incremental adaptation of a robot body schema based on touch events*, 2018. DOI: 10.1109/devlrm.2018.8761022.