Bionic Reflex Control for a Tendon-Optimized Underactuated Anthropomorphic Hand: A Reinforcement Learning Approach

A thesis submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

by

Hirakjyoti Basumatary

(176103030)

Under the Supervision of

Prof. Shyamanta M. Hazarika

to the

DEPARTMENT OF MECHANICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

March 2025

ABSTRACT

The development and use of anthropomorphic robotic hands that mimic human-like dexterity and functionality have had significant impacts on both industrial and healthcare applications. However, despite the considerable advancements in robotic hand technology, existing models fall short of replicating the intricate functionality of the human hand, particularly in the context of prosthetic applications. For prosthetic hands, it may be important to emulate a sense of reflex that mimics the human grasp reflex, commonly referred to as the bionic reflex, to handle objects efficiently and stably. Challenges associated with the bionic reflex include addressing slippage prevention of the grasped object. The intuitive response to slippage involves an increase in grasp force, necessitating controlled grasp force adjustments to avoid deformation or breakage of the object. This thesis focuses on the emulation of bionic reflex control in a tendon-optimized, underactuated anthropomorphic hand. This is achieved by implementing an intelligent slip and deformation prevention controller based on grasp force signals. The novel control pipeline is tailored for underactuated anthropomorphic robotic hands with specific relevance for prosthetic applications.

The actuation parameters of an underactuated anthropomorphic hand are optimized to generate grasps with better quality. An anthropomorphic hand with underactuated kinematics is chosen keeping in mind the cost and the weight factor. The goal is to optimize the underactuated parameters, i.e. pulley radius, spring stiffness, and spring preload angles of the tendon-driven robotic hand for five grasps generated through synergies. The five grasps of daily living activities chosen include power, precision, cylindrical, oblique, and pinch. This is achieved through optimization based on potential grasp robustness (PGR) and potential contact robustness (PCR) metrics.

A slip prevention control paradigm is developed to automatically detect slippage and prevent droppage of the grasped object for the above-designed underactuated anthropomorphic hand. The novelty lies in a. the controller's ability to detect slip and b. optimize the grasping forces automatically for rigid bodies. Slip detection is achieved using a wavelet transformation technique, and a reinforcement learning (RL)-based controller arrives at the optimal grasping forces. The RL agent has been trained by inducing domain randomization to take care of unmodeled dynamics, noisy parameters, and changing environmental conditions. This is expected to

improve the generalization of the learned policy for better disturbance rejection capability. The trained RL agent is able to grasp and lift different objects without any slippage, regardless of the object's unseen properties.

Further, the slippage prevention controller is augmented with deformation prevention to complete the bionic reflex controller. Knowledge of constitutive models or properties of the grasped object plays a crucial role in formulating the bionic reflex. Past endeavors in deformation prevention predominantly relied on model-free and supervised learning approaches to eliminate a priori knowledge of grasped object properties, necessitating human intervention during thresholding and labeling tasks. In this study, an innovative bionic reflex control pipeline is introduced, leveraging RL to eliminate the need for labeling thresholds, thus minimizing human intervention and automating the learning process. The proposed bionic reflex controller has been designed and tested on the underactuated anthropomorphic hand, manipulating deformable objects in the physics simulator with domain randomization incorporated. The findings of this experiment underscore the promise of RL as a potent approach for advancing bionic reflex control within anthropomorphic robotic hands.

Finally, a hierarchical approach of robustifying the pre-trained RL policy, using adaptive control is proposed. RL-trained policies are susceptible to failure in perturbed environments due to dynamic variations. To address this limitation, an approach of augmenting a pre-trained RL policy with an adaptive sliding mode controller, particularly suited for systems with continuous state and action spaces, is explored. Leveraging the invariance property of the sliding mode algorithm against uncertainties, the approach reported here enhances the robustness of the trained RL policies to tackle real-time, diverse, dynamic variations in the simulation environment settings. The results obtained validate the effectiveness of the proposed method in fortifying RL policies against matched disturbances and model uncertainties.

To conclude, this thesis addresses the bionic reflex problem within the framework of a tendon-optimized underactuated anthropomorphic hand. The facilitation of bionic reflex control in robotic hands proves imperative in the overarching goal of enhancing grasping efficiency. The key contribution of the thesis is the design of a control pipeline to implement bionic reflex, minimizing the reliance on human intervention using intelligent autonomous control, viz. reinforcement learning in a physics-based simulation. The results validate that the RL-based bionic reflex control is capable of grasping and lifting objects with minimum slippage and deformation together with maintaining a good grasp quality. The proposed methods in this thesis hold considerable promise for advancing the realm of prosthetic hands and robotics grasping research and thus, will pave the way towards the development of robust, autonomous bionic reflex controllers.

Dedicated to

My beloved friends, well-wishers and family

DEPARTMENT OF MECHANICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
GUWAHATI-781039

Declaration

I, Hirakjyoti Basumatary, hereby declare that the thesis entitled "**Bionic Reflex Control for a Tendon-Optimized Underactuated Anthropomorphic Hand: A Reinforcement Learning Approach**" submitted to the Department of Mechanical Engineering, Indian Institute of Technology, Guwahati in partial fulfillment of the requirements for the award of the degree of Doctor of Philosophy in Mechanical Engineering is based on the bonafide work carried out by me under the supervision of my supervisor, Prof. Shyamanta M. Hazarika. The results embodied in this thesis have not been submitted in part or in full, to any other university or institute for award of any degree or diploma.

Hirakjyoti Basumatary
(Registration No. 176103030)
Department of Mechanical Engineering
Indian Institute of Technology Guwahati
Guwahati, Assam-781039

DEPARTMENT OF MECHANICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
GUWAHATI-781039

CERTIFICATE

This is to certify that the work contained in the thesis entitled "**Bionic Reflex Control for a Tendon-Optimized Underactuated Anthropomorphic Hand: A Reinforcement Learning Approach**", by *Hirakjyoti Basumatary* (176103030), submitted to the Department of Mechanical Engineering, Indian Institute of Technology Guwahati, India, in partial fulfillment of the requirements for the award of the degree of Doctor of Philosophy has been carried out under my supervision and this work has not been submitted elsewhere for the award of any other degree or diploma.

Prof. Shyamanta Moni Hazarika
Professor
Department of Mechanical Engineering
Indian Institute of Technology Guwahati
Guwahati, Assam-781039

ACKNOWLEDGEMENT

I wish to express my sincere gratitude to a number of people who have supported me in a variety of ways throughout the course of this thesis.

First and foremost, I would like to sincerely thank Prof. Shyamanta Moni Hazarika for his invaluable guidance. I am deeply obliged to him for giving me the freedom to work in the field of robotic grasping and manipulation that I wanted. His critical comments and suggestions have been instrumental in completing this dissertation and I am extremely gratified with his endeavor in molding me into a disciplined research student. I would love to cherish my whole life in the memory of his association with me as a student.

Special thanks to all my doctoral committee members: Prof. Atanu Banerjee, Prof. Deepak Sharma, and Prof. Prithwijit Guha for their continuous guidance and suggestions during my research work. They are equally responsible for building confidence inside me through their constructive comments and advice.

I would like to extend my sincere thanks to the head of the department Prof. K. S. R. Krishna Murthy for his kind permission for lab facilities and fellowship support in the Department of Mechanical Engineering, IIT Guwahati.

In this precious moment of my life, I would like to express my deep sense of gratitude to all my friends, well-wishers, and my family, for their love, care, blessings, and constant encouragement throughout my life. Without their support, it would be impossible to complete my PhD journey. I am blessed to have them. I am deeply in debt to my family for the sacrifices they have made to ensure the fulfillment of my dreams. All of my friends, well-wishers and my family members have always been inspiring, supporting, and teaching me to understand the true value of human life

Lastly, I thank the Almighty for blessing me with this precious life.

Hirakjyoti Basumatary
IIT Guwahati
March, 2025

Contents

List of Figures

List of Tables

Acronyms

$AFNITSM$	Adaptive Fast Non-singular Integral Terminal Sliding Mode.
$ASMC$	Adaptive Sliding Mode Control.
$DDPG$	Deep Deterministic Policy Gradient.
DL	Deep Learning.
DNN	Deep Neural Network.
DR	Domain Randomization.
DRL	Deep Reinforcement Learning.
DWT	Discrete Wavelet Transform.
FEM	Finite Element Method.
MDP	Markov Decision Process.
$MFRL$	Model Free Reinforcement Learning.
RL	Reinforcement Learning.
$SOAITSMC$	Second-order Adaptive Integral Terminal Sliding Mode Controller.

Chapter 1

Introduction

The human hand is a potential tool to sense and manipulate objects with dexterity. It helps to accomplish complicated movements with precision [1]. The loss of an upper limb is a devastating injury that severely limits one's ability to perform activities of daily living (ADLs) [2]. Prosthetic hands allow amputees to interact with the outside world and aid in restoring their physical appearance and lost functionalities [3]. Presently, the primary objective is to develop a prosthetic hand with bidirectional information transmission that resembles a human limb [4]. The basic pipeline of a bionic prosthetic hand is shown in Fig. 1.1, where the amputee's biosignals are processed to detect grasp intents, and after execution of desired grasps, the sensor signals are sent back to the user as modality cues for haptic sensory feedback. Due to sensory feedback, one of the intricate features of the human hand is its ability to rapidly adjust the grasping force in response to grasped object information which includes object mass, stiffness, and friction. This is known as the grasp reflex. Bionic reflex, the emulation of the human grasp reflex in a prosthetic hand, is an active research area [5]. It involves slip prevention through control of forces exerted based on the grasped object properties to minimize deformation. Sensory feedback for the prosthetic hand can be achieved mainly in two ways: (i) by providing invasive or non-invasive feedback to stimulate the peripheral/central nervous system (user-in-the-loop) and (ii) by providing feedback to the controller to control grasp performed by the prostheses (user-outside-the-loop). In the latter case, the sensory loop stays within the prostheses and may not involve the user [6]. The basic block diagram of the user-in-the-loop and user-ouside-the-loop prostheses is shown in Fig.

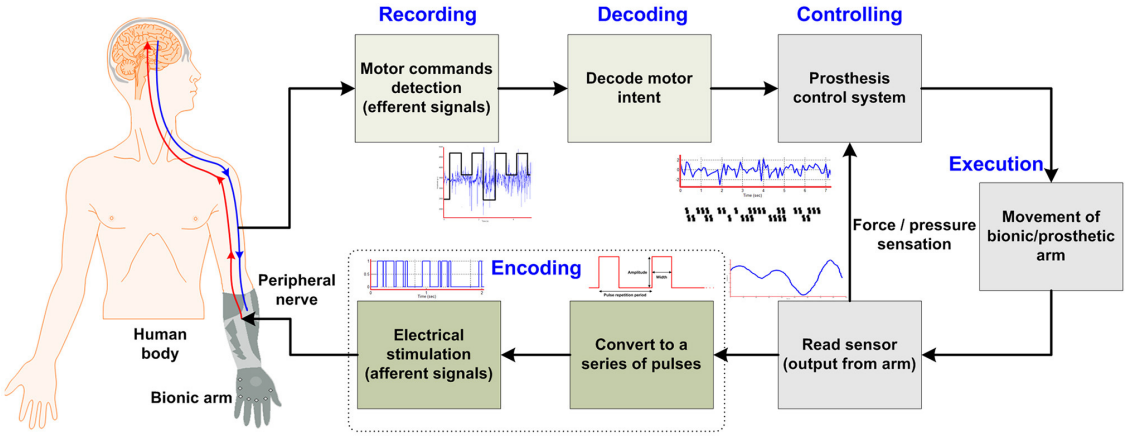


Figure 1.1: Block diagram of a typical user-in-the-loop bionic hand prosthesis [7]

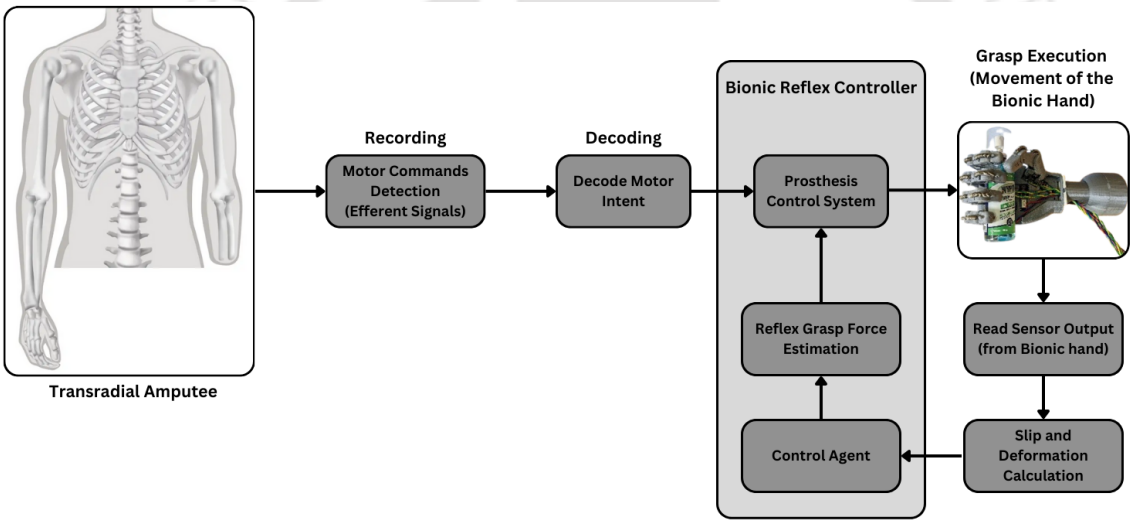


Figure 1.2: Block diagram of a typical user-outside-the-loop bionic hand prosthesis

1.1 and Fig. 1.2, respectively.

For a user-in-the-loop prosthetic hand, the finite amount of time (∼100ms) between the user's desire generated through the high-level sensory pathways to manipulate the prosthesis and their reaction time to regrasp during slippage of the object may cause a delay. To elaborate, for user-in-the-loop prosthesis, getting sensory input based on slippage, acquiring bio-signals, and decoding the biosignal for the updated grasp execution takes time. Hence, by drawing inspiration from the shorter response time of the autonomous spinal reflex arc in humans and adopting a low-level reflex control strategy may be a good choice to compensate for the delay time

caused by biological signal decoding techniques [8]. A low-level controller is part of the control framework in a hierarchical controller, which looks after the execution section, associated with error calculations and feedback control loops (in my case, the final grasp force control loop) [9]. This work implements such a controller in the user-outside-the-loop prosthesis to generate an autonomous slip and deformation control while grasping. In the user-outside-the-loop controller, the control command stays within the prosthesis with a higher degree of autonomy without involving the user. This way, it may be possible to achieve the bionic reflex reaction time to be as close as that of the human hand. Hence, implementing bionic reflex by understanding extensively the human grasp reflex and control may be a good way to reduce the cognitive burden of operating a prosthetic hand. The proposed work aims to work on the bionic reflex control system with the feedback loop within the prosthesis.

1.1 Motivation and Objectives

A majority of clinical prostheses operate on open-loop control, primarily based on visual feedback. The absence of reflex capabilities in most prosthetic hands, which are routinely exhibited by a human hand, often leads to a frustrating situation for users. Even if visual feedback may drive a reflex action, still the response time may not be enough to filter, extract and classify the newly extracted bio-signals in a user-in-the-loop prosthesis. Though the problem of slippage prevention has been looked at previously, not much work has been reported on preventing slip in an anthropomorphic underactuated hand, particularly without apriori knowledge of object properties. To tackle this issue of apriori knowledge of grasped object properties, data-driven and supervised machine learning techniques have been utilized. However, such model-free and supervised machine learning algorithms may not be intelligent enough to classify slippage for untrained models due to training bias inculcated during labeling and thresholding slip signals.

Also, I would like to mention that even if the problem of deformation prevention has been tackled by conventional Hooke's law or impedance controller in the literature, very few works take into account the control of grasped objects with variable stiffness. Feedback from tactile sensors of instantaneous force change can be helpful for object deformation prevention, relinquishing the reliance on a priori knowledge of constitutive models or object properties. Hence, an intelligent and robust bionic

reflex controller with slippage and deformation prevention together may improve the efficiency of anthropomorphic hands.

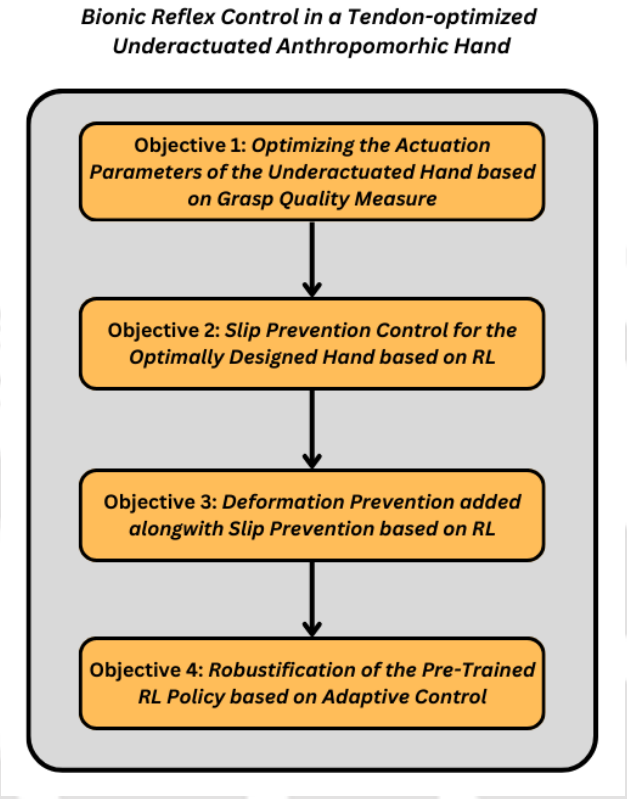


Figure 1.3: Thesis Objectives

Further, tendon-optimized underactuated prosthetic hand mimicking the biomechanical functionality of grasping and efficiently transmitting the grasp forces holds promise for bionic reflex control in hand prostheses due to its advantages like lower weight and lower control requirements. The literature on optimizing the underactuated tendon-driven mechanism of anthropomorphic hands based on grasping capabilities is limited.

Therefore, the above-existing issues and unexplored research directions of bionic reflex control in underactuated anthropomorphic prosthetic hands motivate us to develop improved techniques for it. The main aim of this dissertation is the development of a novel control pipeline to emulate bionic reflex control autonomously for an underactuated anthropomorphic hand. In light of the motivation above, the broad aim of the present work has been divided into the following objectives (Please

see Figure 1.3):

1. **Optimize the design of underactuated anthropomorphic hand based on grasp quality measure:** Enhancing grasp quality metric based on optimization of design variables of the actuation parameters of the tendon topology, i.e., pulley radius, spring stiffness, and pre-load angles.

2. **Autonomous bionic reflex controller based on reinforcement learning:**

   (a) **Slip Prevention-based Bionic Reflex Grasp Controller**: Intelligent slippage prevention controller based on wavelet transformation technique and reinforcement learning.

   (b) **Deformation Prevention-based Bionic Reflex Grasp Controller**: Autonomous deformation prevention controller on top of the slip prevention controller to augment the bionic reflex control to enable efficient grasping.

3. **Robustifying the pre-trained RL policy for bionic reflex control**: A hierarchical approach of robustifying the pre-trained RL policy by augmenting it with an adaptive controller for better disturbance rejection capability.

1.2 Thesis Organization

The results of the investigations are elucidated and summarised in the thesis in the form of chapters. Fig. 1.4 shows the area of work-wise organization of the thesis. The thesis is divided into seven chapters.

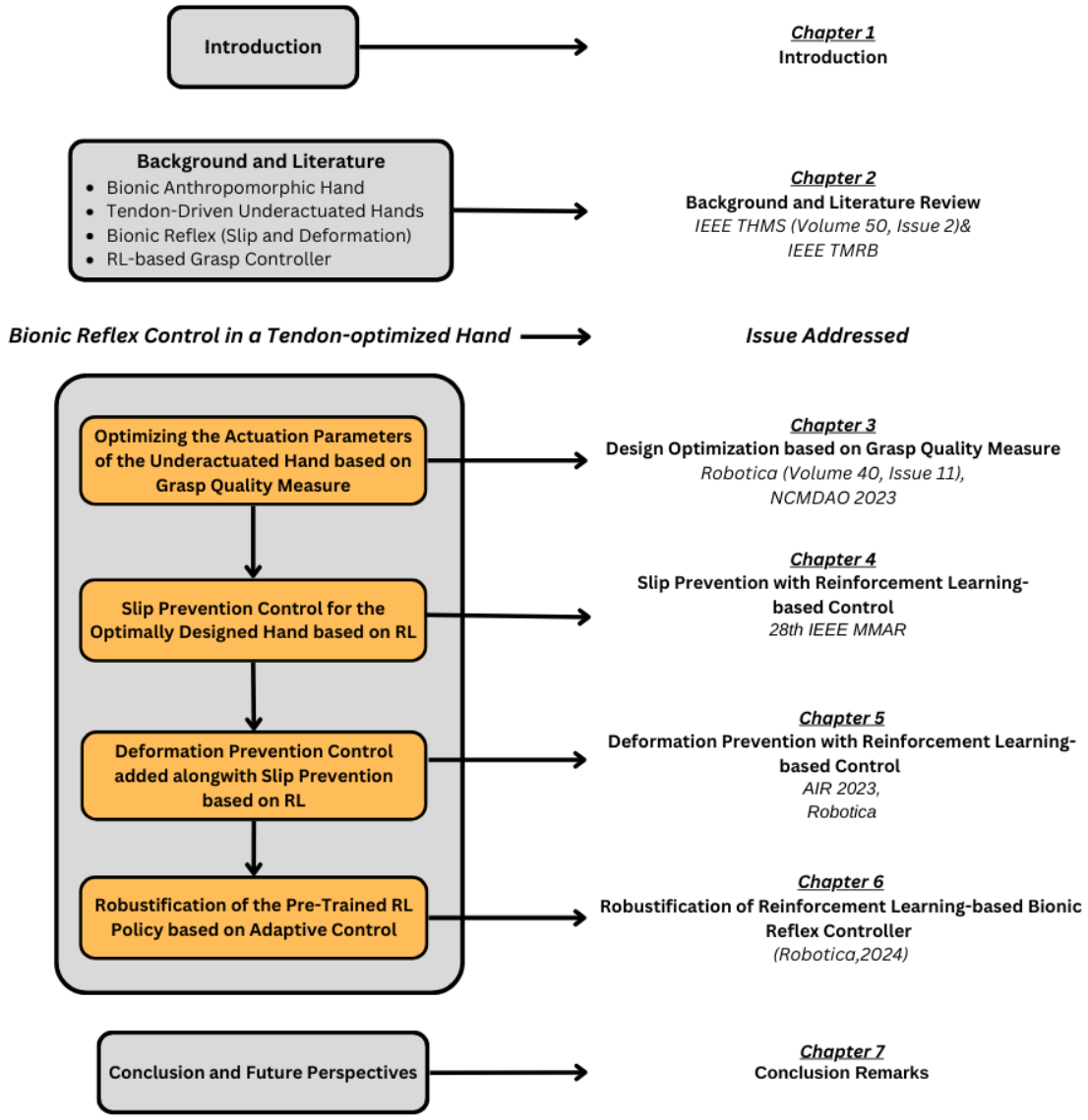


Figure 1.4: Area of my work-wise organization of the thesis

- **Chapter-1: Introduction**
  Chapter 1 presents an overview of the PhD area, motivation, research objectives, and overall outline of the thesis.

- **Chapter-2: Background and Literature Review**
  Chapter 2 discusses the problem of bionic reflex control in underactuated anthropomorphic hands. A brief review of the existing slippage and deformation

detection techniques for implementing bionic reflex is discussed. Further, a feasible design of the underactuated hand is important to implement such a biomimetic control strategy. Hence, literature on the state-of-the-art optimization of the hand design is reviewed. Finally, an autonomous grasp control technique based on reinforcement learning is discussed to be used as a tool for bionic reflex control.

- **Chapter-3: Design Optimization based on Grasp Quality Measure**
  Chapter 3 presents a design methodology for optimizing the anthropomorphic underactuated hand based on the grasp quality measure. For my task at hand, I generated force closure and robust grasps, which could keep the grasped object stable. Hence, the grasp robustness measure was utilized as a quality metric framed into an objective function to optimize the design of the underactuated parameters of the hand. This approach would eventually generate better grasps and a benchmark to implement bionic reflex control.

- **Chapter-4: Slip Prevention with Reinforcement Learning based Control**
  Chapter 4 introduces the slip prevention-based bionic reflex control based on the reinforcement learning approach. The bionic reflex problem is addressed in two steps/subtasks. One is slippage prevention, and the other is deformation prevention. In this chapter, the reinforcement learning-based bionic reflex controller is formulated and evaluated on rigid objects to determine its effectiveness in preventing slippage. Further, domain randomization was introduced during the training phase of the reinforcement learning agent. The addition of domain randomization was intended to improve the generalization capacity of the RL policy across a wide range of parameter variations. Further, the results are validated using grasp quality metrics.

- **Chapter-5: Deformation Prevention with Reinforcement Learning based Control**
  Chapter 5 further investigates bionic reflex control by introducing deformation prevention. This chapter implements the reinforcement learning-based bionic reflex controller on deformable objects to efficiently fulfill the second step/subtask of the bionic reflex problem, i.e., deformation prevention. The agent here is trained to prevent slippage of the grasped object and minimize its

deformation. Domain randomization is also introduced here by randomizing the physics parameters of the grasped object, and the results are validated using grasp quality metrics.

- **Chapter-6: Robustification of Reinforcement Learning-based Bionic Reflex Controller**

  Chapter 6 builds upon Chapters 4 and 5 and describes the use of adaptive sliding mode control to improve the robustness of reinforcement learning-based controllers. The control framework translates into a technique of augmenting and robustifying the reinforcement learning policy deployment following training in the nominal environment. This is a hierarchical deployment strategy for reinforcement learning policy with a traditional robust adaptive controller in the control pipeline. Here, the reinforcement learning policy is utilized as a planner to generate the desired action for traditional adaptive controllers, and the latter enables us to employ conventional stability and verification methods. Finally, the bionic reflex controller trained by this augmented adaptive controller is able to reject both model disturbances and matched uncertainties. This holds the promise of efficient policy generalization capability and Sim2Real transfer capability.

- **Chapter-7: Concluding Remarks**

  Chapter 7 draws the significant conclusions of the thesis and recommendations for future works.

Chapter 2

Background and Literature Review

2.1 Bionic Anthropomorphic Hand

In the field of robotics, traditional end effectors often fall short when high levels of dexterity and versatility are required, making humanoid robot hands an attractive alternative [10]. Humanoid robot hands/anthropomorphic hands in the form of prosthetic hands have played a significant role by augmenting amputees who have lost their arm due to disease or disability [11]. Despite the progress in prosthetic hand technologies, only 50-60% of amputees wear one. The rejection rate is also estimated to be as high as 40% [12]. The development of bionic hands has not yet percolated to amputees, mainly due to the high price associated with them [13]. On the other hand, upper limb loss leads to many patients being unable to return to their previous occupation or work, in general, [14]. In view of the above, restoration of hand functionalities could lead to socio-economic impact. Therefore, research in the area of bionic hand assumes importance.

A word cloud of the most popular topics in current research in bionic prosthetic hands is given in Fig. 2.1. From the word cloud analysis, it is seen that most of the prominent topics of the word cloud relate clearly to the subject area of hand prostheses, e.g., “prosthetic,” “hand,” “upper-limb,” “anthropomorphic,” “bionic,” “transradial,” “transhumeral,” “forearm,” etc. Words like “control,” “myoelectric,” “emg,” “signal,” etc., being significant, convey the message that much

research is focused on the myoelectric control system (MCS). A myoelectric-based control system can be classified into EMG-based pattern recognition (EMG-PR) and EMG-based non-pattern recognition (EMG-non-PR). Therefore the occurrence of the word "pattern-recognition" is evident. Further, terms relating to the research field of control include "feedback," "force," "real-time," "adaptive," "fuzzy," "position," and "robust," showing that the subject area of control in the bionic hand is very active. These words, along with other prominent words like "stiffness," and "slip," provide information that research has been published on bionic reflex to implement slippage avoidance and deformation prevention control. A rich group of concepts from the mechanical aspects of the design and development of prosthetic hand is reflected from words like "under-actuated," "mechanism," "novel," "actuation," "sensor," "joint," "finger," "kinematic," "flexible," "mechanical," "compliant," "optimization.' Research has been published on the joint coupling of tendon-driven mechanisms (TDM) and topology optimization, which have been visible from the significance of the word "tendon-driven" in the word cloud [15]. Hence, in this thesis, based on the word cloud analysis, I focus on a few popular topics and aim my objectives towards underactuated tendon-driven anthropomorphic hand control for slippage and deformation prevention.

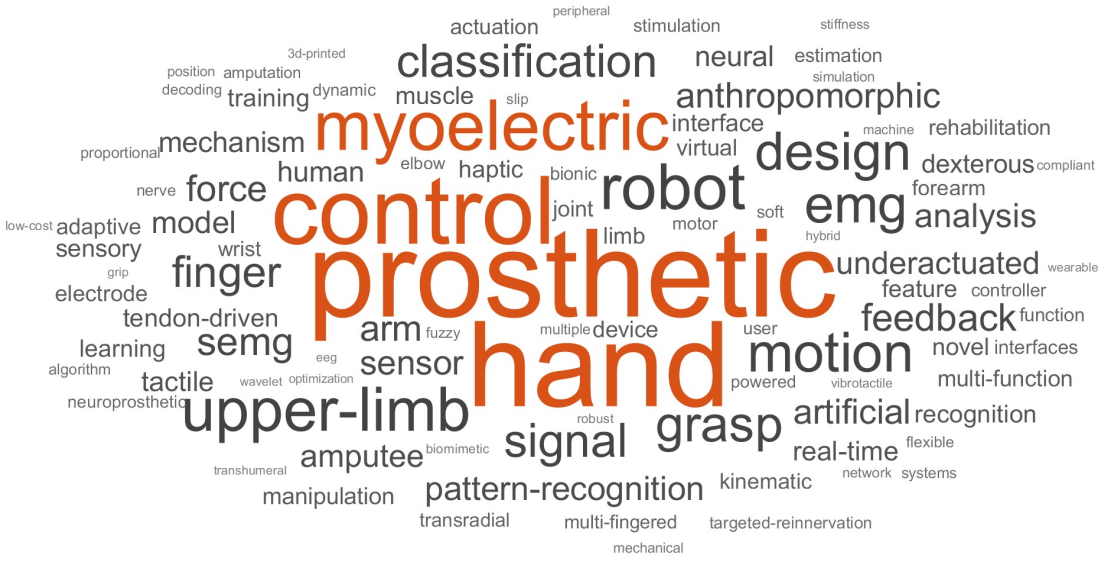


Figure 2.1: Word-cloud of the most popular topics in bionic hand research [15].

2.1.1 What is Bionic Reflex ?

As discussed in the previous chapter, for a user-outside-the-loop prosthetic hand, one of the important features is the ability to grasp without involving the user. Here, at the low level, a feedback-based controller helps generate optimal grasp force and manipulation, thus aiding bionic reflex. "Bionic reflex", which is feedback reflecting emulation of human grasp reflex in prosthetic hand, generally involves preventing slip and deformation of the grasped object, is essential in a prosthesis [16]. I believe an efficient low-level bionic reflex controller may aid in achieving a sense of embodiment for amputees with user-outside-the-loop prostheses.

2.1.2 Bionic Reflex Applications

The implementation of efficient bionic reflex has numerous applications. While grasping any object with a robotic hand, there is a chance of the object slipping due to external disturbances. Grasped object slippage is a quick and transient event caused by the interruption of touch between the end-effector and the grasped object during any grasping maneuver, leading to the contact forces sliding away from the friction cone. In order to appropriately regain stability, the grasp force must be adjusted [17]. With the aid of a slippage prevention controller implemented in the grasping pipeline, slippage of the grasped objects can be prevented. However, the challenges of slippage due to external disturbances and suboptimal grasp forces leading to object deformation are significant obstacles affecting the efficiency and safety of robotic grasping applications [18]. Addressing the risks of slippage and deformation is paramount to the evolution of robotic grasping. In the realm of prosthetic and robotic hands, this quest leads us to the frontier of bionic reflex, an emulation of the human grasp reflex control.

2.1.3 Current Challenges

Challenges of bionic reflex control applications involve:

- *Optimized design of the robotic hand:* An optimal design of the robotic hand is essential to generate efficient grasps. Further, many things need to be taken care of, like anthropomorphic design, lightweight, simplicity, etc. All these factors push the agenda of designing an underactuated hand as a starting

point to conduct research on the bionic reflex. However, the optimized design of the anthropomorphic underactuated hand should cater to being able to grasp objects efficiently even in the presence of disturbances to provide robust grasps.

- *Autonomous slip prevention framework*: There are currently different approaches to detecting slip in a robotic hand. With the advancement of modern-day sensor technologies, it has become easy to detect slippages while grasping objects. However, the discerning of slip features should take place automatically to help in generalization to unknown testing objects. Currently, the majority of slip detection techniques require thresholding the raw signals or labeling the training dataset, thus requiring human intervention in slip prevention design.

- *Intelligent deformation prevention system*: For the completeness of the bionic reflex framework, it is important to generate optimal grasp force while avoiding slip so that the grasped objects do not undergo deformation. To facilitate that, an intelligent deformation prevention control paradigm needs to be included in the bionic reflex pipeline. However, deformation prevention relies on knowing the object properties beforehand, or similar to slippage avoidance, requires thresholding or labeling training set, necessitating human intervention in the control design stage.

- *Robust adaptive grasping controller*: Maintaining a stable grasp on objects is a critical challenge in the realm of robotic hands and upper-limb prostheses. However, external disturbances, like matched disturbances to the control input, often disrupt grasp stability, leading to disturbances and the grasped object slipping away or breaking. The bionic reflex controller is required to be robust to disturbances and adaptively should be able to handle uncertainties for better generalization capability.

2.2 Tendon driven underactuated mechanism

2.2.1 Hardware Topologies

Humanoid robotic hands integrate a spectrum of actuation systems to emulate the complex and dexterous motions of the human hand, with underactuated mechanisms standing out as a key strategy to simplify design while enabling adaptive grasping. In these systems, fewer actuators than degrees of freedom are employed, allowing mechanical compliance and efficient distribution of forces across multiple joints. Tendon drives—using lightweight cables fabricated from metal, plastic, or nylon—are especially prominent in underactuated designs due to their ability to transmit motion over long distances and facilitate remote actuator placement, which reduces overall inertia. By contrast, gear drives offer high transmission accuracy and significant torque amplification; however, their complexity and stringent assembly requirements typically limit their use in underactuated configurations that prioritize simplicity. Additionally, link and fluid drive systems provide alternative pathways for motion transmission, although they often face challenges such as cumulative errors and increased weight. Overall, the emphasis on underactuation in tendon-driven systems reveals a promising route for achieving enhanced grasp adaptability and efficient control in humanoid robotic hands, balancing the trade-offs between performance and mechanical complexity [19]. There are different tendon driven mechanisms. [20] presents a comprehensive framework for tendon-driven mechanisms, demonstrating that a wide variety of such systems exist—from fully controllable to underactuated designs—that can be decomposed into subtendon modules for optimized performance. The authors systematically classify these mechanisms based on their tendon routing and bias force properties, highlighting their versatility in robotic and prosthetic applications.

2.2.2 Optimization of the underactuated parameters

In robotics, the large number of degrees of freedom in a multi-fingered hand poses a great challenge in terms of design (large number of actuators, complex transmissions) and control (coordination of the degrees of freedom in the performance of a task). These challenges have motivated the implementation of underactuation mechanisms in robotic hands. An underactuated hand mechanism in prosthetic robotic

hands refers to a design where the number of actuators (motors or mechanisms that provide controlled movement) is fewer than the degrees of freedom (DOFs) of the hand. In simpler terms, it means that the hand has fewer motors or control inputs than the number of joints or movable parts it possesses [21]. Underactuated hands rely on passive mechanical elements, such as springs, tendons, or linkages, to exploit natural dynamics and achieve a wide range of grasping and manipulation tasks with minimal control inputs. By leveraging passive dynamics and exploiting the inherent compliance of materials, underactuated hands can achieve complex movements with fewer actuators, reducing weight, complexity, and power consumption compared to fully actuated designs. This can be particularly advantageous in prosthetic applications, where lightweight, efficient designs are desirable to enhance user comfort and functionality.

A lot of research has been conducted on the optimization of underactuated grippers [22–26]. However, this form of optimization is notably less pervasive in the context of anthropomorphic hands. Existing optimization efforts predominantly focus on specific design or geometry considerations, with some directed towards kinematic properties [27] or grasp workspace optimization [28]. Yet, optimization based on grasping capabilities is indispensable for establishing a stable grasp predicated upon force closure principles and facilitating robust disturbance handling. This constitutes a research area aimed at optimizing grasp efficiency and achieving functional biomimetic solutions. Moreover, such optimization strategies must also accommodate the underactuated nature inherent in anthropomorphic hand mechanisms.

2.2.3 Grasp Quality Metrics

Grasping capabilities can be measured by a metric known in the literature as grasp quality measure [29]. The grasp quality measure is an index that quantifies the goodness of a grasp. The quality of a grasp depends on the following properties: (i) disturbance resistance, (ii) dexterity, (iii) equilibrium, and (iv) stability. Given a robotic hand and a grasped object, there are an infinite number of possible grasps that can satisfy one or all of the above properties. The quality measures for robotic hands for dexterous manipulation can be broadly classified into two categories: (i) geometrical grasp quality criteria based on contact point location and (ii) grasp quality criteria based on robot hand properties. Under the first category, some

of the quality measures include force closure criterion, largest minimum resisted wrench, volume of the ellipsoid in the wrench space, minimum singular value, grasp isotropy index, minimization of grasping forces, etc. Under the second category, the grasp quality measures that fall are criteria depending on the hand geometrical properties, criteria based on hand-object jacobian [30].

Limited literature exists concerning grasp quality metrics tailored specifically for underactuated hands. Pozzi et al. [31] proposed two metrics for evaluating grasp quality in underactuated hands: (i) a modified variant of the largest-minimum resisted wrench to address underactuation challenges, and (ii) two robustness indices suitable for assessing grasp quality. One of these indices, the contact robustness metric, termed potential contact robustness (PCR), evaluates the contact's ability to withstand violation of contact constraints. The other metric, grasp robustness, referred to as potential grasp robustness (PGR), assesses the grasp's capability to overcome object immobilization constraints [32]. Such quality metrics offer valuable insights that can inform the implementation and optimization of underactuated anthropomorphic hand designs.

2.2.4 Control strategies for underactuated hands

Robotic finger control strategies have evolved from simple open-loop schemes to sophisticated closed-loop systems that utilize real-time sensor feedback for precise motion and force regulation. Traditional approaches, such as PID control, offer straightforward tuning and are effective in maintaining desired setpoints, yet they often lack robustness in the presence of unpredictable disturbances and varying load conditions. To address these limitations, advanced methods like impedance control and hybrid force–position control have been developed; these techniques adjust the virtual mechanical impedance of the finger to dynamically match the interaction forces with the environment, thereby improving adaptability during grasping tasks. Model-based control strategies, which rely on accurate dynamic models to predict system behavior, have also been employed, though their effectiveness can be compromised by the inherent uncertainties and nonlinearities of underactuated, compliant mechanisms. Adaptive control methods take this a step further by continuously tuning control parameters based on sensor feedback, while fuzzy logic controllers provide an alternative by managing uncertainty through rule-based decision making. Recently, machine learning-based control—particularly reinforcement

learning—has garnered significant attention. This data-driven approach enables the learning of optimal control policies directly from interaction data, allowing the system to better cope with the complex, nonlinear dynamics of underactuated robotic fingers and to generalize across a range of grasping scenarios. Given the challenges of modeling compliance and nonlinearity in such systems, reinforcement learning offers a compelling solution by adapting control strategies to diverse environments and tasks, making it an attractive option for robotic grasping applications [33].

2.3 Bionic Reflex Control

As discussed earlier, the term "bionic reflex" typically denotes the replication of the human grasp reflex in the context of robotic or prosthetic hands, primarily entailing the mitigation of slippage and deformation phenomena encountered during object manipulation. Thus, the discourse on bionic reflex control usually refers to the control of grasping force based on the detected slippage and deformation information from the feedback of the embedded sensors [5].

2.3.1 Slip prevention in bionic hands

In the context of robotic or prosthetic hand grasping research, slip prevention refers to the ability of the hand to maintain a stable grasp on an object despite external forces or perturbations that may cause slipping. Slip prevention is crucial for tasks that require holding onto objects securely, such as lifting, manipulation, or transportation. The goal of slip prevention is to ensure that the grasped object remains firmly held within the hand's grasp, minimizing the risk of dropping or losing control of the object. This is particularly important in applications where precise control and dexterity are required, such as in prosthetic hands used for daily activities or robotic hands used in industrial settings.

Detecting slip is one of the crucial tasks to prevent slippage and attain bionic reflex control. Methodologies of slip detection are mainly based on vibration, friction models, contact relationships, vision, and data-driven approaches. Vibration-based methods use sensors like polyvinylidene-fluoride (PVDF) film or accelerometers to detect induced vibrations during slippage. Friction model-based approaches rely on maintaining a low tangential force-to-normal force ratio ($< \mu_s$, static coefficient of

friction). Contact relationship-based methods involve time-frequency analysis (e.g., STFT, DWT) to detect sudden changes. Vision-based techniques employ optical and tactile sensors for detecting displacement, force distribution, and slippage velocities. Data-driven methods leverage supervised learning methods and deep learning for slip detection [34].

Following slip detection, the integration of this information into the control algorithm of the bionic hand is essential [35]. Numerous efforts have been directed towards mitigating slippage using sophisticated control methodologies. These include classical methods such as PID controllers, alongside more advanced techniques like sliding mode control, fuzzy control, and model predictive control, as well as model-based and model-free approaches [36]. Given the propensity for unexpected disturbances to induce unwanted slippages, robust adaptive control methods have been adopted to introduce adaptability into control inputs. For instance, adaptive sliding mode controllers have been successfully deployed to grasp objects securely without slippage [37]. Nevertheless, these approaches often rely on predetermined thresholds for detecting slip events and adjusting grasp stiffness, necessitating human intervention in the labeling process. To address this limitation, the development of model-free controllers is imperative, as they can circumvent the need for thresholding slip signals, thereby reducing human intervention. The integration of such robust algorithms into real anthropomorphic hands for implementation in bionic hands holds promise for enhancing grasping efficiency and reliability.

2.3.2 Deformation prevention in bionic hands

Deformation prevention in robotic or prosthetic hand grasping research refers to the capability of the hand to minimize or mitigate any changes in the shape or structure of the grasped object during manipulation. The prevention of deformation is crucial for maintaining the integrity of delicate or fragile objects and ensuring accurate manipulation tasks. Research in this area encompasses various strategies, including the development of control algorithms and mechanical designs aimed at applying appropriate forces and constraints to the object to prevent deformation. Additionally, advances in sensing technologies play a vital role in enabling the hand to detect subtle changes in object shape and adjust its grasping force and posture accordingly.

Deformation prevention and its control necessitate stiffness detection or deforma-

tion measurement. Existing methodologies for stiffness detection and control in the literature include: (1) intrinsic vibration frequency-based signal processing, (2) time-domain analysis methods, (3) the integration of measuring devices [38] (4) Hooke's Law [39]. Detecting the deformation of grasped objects in robotic or prosthetic hands typically involves integrating various sensing modalities to monitor changes in the object's shape, structure, or mechanical properties. Some common approaches to detecting deformation include force and tactile sensing, vibration-based methods, strain gauge sensors, acoustic emission monitoring, and model-based methods.

The subsequent step entails integrating the signals from deformation detection into the deformation control algorithm. Adaptive control methodologies have been harnessed to preclude deformation in objects exhibiting variable stiffness profiles during grasping and manipulation. Additionally, impedance-based control strategies have found utility in grasping deformable objects within the domain of robotic grasping research. Nevertheless, these methods necessitate comprehensive knowledge of model parameters and system dynamics, which can pose challenges in real-world applications. Alternative approaches to deformation control involve leveraging the kinematics of underlying underactuated mechanisms, compliant mechanisms, and soft grippers. However, the design and implementation of such intricate mechanisms to adapt grasping forces during the manipulation of deformable objects present formidable challenges in terms of complexity.

Addressing these intricacies pertaining to deformation prevention is crucial for enhancing the adaptability and efficacy of anthropomorphic robotic hands across a spectrum of tasks and environments. From delicate surgical instruments to objects with intricate geometries in industrial settings, the ability to prevent deformation is fundamental for achieving precise and reliable manipulation outcomes. This thesis contributes to the field by delving into novel methodologies and approaches aimed at advancing deformation prevention capabilities in anthropomorphic robotic hands. By exploring intelligent techniques and strategies, this research endeavors to enable more accurate and optimal manipulation of objects in diverse applications, thus increasing the generalization capability of the control agent in the field of robotic grasping and manipulation.

2.4 Reinforcement Learning based Grasp Controller

Advancements in machine learning, simulation techniques, algorithm design, and computer hardware have significantly enhanced the capabilities of reinforcement learning (RL). This progress has transformed RL into a robust tool capable of addressing a wide range of complex problems that were previously difficult or unsolvable by other methods. A particularly promising area for RL application is robotic control, where researchers have achieved notable success in tackling various challenging tasks, including navigating rough terrains and performing intricate object manipulation [40]. Reinforcement learning-based grasp control represents a pivotal frontier in the realm of robotic manipulation, offering a paradigm shift from traditional rule-based approaches to adaptive, data-driven methodologies. In this framework, robotic hands autonomously learn to grasp objects by interacting with their environment, guided by reinforcement signals that indicate the success or failure of grasp attempts. By leveraging reinforcement learning algorithms, robotic hands can acquire complex grasping strategies that are robust, versatile, and capable of generalizing across different objects and environments. This approach holds promise for overcoming challenges associated with variability in object shape, size, and material properties, as well as uncertainties in sensor measurements and environmental conditions. Moreover, reinforcement learning-based grasp control enables continual improvement and adaptation over time as the robotic system accumulates experience and refines its grasp strategies through trial and error. By delving into this frontier, this thesis aims to contribute to the advancement of reinforcement learning techniques for grasp control, with the overarching goal of enhancing the efficiency, reliability, and adaptability of robotic manipulation in real-world applications.

2.4.1 Domain Randomization

The core challenge for machine learning-based grasp quality heuristics is the difficulty and expense of assembling labeled real-world data. Hence, it is better to utilize reinforcement learning (RL) compared to supervised or unsupervised learning, given the limited amount of data, due to various advantages like trial and error learning, adaptive control, exploration, generalization, and continuous learning. Further, to increase the availability of training data for different experimental scenarios a technique known as domain randomization can be utilized. Domain ran-

domization is a technique for learning models that work in a test domain (which may be real-world, Sim-to-Real, or simulated, Sim-to-Sim) after only being trained on low-fidelity simulated data by randomizing the physics parameter of the simulator [41]. Domain randomization in grasping and manipulation research has been utilized in data generation for new object synthesis in vision-based grasping [41], data generation for reach-and-grasp grasping trajectories [42], pose estimation for vision-based grasping [43], unsupervised feature learning for visual inputs of grasped object properties [44], for sim-to-real transfer for efficient deployment and bridging the reality gap [45]. Hence, in this thesis work, utilizing the advantages of RL and domain randomization, I plan to design an efficient bionic reflex controller, which would have better generalization capability in handling slippage and deformation in different grasping scenarios.

2.4.2 Robustification of Learned Agent

Robustification is a type of optimization in which the input variables and parameters of a system are made less susceptible to the effects of random variability or noise. Robustifying the RL policy helps it to generalize against a wide range of variabilities and disturbances. However, a reinforcement learning agent trained well in the simulator may not perform well in the real-life scenario due to a mismatch of simulated and real control-law performances caused by the inaccurate representation of the real environment in simulation. To help improve the generalization capability of the RL agent to perform well in the testing environment (which may be unknown environment conditions in the real world, Sim-to-Real or simulated world, Sim-to-Sim), it may be trained under varying parameter variances and situations. Such a generalization can be addressed using techniques like mainly domain randomization (DR), adversarial reinforcement learning, learning to adapt (meta-learning), transfer learning, post-training augmentation, and knowledge distillation [46–48].

Domain randomization is a technique used in reinforcement learning (RL) and robotics to improve the generalization capabilities of learned policies across different environments or domains. Domain randomization addresses this issue by augmenting the training data with variations in the environment's parameters, such as lighting conditions, object textures, dynamics, or physics properties. By exposing the agent to a diverse range of simulated environments during training, domain randomization aims to encourage the learning of policies that are robust and adapt-

able to changes in the environment. Adversarial reinforcement learning (RL) is a framework that involves training an RL agent in a competitive environment against one or more adversaries. In traditional RL, the agent learns to maximize its reward by interacting with an environment that is assumed to be stationary and benign. However, in adversarial RL, the environment may actively oppose the agent's goals or attempt to deceive it. It can help in creating agents that are resilient to unexpected changes and adversarial attacks, leading to more reliable and trustworthy AI systems. Meta-learning, also known as learning to learn, is a subfield of machine learning that focuses on developing algorithms and methodologies capable of learning from and adapting to multiple learning tasks or environments. The core idea behind meta-learning is to enable models to acquire knowledge or learn new tasks quickly and efficiently, often with limited data or training examples. Transfer learning is a machine learning technique where a model trained on one task is reused or adapted for a different but related task. In traditional machine learning, models are typically trained from scratch on specific datasets for each new task. However, transfer learning leverages knowledge gained from one task to improve learning and performance on another task, especially when the two tasks share similarities. Knowledge distillation is a technique used in machine learning to transfer knowledge from a large, complex model (teacher model) to a smaller, more lightweight model (student model). The goal of knowledge distillation is to compress the knowledge stored in the teacher model into a more compact form that can be efficiently utilized by the student model, improving generalization. In the post-training augmentation-based methodology of robustifying pre-trained RL policy, an augmented robust controller is added in a hierarchical approach to the RL policy, to compensate for any disturbances. This approach helps to adaptively tackle disturbances during the testing phase.

2.5 Chapter Summary

This chapter includes an overview of the bionic reflex control research. It elaborates on the different applications and challenges of bionic reflex control. The chapter discusses the major methods and state-of-the-art in tendon-driven underactuated mechanism, bionic reflex, and reinforcement learning-based grasp control. Research areas on robotic grasping are examined in the context of slippage and deformation

prevention. I believe the bionic reflex control problem for robotic hands is an elusive problem, and the ability to create an autonomous and robust controller would significantly improve grasping.

Chapter 3

Design Optimization based on Grasp Quality Measure

3.1 Introduction

Anthropomorphic hands may help improve the embodiment of upper-limb amputees. However, the cost (and possibly the weight) of a prosthetic device increases as the complexity and functionality increase [49]. One approach that could help bring down not only the weight but also the cost is through underactuation. Underactuation usually means a mechanism with lower actuators than the total degrees of freedom (DoF) [50]. Consequently, the system cannot be commanded to follow arbitrary trajectories in configuration space. Nevertheless, the benefits of underactuation, such as lower weight and lower control requirements, outweigh the above restriction [51]. Designing an efficient, underactuated anthropomorphic prosthetic hand is a challenging task. The challenge comes from the perspective of designing an underactuated mechanism maintaining the functional biomimesis. In my case, I intend to optimize the design of the underactuated parameters keeping in perspective its ability to grasp efficiently. To give a notion of efficient grasping, I would require a metric to measure the efficiency of a grasp. Hence, I use grasp quality measure as a parameter to depict efficient grasping. Grasp quality measure refers to the goodness of the grasp that satisfies certain properties like disturbance resistance, dexterity, equilibrium, and stability [29].

Underactuation in robotic hands has a close link with postural synergy. Syner-

gies of the hand usually refer to those variables that can demonstrate and recreate the intricate human hand movements and configurations, even if reduced in dimension compared to the degrees of freedom of the hand. Synergy-based control of an underactuated hand aids in accomplishing the different grasp types by helping control the synergies directly, rather than controlling individual degrees of freedom (DOFs). Implementing the synergies on an anthropomorphic robot hand can control the joints with fewer inputs [52]. The number of synergies usually refers to the number of actuated joints in an underactuated hand [53]. Utilizing a synergy coupling matrix can help achieve underactuation and produce grasps with high quality, which aids in maintaining the functional biomimesis.

In this chapter, I optimize the actuation parameters of the underactuated hand based on grasp quality measures. I consider a soft synergistic compliant grasping model and provide an optimization framework that aids in getting the optimal actuation parameters of an underactuated anthropomorphic hand based on the grasp robustness measure. Postural synergies could be exploited by multi-fingered underactuated anthropomorphic hands to perform a multitude of grasps. Here the focus is on the five important grasp types of daily living activities, i.e., power grasp (G_{pow}), cylindrical grasp (G_{cyl}), precision grasp (G_{pre}), pinch grasp (G_{pin}), and oblique grasp (G_{obl}),(where G represents Grasp Matrix), as shown in Fig. 3.1. These grasps are considered as they are involved in more than 70% of the daily living activities [54]. I validate the proposed hand design through grasping experiments in a simulated environment and evaluate the grasp wrench space. The rest of the chapter is as follows: Background and related work is discussed in Section. 3.2. Section. 3.3 presents an overview of the problem formulation. The overall design methodology is reported in Section. 3.4. The results of the numerical simulation and its validation are discussed in Section. 3.5. Finally, I provide a summary of the work in Section. 3.6.

3.2 Background and Related Work

While researchers have optimized the design of underactuated non-anthropomorphic grippers on many instances [22–26], such optimization is far less prevalent among anthropomorphic hands. Recently, there has been research on the optimization of anthropomorphic hand design for fully actuated hands based on a novel perfor-

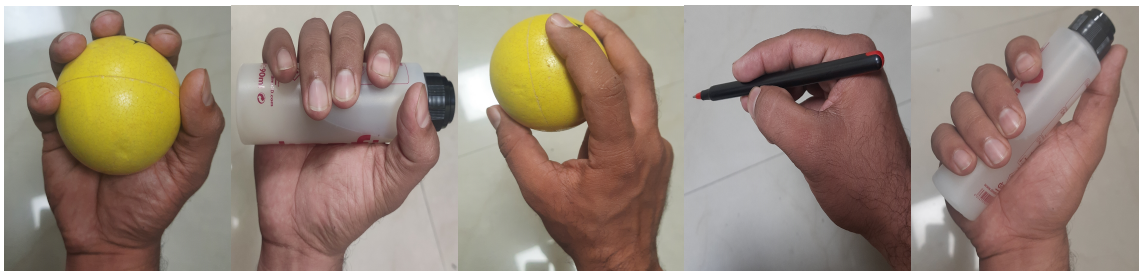


Figure 3.1: Human hand grasp types: (a) Power grasp (b) Cylindrical grasp (c) Precision grasp (4) Pinch grasp (5) Oblique grasp

mance index for its kinematic properties referred to as "interactivity of fingers" [27]. There is also literature on kinematic optimization of underactuated hands based on maximizing the workspace intersection and size of the largest graspable object [28]. However, it is concerned with kinematics rather than design based on grasping capabilities. A few works can be seen across literature where grasping capabilities were improved based on the optimization of the routing mechanism and location of pulleys of the tendon-driven mechanism. Gosselin et al. [55] optimized the tendon routing using a constraint gradient-based technique of an under-actuated anthropomorphic robotic hand with differential mechanism. However, the grasp quality was not checked during its evaluation. Inouye and Valero-Cuevas [56] in their paper optimized the position of joint centers, tendon routings, and maximal tendon tension. The grasping capabilities of the human hand for index finger and thumb precision were calculated and compared to the capabilities of thousands of different tendon-driven anthropomorphic designs. However, the hand was a tendon-driven hand and not an underactuated one. Moreover, the work and the MIV grasp quality were analyzed in only the two-fingered grasp of an anthropomorphic hand. Wen et al. [22] optimized the size and location of the pulley for maximum transmission of tendon forces in a three-fingered robotic hand.

Optimization based on grasping capabilities is necessary to generate (a) a stable grasp based on force closure properties and (b) to enable disturbance handling. This is an area of active research, particularly to arrive at functional biomimesis. Recently, Chen et al. [57] performed design optimization based on the unbalanced torque method of the mechanically realizable manifold in posture and torque space by checking implementable grasp synergies. However, there remains the requirement of optimizing the design of the underactuated hand based on some grasp quality

measures, as I am not aware of any reported in the literature. As discussed in the literature, there are many grasp quality measures [29]. However, few talk about grasp quality measures for underactuated hands. Pozzi et al. [31] gave two measures of grasp quality for an underactuated hand: (i) a modified version of the largest-minimum resisted wrench to handle the underactuation; (ii) two robustness indices that can be used to measure grasp quality. One of the robustness indices is the contact robustness measure, known as potential contact robustness (PCR), and the other is the grasp robustness measure, called potential grasp robustness (PGR). Contact robustness is the ability of the contact to resist itself from violating contact constraints, whereas grasp robustness is the ability of the grasp to overcome the object immobilization constraint [32]. In this work I will be utilizing PGR as the objective function to be optimized for designing the underactuation parameters.

3.3 Problem Formulation

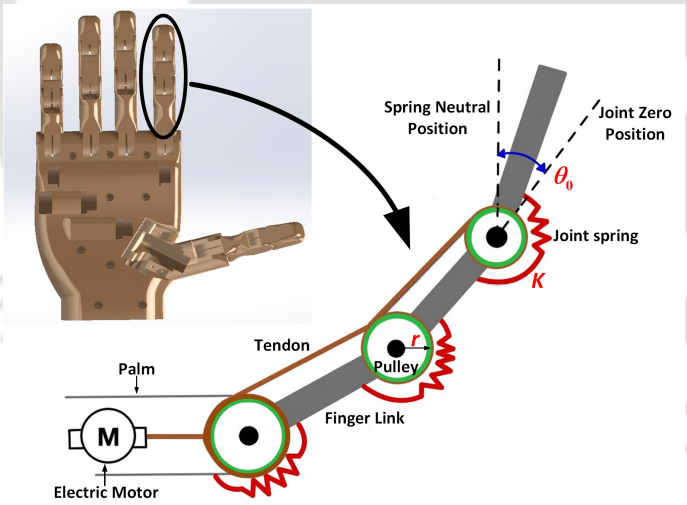


Figure 3.2: Schematic of an underactuated finger. Design parameters to be optimized are shown in red, i.e., r, K and θ_0

The goal is to optimize the underactuated parameters for the five grasps generated through synergies. I aim to achieve it through the optimization of grasp quality measures of the grasps. Hand kinematics of under-actuated hands usually include the specification of number, shape, size of fingers, tendon routing mecha-

nism, size of the tendon moment arms, etc. The design space (all the possible design solutions) of the anthropomorphic hand was narrowed down by considering a hand with pre-defined kinematics and structure. I consider the kinematics, where the tendons are routed around the pulleys. The flexion is achieved under the action of the tendon-driven mechanism, and the springs help in the extension of the fingers. The schematic of the underactuated finger and the design parameters to be optimized are shown in Fig. 3.2.

The author has selected the above underactuated tendon-driven finger mechanism because it effectively leverages both active and passive tendon elements to achieve adaptive grasping with reduced actuator count. Unlike some designs in [20]—which include fully actuated, semi-actuated, and passive configurations— the underactuated mechanism, which I adopted, integrates passive bias forces through strategic tendon routing and elasticity. This design enables a single actuator to modulate multiple joints by relying on the inherent compliance of the tendons, thus decoupling joint motion and allowing the system to adapt its grasp to irregular objects, as well as coupling joint torques for efficient actuation transmission. In typical underactuated tendon-driven mechanisms, a single actuator (or a small set of actuators) generates a tendon force that is distributed to multiple joints through a routing matrix. This design inherently couples the joint torques—meaning that the same actuator force simultaneously contributes to the torques at several joints. However, due to the compliant elements (like springs) incorporated at the joints, the resulting joint motions can adapt somewhat independently to external forces or object geometry, effectively "decoupling" the motion even though the torques remain coupled. Additionally, the approach offers advantages in weight reduction, energy efficiency, and reliable actuation.

The actuation parameters of the design space I choose to optimize include (i) the pulley radius (i.e., the tendon moment arms of the joints): which helps in the efficient transmission of forces; (ii) the spring stiffness (passive actuators at the joints): which helps in extension and maintaining joint-stiffness; and, (iii) the spring pre-load angles: which helps maintain a feasible posture for grasp preshape before attempting different grasp types. The optimized parameters are expected to produce the most robust grasps for all the grasp postures.

Table 3.1: Nomenclature

Notation	Definition
$\boldsymbol{w} \in R^6$	external wrench
$\boldsymbol{\tau} \in R^{n_q}$	joint torque
n_q	number of hand joints
n_c	number of contact points
$n_\lambda = 3n_c$	single point contact with friction
$\boldsymbol{\lambda} \in R^{n_\lambda}$	generic contact forces
$G \in R^{6\times n_\lambda}$	grasp matrix
$J \in R^{n_\lambda \times n_q}$	hand jacobian
$K_c \in R^{n_\lambda \times n_\lambda}$	contact stiffness matrix
$\boldsymbol{c^h}$	contact location on the hand
$\boldsymbol{c^o}$	contact location on the object
$\delta\boldsymbol{x}$	change in $\boldsymbol{x}$; where $\boldsymbol{x}$ is a variable
$\boldsymbol{q} \in R^{n_q}$	joint displacement
$\boldsymbol{u}$	object position
$K_q \in R^{n_q \times n_q}$	joint stiffness matrix
$S \in R^{n_q \times n_z}$	synergy matrix
$S_a \in R^{n_q \times n_{z_a}}$	synergy matrix of active tendons
$S_p \in R^{n_q \times n_{z_p}}$	synergy matrix of passive tendons
$\boldsymbol{\sigma}$	Langragian co-ordinate contact forces
K_{za}	active joint stiffness
$\boldsymbol{z_a}$	actual actuator position
$\boldsymbol{z_p}$	passive input position
$\boldsymbol{z_{ra}}$	reference actuator position
K_{zp}	passive joint stiffness
K_z	synergy stiffness
$\boldsymbol{z_r}$	synergy reference value
μ_i	friction co-efficient of i^{th} contact
$\boldsymbol{\lambda_{i,n}}$	normal contact force i^{th} contact
$\boldsymbol{\lambda_{i,t}}$ and $\boldsymbol{\lambda_{i,o}}$	tangential contact force i^{th} contact

3.3.1 Grasp Quality Measure-Based Optimization

3.3.1.1 The Complaint Grasp Model

The mathematical model of grasping must represent the behavior of the hand and the grasped object while applying varied loading conditions. The most desirable quality while grasping is the capacity to maintain grasp stability in the face of unknown disturbance forces and moments applied to the object. The rigid-body grasping model is the basic mathematical model of grasping. According to this model, in static equilibrium conditions, utilizing the principle of virtual work principle, I can write the equation of object and hand equilibrium, respectively, as [58]:

$$\boldsymbol{w} = -G\boldsymbol{\lambda} \tag{3.1}$$

$$\boldsymbol{\tau} = J^T\boldsymbol{\lambda} \tag{3.2}$$

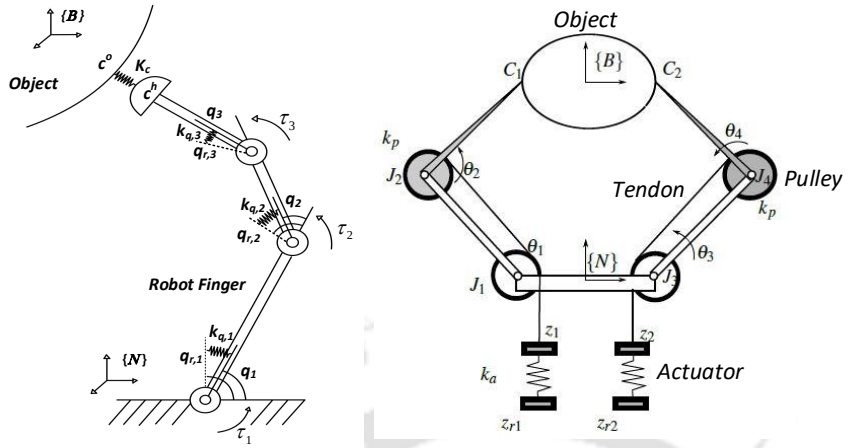
where, $\boldsymbol{w}$ and τ are described in Table. 3.1. For ease of readability, symbols used in this chapter are listed in Table. 3.1.

Extending the rigid-body grasping model to include compliance and design hands that maintain a secure grasp with fewer joints, thus providing greater mechanical robustness and minimizing the flexibility of grasp planning [59]. Analyzing a grasp through the complaint grasping model increases the dexterous manipulability and robustness of static grasps as it helps implement desired compliance behaviors while grasping an object. In modeling the compliant grasp, a set of hypothetical springs is introduced at the contact locations between the object and the fingers. Fig. 3.3(a) shows the complaint joints, complaint contacts, and the different vectors like joint variables $\boldsymbol{q}$, $\boldsymbol{q_r}$, $\boldsymbol{\tau}$, etc. In modeling the compliant grasp, a set of hypothetical springs is introduced at the contact locations between the object and the fingers. The contact force variation $\delta\boldsymbol{\lambda}$ from the initial equilibrium position is given as

$$\delta\boldsymbol{\lambda} = K_c(\delta\boldsymbol{c^h} - \delta\boldsymbol{c^o}) = K_c(J\delta\boldsymbol{q} - G^T\delta\boldsymbol{u}) \tag{3.3}$$

The actual joint variables, $\boldsymbol{q}$, may differ from the reference one, $\boldsymbol{q_r}$, if the hand structural parameters are not perfectly stiff. This can be modeled as:

$$\delta\boldsymbol{\tau} = K_q(\delta\boldsymbol{q_r} - \delta\boldsymbol{q}) \tag{3.4}$$



(a) (b)

Figure 3.3: Compliant Grasping (a) Compliant joints and compliant contacts with visualization of points, vectors and geometry, adapted from [59] (b) Complaint underactuated hand with passive elastic elements in the joints J_2 and J_4, actuated by two elastic tendons, adapted from [58].

Further, to include the concept of underactuation, the concept of postural synergies from neuroscience can be utilized. According to postural synergies, the most commonly seen postures in human hand movement can be defined by a reduced dimension of principal directions in hand configuration space. The postural synergy model, on the other hand, is good for creating pre-shaping grasp phases but not for predicting how grasping forces are formed after contact. As a result, a soft synergistic model with compliance is used to account for forces emerging from resistance to interpenetration of bodies. This model allows a feasible grasping model for an underactuated robotic hand. Fig. 3.3(b) shows an example of a compliant underactuated hand grasping an object. In this figure, the 4-DOFs hand is actuated by two actuators through a tendon-pulley transmission mechanism and has passive elastic elements (springs) at two joints J_2 and J_4.

Considering the linear approximation of all the possible displacements around an equilibrium configuration while grasping an object and the synergistic underactuation, the fundamental grasp equation (FGE) can be summarized, which takes into account all the above conditions [60]. For further details of the derivation of FGE, refer to Appendix A.

$$A\boldsymbol{\delta x} = \boldsymbol{\delta y} \tag{3.5}$$

$$A = \begin{bmatrix} -G & 0 & 0 & 0 & 0 & 0 & 0 \\ J^T & K_{J,u} & -I & K_{J,q} & 0 & 0 & 0 \\ 0 & 0 & S^T & 0 & -I & K_{S,z} & 0 \\ C_c & G^T & 0 & -J & 0 & 0 & 0 \\ 0 & 0 & -I & -K_q & 0 & 0 & K_q \\ 0 & 0 & 0 & 0 & C_z & I & 0 \\ 0 & 0 & 0 & 0 & 0 & S & -I \end{bmatrix} \tag{3.6}$$

$$\boldsymbol{\delta x} = [\delta\boldsymbol{\lambda} \;\; \delta\boldsymbol{u} \;\; \delta\boldsymbol{\tau} \;\; \delta\boldsymbol{q} \;\; \delta\boldsymbol{\sigma} \;\; \delta\boldsymbol{z} \;\; \delta\boldsymbol{q_r}]^T \tag{3.7}$$

$$\boldsymbol{\delta y} = [\delta\boldsymbol{w} \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; \delta\boldsymbol{z_r} \;\; 0]^T \tag{3.8}$$

The linear system given by Eq. 3.5 can be solved to find $\delta\boldsymbol{x}$, if the square matrix A is not singular. Usually, matrix A is typically invertible. Conditions of its non-invertibility arise only when: condition (1) indeterminacy of the grasp, i.e., nullspace of the transpose of grasp matrix ($\mathcal{N}(G^T)$) is non-trivial, condition (2): defectiveness of the grasp, i.e., nullspace of the transpose of hand jacobian ($\mathcal{N}(J^T)$) is non-trivial. The above conditions arise when the total contacts are too limited, and the object displacement is not fully constrained by the contacts arising between the hand and the environment. Given $\delta\boldsymbol{z_r}$ and/or for a given $\delta\boldsymbol{w}$, a unique solution exists for Eq. 3.5, if the matrix, A is not singular. If I assume $\delta\boldsymbol{w} = 0, \delta\boldsymbol{z_r} \neq 0$ and solve, then it means for a given grasping system, what are the permissible controllable forces and motions acting on the hand and arm reference configuration. Further, the solution for the assumption $\delta\boldsymbol{w} \neq 0, \delta\boldsymbol{z_r} = 0$, demonstrates how much amount of external disturbances the robotic hand can withstand. External disturbances are represented as changes in the wrenches applied to the grasped object [61].

3.3.1.2 Grasp Quality Measures: Grasp Robustness and Contact Robustness

Considering i-th contact, the contact forces $\boldsymbol{\lambda_i}$ are required to satisfy unilateral force constraint and Coulomb friction constraint. Eq. 3.9 and Eq. 3.10 specify the above conditions respectively. These constraints help in avoiding detachment and slip of contact during object grasp.

$$\boldsymbol{\lambda_{i,n}} \geq 0 \tag{3.9}$$

$$\sqrt{\boldsymbol{\lambda_{i,t}}^2 + \boldsymbol{\lambda_{i,o}}^2} \leq \mu_i \boldsymbol{\lambda_{i,n}} \tag{3.10}$$

Given an object and a hand, there are usually many grasps that satisfy a desired grasp property. The quality of a good grasp is determined by its ability to take into account disturbance resistance, dexterity, equilibrium, and stability. As a result, to choose the optimal grasp, a quality measure, or an index that quantifies the goodness of a grasp, is used. This measure is commonly referred to as grasp quality measure in the literature [29]. There are other measures for grasp quality, like potential contact robustness (PCR) and potential grasp robustness (PGR), which help in evaluating the contact robustness and grasp robustness, respectively. Contact robustness refers to the distances or how far a contact is from the violation of any contact constraint, whereas grasp robustness concerns how far the grasp is from overcoming the object immobilization constraint, i.e., the ability to resist external disturbances. Details regarding PCR and PGR are included in Appendix B. PGR measure is considered better than the PCR measure because it considers the fact that even if some contact constraints are not satisfied, a hand can perform a stable grasp, whereas PCR is a conservative measure as it states that only if contact forces are applied at all the contact points, the grasp is stable [32, 62]. An improved method of evaluating PGR is discussed in the literature [63], which requires analytically solving a problem that involves minimizing the cost function $V(y)$ described in the next section. This cost function is determined by contact properties such as contact surface geometry and friction coefficient, as well as contact forces. Based on this, I have framed the objective function, which is discussed in the next section.

3.3.1.3 Objective Function

The constraints of contact forces that need to be satisfied while grasping, as defined in Eq. 3.9 and Eq. 3.10, can be expressed as $\boldsymbol{\sigma_{i,f}} = \eta_i \|\boldsymbol{\lambda_i}\| - \boldsymbol{\lambda_{i,n}} < 0$ where $\eta_i = \left(\sqrt{1+\mu_i{}^2}\right)^{-1}$. Also, if I impose bounds on the contact force magnitude, then I can introduce two more constraints: (1) $\boldsymbol{\sigma_{i,m}} = \boldsymbol{f_{i,\min}} - \boldsymbol{\lambda_{i,n}} < 0$ (2) $\boldsymbol{\sigma_{i,M}} = \|\boldsymbol{\lambda_i}\| - \boldsymbol{f_{i,\max}} < 0$, where $\boldsymbol{f_{i,\min}}$: lower bound and $\boldsymbol{f_{i,\max}}$: upper bound on the contact forces. Expressing the above constraints into a single inequality equation gives

$$\boldsymbol{\sigma_{i,k}} = \eta_{i,k} \left\| \boldsymbol{\lambda_i} \right\| + \gamma_{i,k} \boldsymbol{\lambda_{i,n}} + \delta_{i,k} < 0 \tag{3.11}$$

where, $i = 1, ..., n_c$: no. of contacts; k(constraint type) $= f, m, M$ and $\eta_{i,k}, \gamma_{i,k}$, *and* $\delta_{i,k}$ represents constant parameters defined in Table 3.2.

Table 3.2: Co-efficients of constraints

Constraint type	$\eta_{i,k}$	$\gamma_{i,k}$	$\delta_{i,k}$
Friction cone($k = f$)	η_i	-1	0
Minimum normal force ($k = m$)	0	-1	f_i, min
Maximum force module($k = M$)	1	0	$-f_i, max$

A term $\Omega^{\in}_{i,k} \subset R^h$ representing the set $\boldsymbol{y}$ for a given external wrench $\boldsymbol{w}$ is defined, satisfying the constraint in Eq. 3.11 by a margin $\in$: $\Omega^{\in}_{i,k} = \{\boldsymbol{y} | \boldsymbol{\sigma}_{i,k}(\boldsymbol{g}, \boldsymbol{y}) < -\in\}$

For each contact, i and constraint, k the below function can be defined:

$$V^{\in}_{i,k}(\mathbf{w}, \mathbf{y}) = \left\{ \begin{array}{ll} (d\boldsymbol{\sigma}^2_{i,k})^{-1} & \boldsymbol{y} \in \Omega^{\in}_{i,k} \\ a\boldsymbol{\sigma}^2_{i,k} + b\boldsymbol{\sigma}_{i,k} + c & \mathbf{y} \notin \Omega^{\in}_{i,k} \end{array} \right\} \tag{3.12}$$

Friction constraints and contact force magnitude limitations are taken into account in the above cost function. The objective function for the grasp quality measure-based optimization based on the above cost function can be defined as:

$$V(\boldsymbol{w}, \boldsymbol{y}) = \sum_{i=1}^{n_c} \sum_{k=f,m,M} V^{\in}_{i,k}(\boldsymbol{w}, \boldsymbol{y}) \tag{3.13}$$

Then,

$$PCR = \frac{1}{V(\hat{\boldsymbol{y}})} \tag{3.14}$$

and,

$$PGR = \max_{C_j} \frac{1}{V(\hat{\boldsymbol{y}}, C_j)} = \max_{C_j} PCR(C_j) \tag{3.15}$$

$$subject\ to \quad \mathcal{N}(K(C_j)G^T) = 0 \tag{3.16}$$

3.4 Design Methodology

3.4.1 Problem Decomposition

The problem is a multi-objective optimization within the identified design space. Since it is a high-dimensional optimization problem, it is not feasible to search for all the design variables simultaneously. Hence, I solve the optimization problem by decomposing it into a two-step process:

1. Optimization of the contact force distribution based on grasp quality
2. Optimization of the kinematic behavior based on unbalanced torque method

The motive for the decomposition is to mainly split the optimization problem for both the pre-contact and post-contact behavior. The post-contact contact force distribution based on the grasp quality to optimize the moment arms and spring stiffness is undertaken in the first step. Secondly, for the pre-contact kinematic behavior, I take into account the optimization of the spring pre-load angles for optimal grasp preshaping. The details of the two-step optimization are referred to in the next section.

3.4.2 Optimization Methodology

The flow chart of the proposed optimization framework is shown in Fig. 3.4. The optimization starts with pre-specified desired grasps and hand kinematics. First, in Step 1, I solve the optimal contact force distribution post-contact in the optimization framework as detailed in Algorithm 1. I then solve for the optimal parameters that contribute to the kinematic behaviors for grasp pre-shapes in Step 2, explained in detail in Algorithm 3 and Algorithm 2. Finally, I get the optimal actuation parameters for the underactuated mechanism for both contact force distribution and posture shaping based on grasp quality measures. The evolutionary optimization algorithm utilized to obtain the optimal parameters was simulated annealing. It is a meta-heuristic optimizer for global optimization in a large search space of an optimization problem. Normal heuristic optimizer like gradient ascent finds better neighbor solutions than current values and stop reaching an optimal value that does not have neighbors with better solutions. Thus, arriving local optima, but a better solution may be located in other locations, which may be a global optimum that

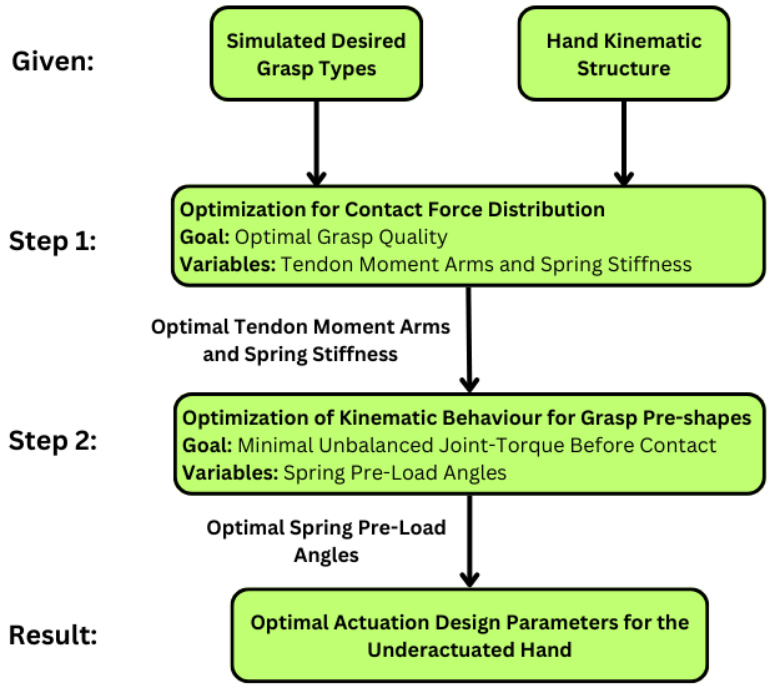


Figure 3.4: Flowchart of the optimization

could be different from the current one. Meta-heuristics utilizes the nearby solutions to explore the search space, and while they, too, search for better neighboring solutions like the simple heuristics, they accept inferior neighbors to avoid getting stuck in local optima. Hence, they are able to find candidates for global optima in the stipulated time.

3.4.2.1 Optimization based on contact force distribution (after contact)

For optimizing the parameters based on contact force distribution after contact, I used grasp robustness as an evaluation property of the generated grasp. Given the grasp space, I perform the optimization given in Algorithm 1.

The algorithm is an optimum global search over different tendon pulley radii (tendon moment arms) and spring stiffnesses for the five grasps that are considered. I calculate the PGR measure for all five grasps. The objective function for the optimization framework is to minimize the overall coefficient of variation (CV), also known as relative standard deviation (RSD), of the PGRs of the five grasps taken together, PGR_{cov}. Minimizing the CV or RSD proves to be beneficial as it aids in maximizing the PGR values to be located near the mean of the PGRs of all the

Algorithm 1 Optimization of the tendon moment arms and spring stiffness

0: Initializing population of random candidates
1: **for** each combination of r and K **do**
2: **for** $k \in \{G_{pow}, G_{pre}, G_{cyl}, G_{obl}, G_{pin}\}$ **do**
3: Calculate PGR_k (discussed in Appendix. 7.7.3)
4: **end for**
Calculate the co-efficient of variation of the PGRs, PGR_{cov}
5: **end for**
6: Select the combination of r and K that gives the minimum of all the PGR_{cov}s obtained, i.e.,
$r_{opt}, K_{opt} = min(PGR_{cov})$

grasps so that outliers have a lesser effect on the overall average metric.

3.4.2.2 Optimization based on unbalanced torque (before contact)

For the calculation of the spring pre-load angles, I used the same method of optimization, i.e., unbalanced joint torque method before contact, as described in [57]. In an underactuated mechanism, where a tendon connects different joints together, the norm of the unbalanced joint torques (before contact) is given as:

$$\delta\boldsymbol{\tau_b} = R(r_1, r_2, ...)\boldsymbol{t} - \boldsymbol{\tau}_s \tag{3.17}$$

where, $\boldsymbol{t}$: unknown tendon tension vector,
R: function of $r_1, r_2, ...$ (the optimized tendon moment arms from Algorithm 1),
$\boldsymbol{\tau_s} = [K_1(\theta_1 + \theta_{01}), K_2(\theta_2 + \theta_{02})...]^T$: spring torques calculated by spring parameters and hand poses. The goal is to find $\boldsymbol{t}$, which would give the least norm of unbalanced joint torque before contact.

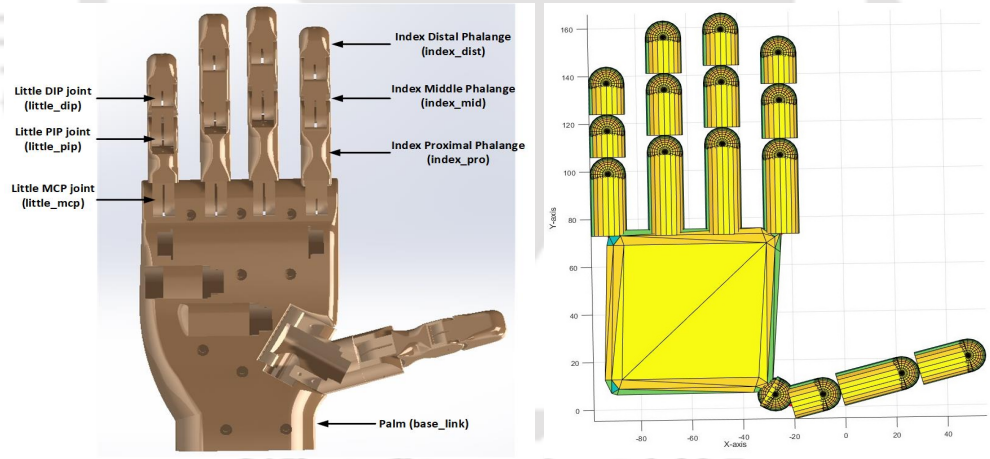
A two-layer framework is utilized to search for the optimal pre-load angles. The outer layer performs an exhaustive search over the pre-load angles. The inner layer, which is a non-convex problem, computes the norm of unbalanced joint torques. The pseudo-code of the optimization framework is shown in Algorithm 3 and Algorithm 2.

Algorithm 2 Calculation of spring pre-load angles

0: Initialize population of random candidates
1: **for** each combination of (θ_0) **do**
2: **for** each grasp type, i **do**
3: Compute $q_i = \|\delta \boldsymbol{\tau_b}\|$ using Algorithm 3
4: **end for**
5: Compute the overall stability metric $Q = \|[q_1, q_2, ...]\|$
6: **end for**
7: Output: θ_{0opt}, i.e., optimal $\theta_0 s$

Algorithm 3 Calculation of stability metric pre-contact

1. find: tendon tensions $(\boldsymbol{t})$
2. minimize: $\|\delta \boldsymbol{\tau_b}\|^2 = \boldsymbol{t}^T R^T R \boldsymbol{t} - 2\boldsymbol{\tau_s^T} R \boldsymbol{t} + \boldsymbol{\tau_s^2}$
3. subject to: taut tendons, i.e., $\boldsymbol{t} \geq 0$



(a) CAD Model (b) Kinematic Model

Figure 3.5: Hand model (a) Solidworks CAD model (b) MATLAB Kinematic model.

3.5 Results and Discussions

A CAD model of the anthropomorphic hand was made in Solidworks (see Fig. 3.5(a)). The measurement of the joint limits and material properties, i.e., poly-lactic acid (PLA plastic), have been added to the assembly. Then, a replica of the

hand model was prepared on MATLAB for further analysis, as shown in Figure 3.5(b).

At first, I built different objects and actuated the joints to generate different grasps, as shown in Fig. 3.6. Then, I utilized the synergy coupling matrix given in Eq. 3.18 to replicate the coupling between the joints of the under-actuated anthropomorphic hand. The synergy coupling matrix is constant and is inspired by the example of the underactuated mechanism in [64].

$$S = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{r1(t)} & -\frac{r2(t)}{r1(t)} & -\frac{r3(t)}{r2(t)} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{r1(i)} & -\frac{r2(i)}{r1(i)} & -\frac{r3(i)}{r2(i)} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{r1(m)} & -\frac{r2(m)}{r1(m)} & -\frac{r3(m)}{r2(m)} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{r1(r)} & -\frac{r2(r)}{r1(r)} & -\frac{r3(r)}{r2(r)} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{r1(l)} & -\frac{r2(l)}{r1(l)} & -\frac{r3(l)}{r2(l)} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix} \tag{3.18}$$

where $r1$, $r2$, and $r3$ are the pulley radii of the proximal, middle, and distal joints, respectively, and t, i, m, r and l represent thumb, index, middle, ring and little fingers respectively.

3.5.1 Numerical Simulation

The results were calculated using the proposed methods of optimization, i.e., based grasp quality measure based methodology and also the one referred to in [57]. The latter was done to compare and validate the results obtained through both the different methodologies (as discussed in the next section). The upper and lower bounds of the pulley sizes depended on the anthropomorphic hand design, whereas the bounds of spring stiffness depended on the mounting area and the allowed extension. The bounds on the pulley sizes were kept between 5 to 9 mm, and that of spring stiffness was kept at 0.05 to 0.10 N/mm. Further, the spring pre-load angles are bounded with respect to the maximum allowed torsional angle for the ability to attain different grasp shapes and generate enough restoring torque. The results of the optimization of pulley radius and spring pre-load angles through grasp quality methodology are shown in Table 3.3 and that of unbalanced torque is shown in

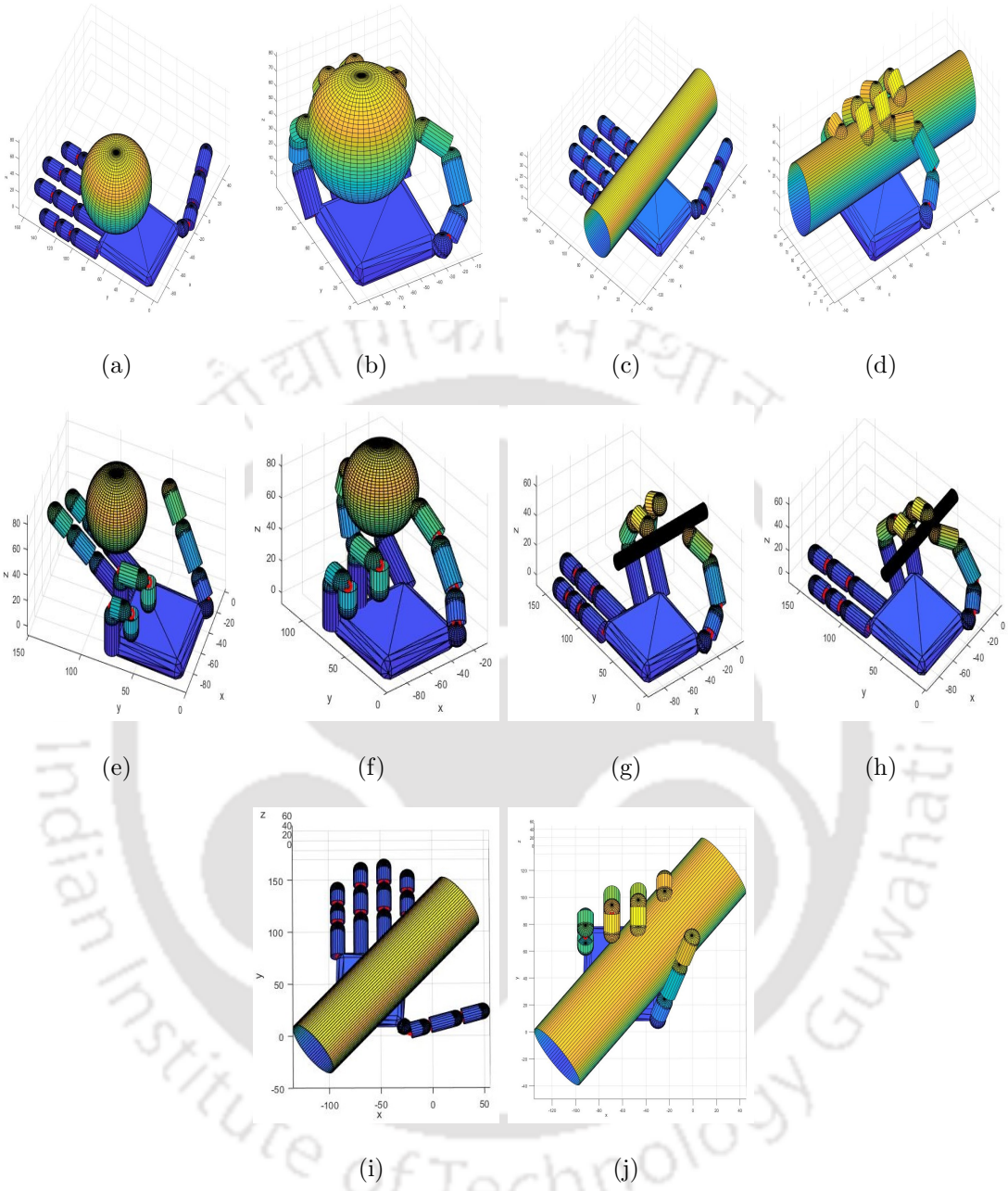


Figure 3.6: Grasps generated in MATLAB (a) Power (initial) (b) Power (final) (c) Cylindrical (initial) (d) Cylindrical (final) (e) Precision (initial) (f) Precision (final) (g) Pinch (initial) (h) Pinch (final) (i) Oblique (initial) (j) Oblique (final)

Table 3.4. The spring stiffness obtained through the former and later optimization methodologies were 0.10 N/mm and 0.05 N/mm, respectively. Fig. 3.7 shows the simulated annealing results of the optimized pulley radius of the underactuated an-

thropomorphic hand obtained as per Step 1, Fig. 3.4, flowchart of optimization. Fig. 3.8 shows the results of the optimized spring pre-load angles of the underactuated anthropomorphic hand obtained as per Step 2, Fig. 3.4, flowchart of optimization. The evolutionary algorithm results show the Pareto optimality or Pareto efficiency of the candidate optimal solutions, as no other change in the magnitude of design variables leads to improved optimal function value.

Table 3.3: Optimized parameters of the underactuated anthropomorphic hand through PGR

Parameter	r_{tmcp}	r_{tpip}	r_{tdip}	r_{imcp}	r_{ipip}	r_{idip}	r_{mmcp}	r_{mpip}	r_{mdip}	r_{rmcp}	r_{rpip}	r_{rdip}	r_{lmcp}	r_{lpip}	r_{ldip}
Value (mm)	5	7	9	5	9	8	9	7	9	5	5	5	5	5	9
Parameter	θ_{tmcp}	θ_{tpip}	θ_{tdip}	θ_{imcp}	θ_{ipip}	θ_{idip}	θ_{mmcp}	θ_{mpip}	θ_{mdip}	θ_{rmcp}	θ_{rpip}	θ_{rdip}	θ_{lmcp}	θ_{lpip}	θ_{ldip}
Value (rad)	0.1	0.08	0.4	0.25	0.20	0.11	0.06	0.01	0.58	0.00	0.02	0.36	0.00	0.32	0.07

Table 3.4: Optimized parameters of the underactuated anthropomorphic hand through unbalanced torque method

Parameter	r_{tmcp}	r_{tpip}	r_{tdip}	r_{imcp}	r_{ipip}	r_{idip}	r_{mmcp}	r_{mpip}	r_{mdip}	r_{rmcp}	r_{rpip}	r_{rdip}	r_{lmcp}	r_{lpip}	r_{ldip}
Value (mm)	7	7	7	9	7	9	6	9	7	6	7	8	9	7	8
Parameter	θ_{tmcp}	θ_{tpip}	θ_{tdip}	θ_{imcp}	θ_{ipip}	θ_{idip}	θ_{mmcp}	θ_{mpip}	θ_{mdip}	θ_{rmcp}	θ_{rpip}	θ_{rdip}	θ_{lmcp}	θ_{lpip}	θ_{ldip}
Value (rad)	0.01	0.00	0.40	0.23	0.13	0.40	0.38	0.08	0.78	0.00	0.00	0.00	0.06	0.00	0.31

3.5.2 Validation based on Grasp Wrench Space

For validation of the results obtained through both methodologies, I provided a physics engine-based simulation of the hand grasping an object in the ROS-Gazebo environment. The gazebo world environment is shown in Fig. 3.9. To simulate the actuation through the tendons, I enforce under-actuation by coupling joint torques $\boldsymbol{\tau} \in R^{3d}$, where d is the number of fingers of the hand, according to

$$\tau = Rf_t - Kq - \boldsymbol{\tau_d}\dot{q} \tag{3.19}$$

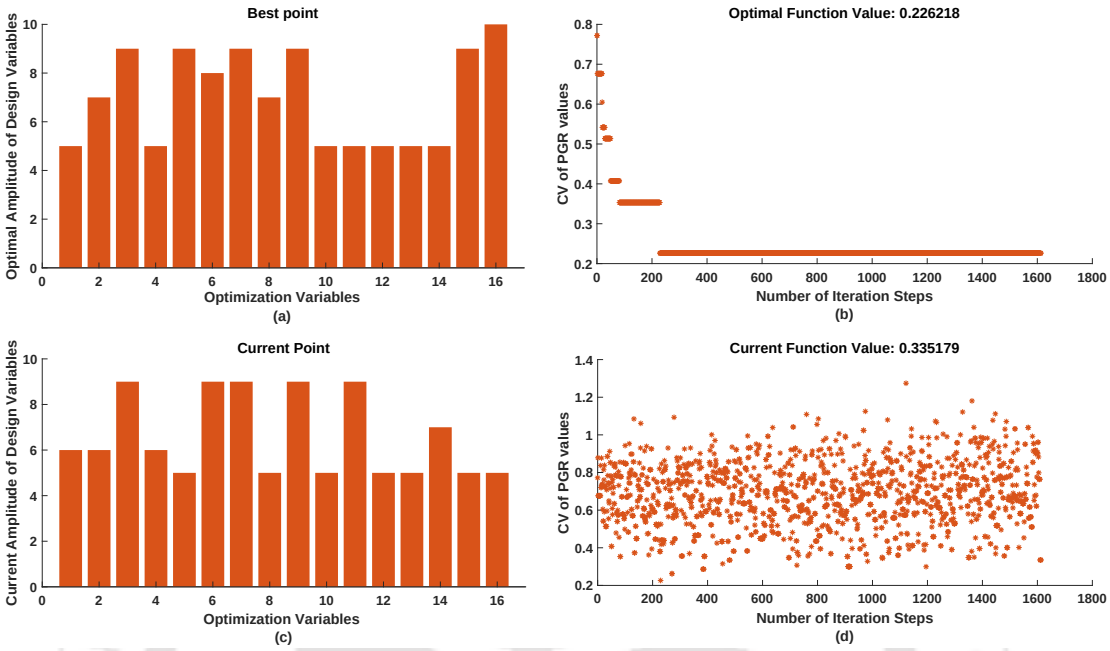


Figure 3.7: Optimization plot of the various Pareto-dominant solutions of pulley radius, Refer Table 3, Row 1 (Optimization performed as per Step 1, Fig 4: Flowchart of the optimization): (a) Magnitudes of Optimal Design Variable (1-15: Pulley Radius (units: mm), 16: Spring Stiffness (units: N/m, values: multiplied by 10)) (b) Optimal values of CV of PGRs (c) Magnitudes of Current Design Variable (1-15: Pulley Radius (units: mm), 16: Spring Stiffness (units: N/m, values: multiplied by 10)) (d) Current values of CV of PGRs

where R is a $3d \times d$ matrix determining the distribution of the tendon forces on the joints. Matrix K is a $3d \times 3d$ diagonal stiffness matrix where its values are defined by the coefficients of the springs mounted on each joint. Vectors $q \in R^{3d}$ and $\dot{q} \in R^{3d}$ are the joint angles and angular velocities, respectively, measured within the simulator. $\tau_d(\dot{q})$ is the damping vector and can also include torsional friction at the joints. In this formulation, the d input values are the actuator's angles ϕ, and they determine the $3d$ finger joint torques and, by that, simulate under-actuation. For the same dynamics, I put the same effort into the actuator for both the optimized hands. Contact sensors were embedded in the links of the fingers. The embedded sensors helped us to generate the grasp wrench data and subsequently, the wrench measures were received. The final grasp configuration is shown in Fig. 3.9(b).

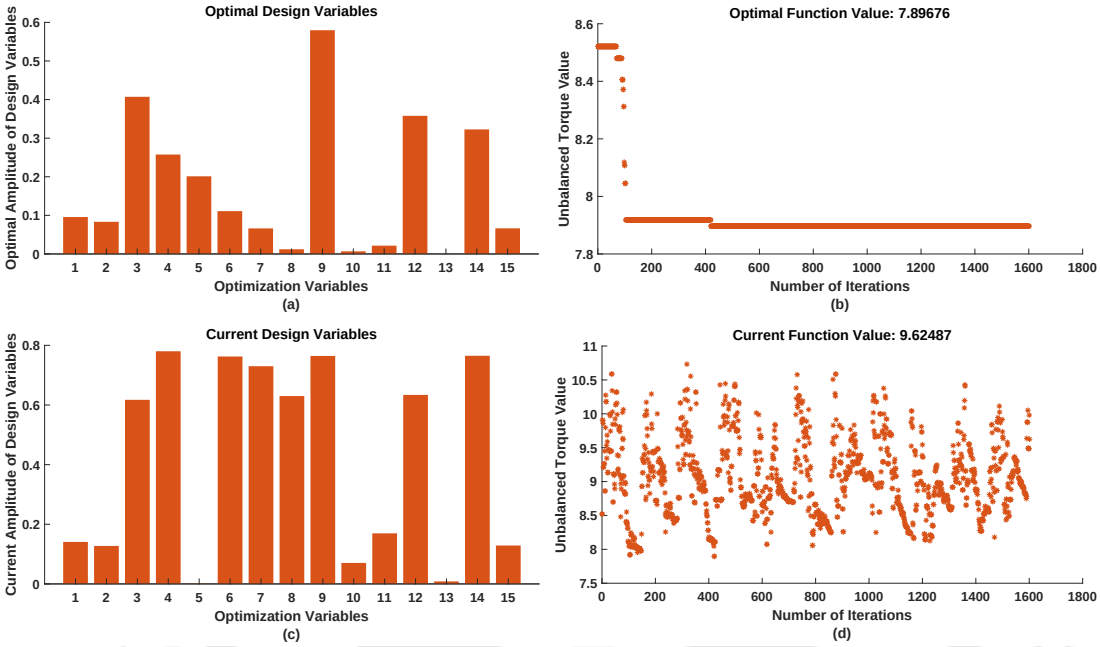
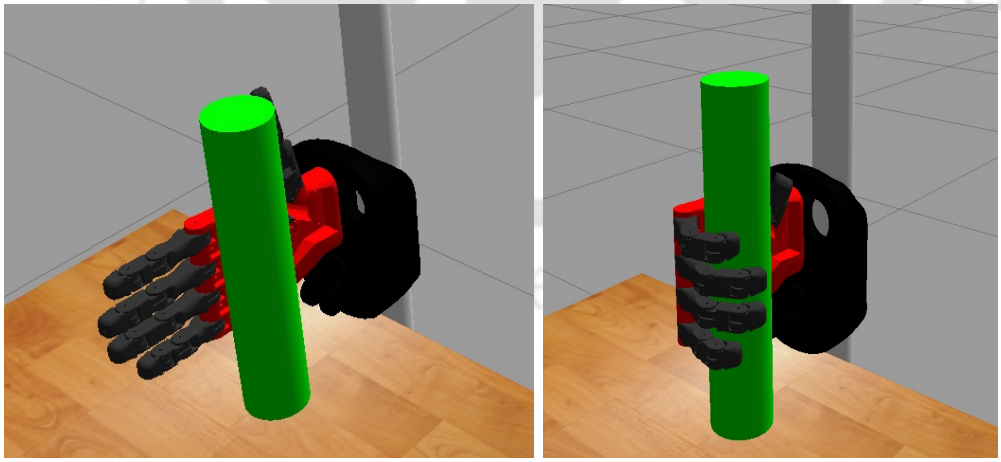
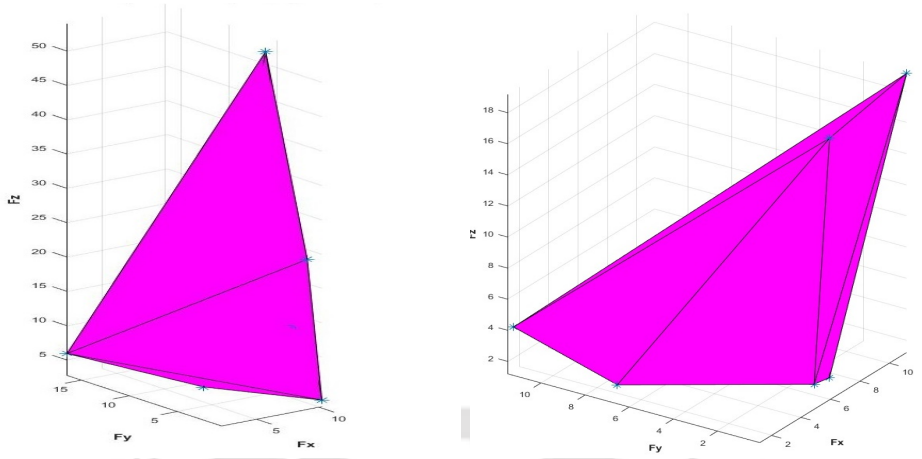


Figure 3.8: Optimization plot of the various Pareto-dominant solutions of Spring Pre-Load Angles, Refer Table 3, Row 2 (Optimization performed as per Step 2, Fig 4: Flowchart of the optimization): (a) Magnitudes of Optimal Design Variable (b) Optimal Unbalanced Torque Value (c) Magnitudes of Current Design Variable (d) Current Unbalanced Torque Value



(a) ROS Initial Position (b) ROS Final Position

Figure 3.9: Simulated anthropomorphic hand in gazebo simulation engine



(a) GWS of PGR optimized, $Q_\epsilon = 10.15$ (b) GWS of optimization through [57] , $Q_\epsilon = 4.49$

Figure 3.10: Grasp Wrench Space showing the convex hull computed over the primitive wrench space. The grasp quality measure, $Q_\in$ and the number of wrench vectors, n are also shown

Grasp wrench space (GWS) of the grasped forces for the two different underactuated hand designs, obtained through optimization by both the grasp quality measure and unbalanced torque methodology, was used for comparison. From the contact sensors embedded at the fingers, I get the set of primitive wrenches, $\boldsymbol{\omega}$ obtained

$$\omega_0 = \sum_{i=1}^{n} \omega_i = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_{ij} \omega_{ij} \tag{3.20}$$

where, $\alpha_{ij} \geq 0, \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_{ij} \leq 1, \omega_i$ is wrench produced by finger i at the i^{th} contact and can be expressed as the positive linear combination of the wrenches ω_{ij} produced by forces f_{ij} (m friction cone edges)

The convex hull of the wrenches (GWS) is given as:

$$GWS = CH\left(\bigcup_{i=1}^{n} \{\omega_{i1}, ..., \omega_{im}\}\right) \tag{3.21}$$

where, GWS is the convex hull (CH) of the primitive wrenches, $\boldsymbol{\omega}$ [29]. The grasp quality measure is then computed as the radius of the largest ball inscribed inside the convex hull of GWS, which is referred to as the largest-minimum resisted wrench ($Q_\in$).

$$Q_{\in} = \min_{\boldsymbol{\omega} \in GWS} \|\boldsymbol{\omega}\| \tag{3.22}$$

The convex hull of the grasp wrench space (GWS) obtained through the PGR optimization and the unbalanced torque method is shown in Fig. 3.10(a) and Fig. 3.10(b) respectively. The $Q_{\in}$ measure came out to be 10.15 for the PGR optimized hand and 4.49 for the unbalanced torque optimized hand. The Q_{ϵ} measure provides the magnitude of the largest perturbation wrench that the grasp can resist in any direction. Hence, the higher the value of Q_{ϵ}, the more external wrench the grasp can resist with limited contact forces overall wrench directions. Due to its direct relation with the wrench domain, its result can be used to verify and cross-check various grasping tasks [31]. The results show the promise of better grasp quality for the underactuated hand-optimized through grasp quality measure.

3.6 Chapter Summary

In this chapter, I have optimized the actuation parameters of the underactuated tendon-driven hand, viz. pulley radii, pre-load angles, and spring stiffness. At its core, a soft synergistic compliant grasping model was exploited to generate the different human grasps. In contrast to previous work in the literature, I tuned the optimization framework for the underactuated anthropomorphic hand based on the grasp quality measure by utilizing the potential grasp robustness measure as the objective function. The results of the optimization proved the efficacy of the framework in arriving at a solution that produces good grasps. Further, validation of the results was performed on a ROS-Gazebo environment. The optimization method seems to be encouraging, as evident from the results of grasp wrench space analysis. For future work, other grasp types and quality measures can be taken into account for comparative analysis. Further, lower-level grasp control, like slippage and deformation prevention, can be embedded into the optimization objective function so that the optimization framework can automatically take into account the manipulation behavior in the design stage.

Chapter 4

Slip Prevention with Reinforcement Learning Based Control

4.1 Introduction

Slippage of the grasped object is a quick and transient event caused by the interruption of touch between the end-effector and the grasped object during any grasping maneuver, leading to the contact forces sliding away from the friction cone. To appropriately regain stability, the grasp force must be adjusted [17]. The sense of slippage is one of the distinctive features of human tactile perception for grasping and manipulation. Touch sensation permits the identification of different properties of the grasped object, e.g., its shape, size, weight, stiffness, friction, roughness, etc. [17]. According to these characteristics, the grasping force on a robotic hand can be suitably controlled to prevent the object from slipping. One of the main requirements of prosthetic hand users to accept their devices is slippage avoidance [2]. With a wide range of cutting-edge sensors and control approaches proposed to assist during grasping and manipulation, the field of research for slippage detection and control is broad and increasing.

Humans can actively perceive certain properties of the grasped object, owing to the mechanoreceptors present. Hence, if any disturbances occur while grasping and

manipulating the grasped object, they are able to regulate their grasp force levels effortlessly. Such a reaction time to regulate the grasping forces usually ranges approximately around 0.1 to 0.2s [35,65]. They can do this by adapting their grasping force based on the friction coefficient which they have learned or experienced throughout their lifetime [37]. This reflex to adaptively adjust the grasping forces to prevent slippage and deformation is known as the bionic reflex. Hence, learning friction coefficient and object stiffness is essential to grasp unknown objects stably. Commercially available prostheses primarily provide force-sensor data and are devoid of providing information on friction, surface properties, etc. Advanced sensors with high-spatial resolution tactile arrays like BioTac (SynTouch Inc, CA, USA) give a lot of information about object properties. However, they are costly [66] and require exploration to learn the unknown object properties [67]. This incurs training time as well as grasping and manipulation time. Further, it is necessary to transfer these sensed signals to the users as feedback so that they can stabilize the grasped force based on the incoming signal.

Although the prosthetic industry is extensive, it is surprising that most commercial prosthetic hands still lack tactile sensory feedback, making them inadequate for slippage avoidance. Researchers continue to face significant obstacles in their efforts to adapt steady grasp via controlling slippage [35]. Hence, providing slippage prevention capabilities with adequate grasp response time [68] has become the need of the hour for upper-limb prostheses. The research on slip-sensing abilities for robot hands also caters to the same argument. It is important to incorporate anti-slip properties to provide advanced capabilities to hand prostheses, just like their human counterpart, to foster widespread user-acceptance of bionic hands [34].

Explicit sensory feedback based on advanced tactile sensors and stimulation techniques in user-in-the-loop prostheses can make the user perceive incoming disturbances [69]. The delay between the user's desire to manipulate the prosthesis and their reaction time would be greater, even if this feedback may prompt amputees to modify their grasping forces. So, for user-in-the-loop prosthesis, getting sensory input based on slippage, acquiring bio-signals, and decoding the biosignal for grasp execution takes time. Hence, adopting a low-level reflex control strategy is a good choice to compensate for the delay time caused by biological signal decoding techniques [8]. A low-level controller is the part of the control framework in a hierarchical controller, which looks after the execution section, associated with error calculations and feedback control loops (in my case, the final grasp force control loop) [9]. The

user-outside-the-loop prosthesis is better at generating an autonomous slip detection algorithm so that the slippage control command stays within the prosthesis with a higher degree of autonomy without involving the user. This way, I can achieve the slip prevention reaction time close to the humans.

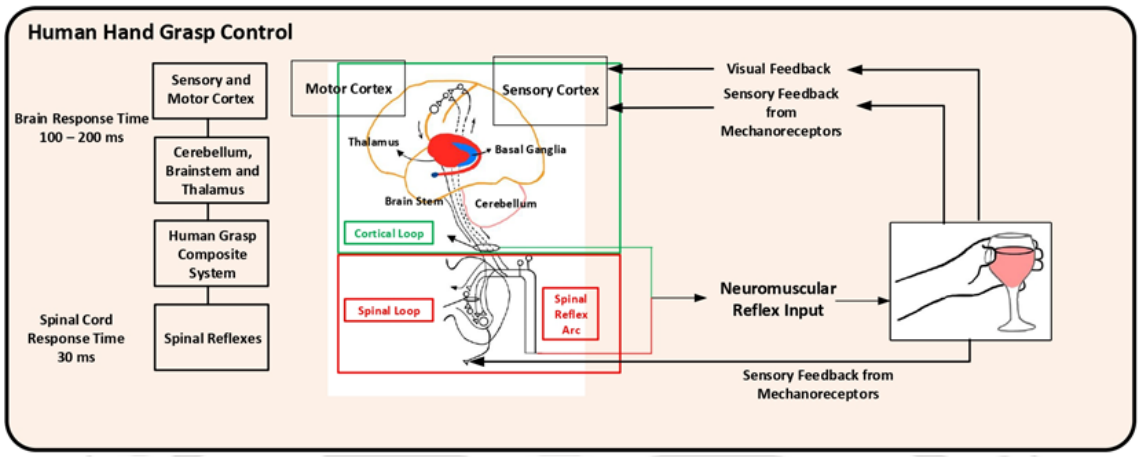


Figure 4.1: Hierarchical motor control scheme for a human finger performing a grasp of a fragile object

As depicted in Fig. 4.1, the motor control for grasping and manipulating any object occurs in a hierarchical order. The brainstem and cerebral structures (cerebellum, basal ganglia, thalamus, etc.) are where the motor commands are regulated, selected, coordinated, and delivered back through the spinal cord. The sensory and motor cortex processes the sensory input and motor outputs. The low-level system is made up of the muscles and sensors located on each finger of the hand. Assuming the user desires to perform a grasp, a composite system forms combining the high-level and lower-level systems to execute the grasp reflexes. Further, the muscle and spinal cord adjust finger movements at the lower level through the spinal reflex arc [9]. The brain stem, association cortices, and cerebral structures control and coordinate the motor instructions and sensory data at the mid-level. Ultimately, sensory data is gathered at a high level, and commands for muscle activity are created [70]. The high-level sensory pathways, which travel to the brain and back, take about 100 ms to react to perturbations, but the lower-level spinal reflexes do so in about 30 ms [71]. It is, therefore, possible to create an intelligent slippage controller for the anthropomorphic hand by drawing inspiration from the spinal reflex's autonomy.

Further, for the grasp reflex controllers to work appropriately, it is important

that they can easily discriminate the frictional properties of unknown objects. It is found that humans display a frictional memory system, increasing the grasped force during object slippage, based on previously encountered frictional characteristics [37]. However, storing such a large dataset of object properties could be more inefficient. Hence, if the frictional features can be estimated through tactile sensing in the user-outside-the-loop prosthesis, then it may achieve slippage prevention at the time duration as its human counterpart. Building a knowledge base of previously encountered objects about object properties and then using threshold-based slip detection based on a data-driven approach is prevalent in the literature. However, they require large datasets. As neural networks are very efficient in learning features, they are used broadly in discerning slip as they compensate for the advanced tactile sensors and hardware non-linearities [72]. To capture the slippage feature using the advantage of the neural network, I incorporate deep learning into the RL solution.

Another important dimension to consider during slippage prevention is grasp force control. The grasping force may be inappropriate, and unexpected disturbances may occur, which may, in turn, cause slip. So, it is essential to design robust adaptive low-level controllers that may avoid slippage during unexpected disturbances, thus stabilizing the grasp [37]. However, a bionic hand trained well in the lab environments may fail in the real world because of its encounter with different scenarios. It is better to train the controller as close to real-life scenarios as possible to account for situations that may occur in the testing world. A controller trained well in the physics simulator under different disturbances may perform well in the testing world, increasing the policy generalization capability to varying environment conditions [47]. However, in the physics simulators, the dynamics may need to be accurately represented; hence, there might be a mismatch between the trained simulated and testing worlds. This gap between the representation is coined as the sim-to-real or sim-to-sim gap depending on the training and testing environments [73]. To facilitate learning efficient robot control policies by increasing the generalization capability, I propose to adopt Domain Randomization (DR) in this work.

The main contribution of this chapter includes (i) the design of an approach for slip detection in the bionic reflex framework, (ii) a model-free novel approach for autonomous slip prevention controller without the requirement of human intervention, (iii) to evidence the reliability of the slippage control algorithm by increasing the generalization capability through DR. The rest of the chapter is as follows: Section.

4.2 provides a brief study of the background and related works, relevant research topics encompassing slip detection and control algorithm techniques, and avenues of DR in Reinforcement Learning (RL). In Section. 4.3.1, I provide an overview of the problem formulation. In Section. 4.4, I describe the methodologies for designing the proposed controller and the RL grasping policy. The results of the proposed slippage controller are summarized in Section. 4.5. In the end, in Section. 4.6, I provide the summary of the current work.

4.2 Background and Related Work

4.2.1 Slip Detection Techniques in Grasping

Slips can be detected through friction-based methods using multi-axial force sensing to determine the friction coefficient. Differentiation and derivative techniques on tactile sensor signals and implementing artificial neural networks and other feature learning-based slip-predictors are also employed to classify slippage [34]. Building a database of previously encountered objects is cumbersome and inefficient. To solve this issue of requiring a large dataset, one-shot learning (OSL) was implemented to handle slippage detection and control [74]. Other methods of detecting slip include vibrations, optics, acoustic signals, velocity, thermal, magnetic, etc. Sensors detecting slippage can be classified into pressure-resistance, optical, piezoelectric, inductive, capacitive, thermal, and ultrasonic sensors [36]. Advanced signal processing techniques such as Fourier transforms, signal power, wavelet decomposition, etc., which retrieve much-hidden information, are employed to detect slippage [35]. Further, filtering techniques utilizing high-pass, low-pass, bandpass, and Kalman filters can be seen in the literature. However, most of the abovementioned techniques detect slippage with a definite ON-OFF procedure by thresholding the perceived tactile sensing signals. As a result, most of the above methods require human intervention in the slip detection stage like thresholding/labeling/knowing object properties, which undermines the ability to automatically detect slips in unknown and unforeseen grasped objects.

4.2.2 Slip Prevention Techniques in Grasping

Slip prevention usually can be of two types: 1) reactive slip prevention and 2) proactive slip prevention [75]. Reactive slip prevention involves controlling slip at a low level by detecting slippage signals. In contrast, a proactive slip informs the user of an oncoming slip that is about to happen before the slip even occurs. There are three broad categories of slip detection research. First is gross slip, where the whole object surface slips; secondly, there is incipient slip, where part but not the whole of the object slips; and finally, there is slip prediction, where tactile features are extracted for prediction of the occurrence of slip [76]. There is research into both model-based and model-free approaches, which further diversifies the techniques for detecting slips. In model-based approaches, friction cones and beam bundle models have been employed to detect slippages. In model-free methods, data-driven approaches like supervised learning and Deep Learning (DL) techniques have been applied [76].

Even if slip may have been detected perfectly, many limitations and challenges exist to integrate the slip signals with the control algorithm of the bionic hand [35]. Many researchers have attempted to prevent slippage using advanced control algorithms. They mainly include closed loop controllers based on linear control models like PID, non-linear control like sliding mode control, fuzzy control, model predictive controls, advanced model-based control methods, and model-free controllers [36]. Typically, the closed-loop control structure of slip prevention includes either desired force control or position control, as shown in Fig. 4.2. However, implementing a position-based controller is usually not preferred, as the deformation of objects varies largely, owing to their different inherent stiffness [36]. In general, force and position control are usually implemented separately in the inner and outer loop, typically known as force control with inner position loop [77, 78] (shown in Fig. 4.3). The problem of deformation, though tackled in the next chapter, I have included this literature here for the completeness of the state-of-the-art as the current chapter only involves work on slippage of rigid bodies.

Unexpected disturbances may introduce unwanted slippages; hence, researchers have adopted robust adaptive control methods. Adaptive sliding mode controllers have been implemented to grip the grasped object without slippage [37]. However, it is implemented by thresholding slippage signals from the force sensor of the gripper. Nevertheless, such robust algorithms must be transformed into real anthropomorphic

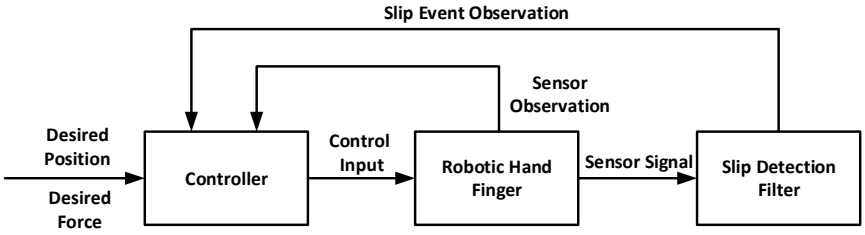


Figure 4.2: Slippage avoidance closed-loop control structure [36]

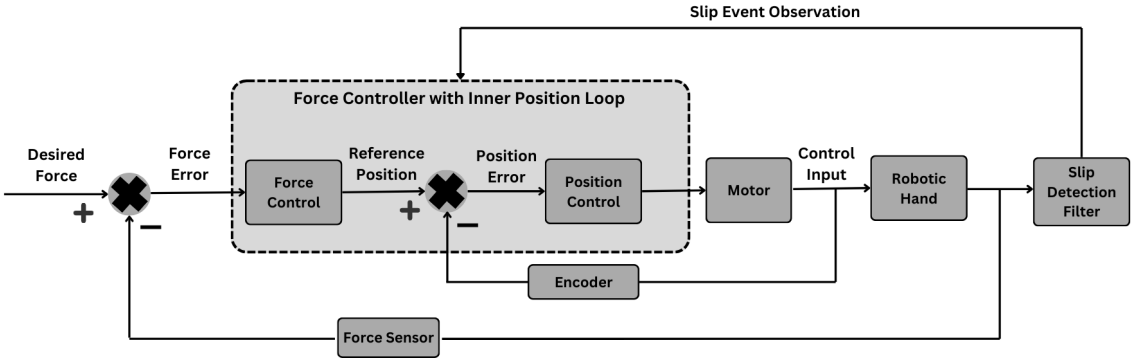


Figure 4.3: Slippage avoidance by force control with inner position loop

hands for implementation in bionic hands.

4.2.3 Reinforcement Learning with Domain Randomization

When it comes to autonomous control, RL comes as a potent tool that can learn their ideal behavior through interaction with the environment. State-action pairings, usually called policies, which the RL agent learns through trial and error, dictate its behavior. For several well-known RL problems, such as pole balancing, mountain car, or grid search, where state and action spaces are small enough, policies can be approximated by tabulation and handled as discrete variables. However, the tabular approach becomes ineffective and unwieldy for more complex systems due to the increasing dimensionality of the state and action spaces [79]. In difficult cases, such as robot control, treatment of the states and actions as continuous variables is a more suitable way to solve the problem [80, 81]. For this, either approximators of the policy are needed, or approximators of a supporting component like a value function, or both may be needed. Deep Reinforcement Learning (DRL), which uses Deep Neural Network (DNN) as approximators can be used to create RL

controllers [82].

DRL, on the other hand, requires a substantial amount of training time when applied to a robot in real-time. Furthermore, due to the intrinsic trial-and-error nature of *RL*, real-world training, especially during the state and action space exploration phase, can result in unsafe robot behavior. Therefore, a quick and secure robot training technique is needed. Simulators can mimic the actual behavior as closely as feasible once they have a model of the robot dynamics. Simulators thus enable the resolution of training-related problems and the detection of incorrect behaviors. Theoretically, simulators make controller training faster and safer by applying learned rules to the actual system. For better generalization capacity, however, it only works when the simulator is provided with a sufficient representation of the actual robot and its surroundings.

Methodologies to improve the generalization capability of the RL agent include: (i) *DR*, (ii) Adversarial *RL*, (iii) Transfer Learning [47]. *DR* is a technique to bridge the sim-to-real reality gap, akin to robust control in control theory with uncertainties present. It focuses on designing controllers that maintain specific properties despite tolerable parameter variations and noise [47]. In adversarial reinforcement learning (ARL), the idea is to enhance robustness and controller transferability by training the controller in various environment models, not just one. ARL achieves diversity by training an adversarial sub-agent to generate challenging models that minimize cumulative rewards [47]. Transfer learning is a machine learning technique where knowledge gained from solving one task is applied to a different but related task, typically by using a pre-trained model as a starting point and fine-tuning it on the new data. It allows models to leverage existing knowledge to improve performance on new tasks, particularly when labeled data is limited. In the current work, I will attempt to improve the generalization ability using *DR*. To the best of the authors' knowledge, such *DR* methodologies to bridge the generalization gap have yet to be applied to slip prevention techniques.

4.3 Problem Scope and Governing Assumptions

4.3.1 Problem Formulation

The goal of this chapter is to design an autonomous slippage-prevention controller to prevent incipient slips during grasping and lifting. As mentioned in 4.1, if I can replicate the spinal reflex loop, then I would be able to achieve an intelligent

slippage controller. I aim to achieve it by designing an RL-based slippage prevention controller as shown in Fig. 4.4. The problem of grasping and lifting the object is presented as a Markov Decision Process (MDP). The detailed formulation of the grasping MDP is given in the next section. Slippage occurring during the lifting of the grasped object has been detected using Discrete Wavelet Transform (DWT). The advantage of such a methodology for slip discrimination lies in the ability to exclude thresholding the slip signals by human intervention. In this current chapter, the formulation of the bionic reflex control has been narrowed down by considering only slippage detection and prevention. Here, I consider rigid body grasping and slippage prevention. The optimized policy will produce the most optimal torques, generating robust fingertip forces for slip prevention.

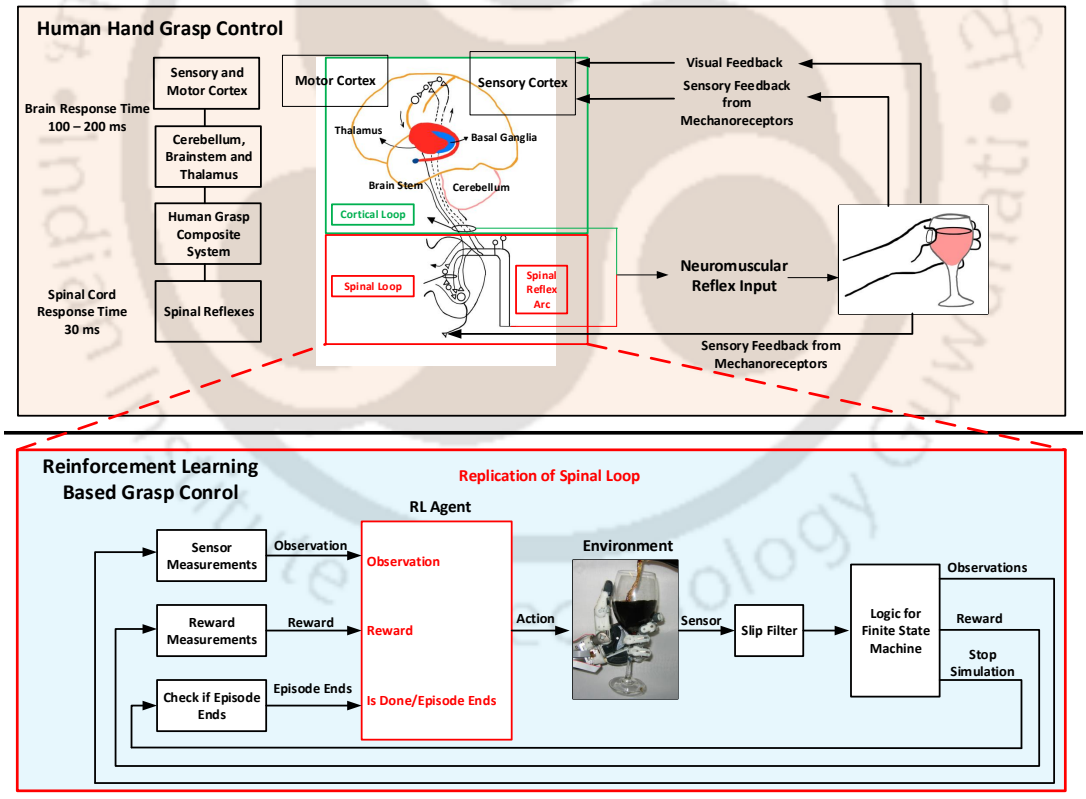


Figure 4.4: The RL based grasp control replicates the spinal loop to integrate it in a user-outside-the-loop prosthesis.

4.3.2 Underlying Assumptions

The development of the reinforcement learning (RL)-based slip prevention framework is grounded on a set of assumptions and simplifications aimed at maintaining the tractability of the training process while ensuring a meaningful level of generalization. This section elaborates on these assumptions and the corresponding implications for real-world deployment.

Finger Material and Contact Surface Properties

In simulation, the gripper fingers are modeled with a uniform material property—typically with moderate friction coefficients representing rubber-like contact behavior. However, in real-world scenarios, finger pad materials can vary widely (e.g., silicone, rubber, hard plastic), significantly affecting contact dynamics. The RL policy trained in simulation assumes a relatively high-friction contact, which supports stable grasping and minimizes slippage during manipulation.

To evaluate the sensitivity of the learned policy to different contact surface properties, domain randomization was employed by varying friction coefficients, mass and gravity within a bounded range during training. This strategy helps the policy become more robust to material variation. Nonetheless, extreme changes in material properties (e.g., transition from rigid to soft contacts) may still require policy fine-tuning or re-training due to their effect on slippage onset and force distribution which is tackled in the next chapter.

Underactuation Topology

The RL framework assumes a specific underactuation topology—typically a tendon-driven mechanism with passive compliance in one or more joints. This topology influences the way forces are transmitted through the finger and how the hand reacts during contact with the object.

The policy is trained under the assumption that the underactuation mechanism remains consistent across different episodes. Although domain randomization introduces variability in material properties, it does not fully capture structural changes in underactuation design, such as switching from a single-tendon to a dual-tendon architecture or from compliant joints to rigid linkages, joint stiffness and actuation delay. Therefore, the generalizability of the trained policy to significantly different underactuation topologies remains an open challenge and warrants future study.

Given that envelope grasping involves shape-based interactions and introduces distinct modeling challenges and design considerations, it was deliberately excluded

from this study to maintain a focused investigation on a specific tendon-driven, compliant underactuation system as mentioned in Chapter 3.

Object Properties and Environment Conditions

Rigid objects with well-defined surface normals are used during training, assuming negligible deformability. Additionally, gravity and contact forces are modeled using the default physics engine with accurate parameterization. Environmental variability such as external disturbances, frictional properties, or variable object mass distribution is partially addressed through stochastic perturbations during domain randomization.

4.4 Design Methodology

4.4.1 Slip Prevention based Bionic Reflex Grasping Policy

As the grasping problem is posed as a MDP and an explicit predictive model of the grasping environment is not defined, I am framing the entire problem as a Model Free Reinforcement Learning ($MFRL$). $MFRL$ learns the policy by directly interacting with the environment; mapping from states to actions [83]. Moreover, the observation states that is required in slip prevention is continuous. Also, the actions to be provided to the joints of the anthropomorphic hand is continuous. Here the state space and the action spaces are both continuous. Hence, I chose Deep Deterministic Policy Gradient ($DDPG$) as the RL algorithm for controlling slip prevention while grasping.

Markov Decision Process (MDP)

The entire RL problem of preventing slippage when grasping is described as a discounted MDP with a finite horizon. An MDP is a discrete-time stochastic control process in which state transitions obey the Markov property, i.e., at each time-step, t; the agent observes something in the environment, $o_t \in O$, takes action $a_t \in A$, and receives a reward $r_t \in R$. The agent acts according to a stochastic policy, $\pi\left(a_t|O_t\right)$, which is a distribution across actions based on several recent observations. The main RL objective is finding a policy, π, that maximizes the expected sum of discounted rewards over a limited trajectory. The optimal approach, i.e., the action value function, is defined as:

$$\pi^* = \arg\max_{\pi} \mathrm{E}_{\tau(\pi)} \left[\sum_{t=0}^{T} \gamma^t r_t \right] \tag{4.1}$$

where $\gamma \in (0, 1)$ is the discount factor, τ is the trajectory distribution under policy π.

Algorithm 1: Slip Prevention based Bionic Reflex Control

Input: Grasping environment in ROS-Gazebo

Output: Optimal joint torques(τ) to lift the grasped object without slippage

begin

- **foreach** *episode* **do**
  - Initialize S (observations of sensors from the grasping environment);
  - **foreach** *step of episode* **do**
    - Generate actions based on **Algorithm 2**;
    - Grasp the object;
    - Lift the object;
    - Slip discrimination during object lifting;
    - **if** *Object Drops* **then**
      - episode ends;
      - reset simulation; (Is Done Checked)
    - **else**
      - **if** *Slip Occurs* **then**
        - $\tau \leftarrow \tau + \Delta\tau$;
        - continue;
      - **end if**
    - **end if**
    - End lifting motion (Grasp Successful);
    - **if** *Grasp Successful* **then**
      - episode ends;
      - reset simulation; (Is Done Checked)
    - **end if**
  - **end foreach**
  - Until Episode Time is terminal
- **end foreach**

end

Algorithm 2: DDPG for estimating optimal policy $\pi \approx \pi_*$ [84]

Input: Algorithm parameters: step size $\alpha \in (0, 1]$, random process N, training batch size n, target network update parameter τ;

Initialize: Critic network $Q(s, a|\theta^Q)$ and actor-network $\mu(s, \theta_\mu)$ with random weights θ_Q and θ_μ, target networks Q' and μ' with weights $\theta'_Q \leftarrow \theta_Q$ and $\theta'_\mu \leftarrow \theta_\mu$, replay buffer R

Output: Optimal Agent Policy

foreach *episode* **do**

Initialize a random process N for action exploration;

Receive initial observation s_t;

foreach *step of episode* **do**

Select action $a_t = \mu(s_t|\theta_\mu) + N_t$ according to the current policy and exploration noise;

Execute action a_t and observe reward r_t and observe new state s_{t+1};

Store transition (s_t, a_t, r_t, s_{t+1}) in R;

Sample a random mini-batch of n transitions (s_i, a_i, r_i, s_{i+1}) from R;

Set $y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1}|\theta^{\mu'})|\theta^{Q'})$;

Update critic by minimizing the loss: $L = \frac{1}{n}\sum_i (y_i - Q(s_i, a_i|\theta^Q))^2$;

Update the actor policy using the sampled policy gradient: $\nabla_{\theta^\mu} J \approx \frac{\sum_i \nabla_a Q(s,a|\theta^Q)|_{s=s_i, a=\mu(s_i)} \nabla_{\theta^\mu} \mu(s,\theta^\mu)|_{s_i}}{n}$;

Update the target networks:

$\theta^{Q'} \leftarrow \tau\theta^Q + (1-\tau)\theta^{Q'}$

$\theta^{\mu'} \leftarrow \tau\theta^\mu + (1-\tau)\theta^{\mu'}$;

$s_t \leftarrow s_{t+1}$

end foreach

Until s is terminal

end foreach

The algorithm for the *RL* framework is discussed below. Algorithm 1 explains the overall bionic reflex framework. The *DDPG* algorithm (Algorithm 2) generates the optimal grasping forces. The presented algorithm, Algorithm 1, aims to enable a robotic system to lift objects without slippage. Operating within a ROS-Gazebo environment, the algorithm iterates through episodes, initializing sensor observations and executing actions based on the Algorithm 2. It proceeds to grasp and lift

objects, monitoring for slippage during the lifting process. If slippage occurs, the algorithm adjusts joint torques to prevent further slipping. Successful grasps lead to the termination of episodes, while unsuccessful attempts trigger simulation resets. This iterative process continues until the terminal time for each episode is reached, ultimately facilitating optimal joint torque generation for slip-free object lifting.

Algorithm 2 outlines a Deep Deterministic Policy Gradient (DDPG) approach to approximate the optimal policy for an agent within a reinforcement learning framework. It begins by initializing the actor and critic networks with random weights and setting up target networks for stability. During each episode, the agent interacts with the environment, selecting actions based on the current policy with added exploration noise. The agent observes rewards and transitions, storing them in a replay buffer for later training. The critic network is updated by minimizing the loss between predicted Q-values and target Q-values calculated using the Bellman equation. The actor policy is updated using the policy gradient, aiming to maximize the expected return. Target networks are periodically updated by softly updating their parameters towards the current networks, enhancing stability during training. This iterative process continues until a terminal state is reached, ultimately yielding an optimal agent policy for the given environment.

4.4.2 Grasping Environment in ROS-Gazebo

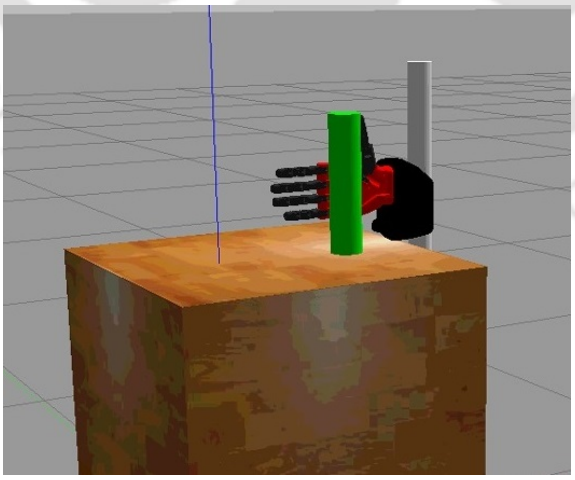
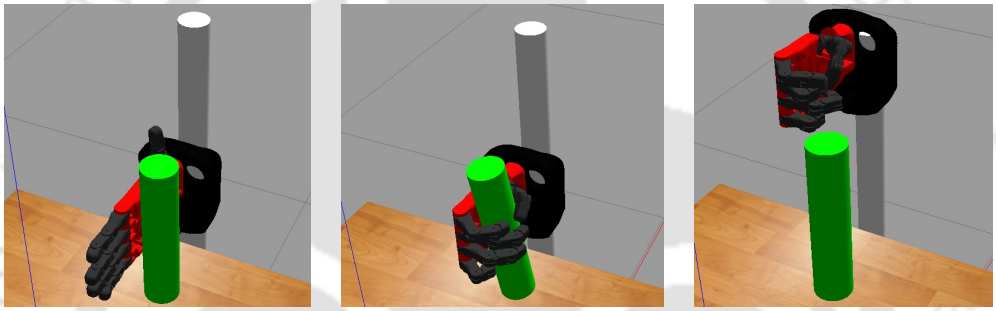


Figure 4.5: Simulation Environment in ROS

The URDF file of the anthropomorphic hand, designed in Chapter 3, is associated

with the cylindrical rail stand in a ROS-Gazebo environment (shown in Fig. 4.5). A table is placed in the grasping world, and a cylindrical object is set down on it. The joint between the anthropomorphic robotic hand's base and the cylindrical rail is prismatic and equipped with effort controllers of the JointPosition controller type from ROS Control to facilitate lifting. The remaining finger joints are revolute joints with JointEffort controllers for effort initiation. The *RL* environment's steps start by loading the grasping world while ROS is launched. The object is initially held for a few seconds (3 seconds) to help it stabilize. The lifting of the object begins at 3 seconds for a short time (7 seconds), and after that, the episode ends. If the hand holds the object for 7 seconds without letting go, the grasp is said to be successful. Dropping will be considered when there are zero contact forces after the lift starts.



(a) (b) (c)

Figure 4.6: Grasp simulation in Gazebo: (a) Initial Grasp Pose, (b) Grasping the object, (c) Slippage while lifting the object

When the ROS world is launched, a ROS node is executed, which is written to initialize the grasp pose (Fig. 4.6(a), grasp the object (Fig. 4.6(b)) and lift it (Fig. 4.6(c)). However, before training, I can see slippage occurs during lifting, and the anthropomorphic robotic hand is unable to lift it successfully (Fig. 4.6(c)).

4.4.3 Design of the Slip Filter

Fig. 4.7 shows the basic pipeline block layout of the standard slip prevention filter described in the literature. The following are the primary steps: 1) a network of filters, 2) rectification of a half-wave, 3) an envelope 4) binary ON-OFF calculation.

In the basic traditional slip filter detection that is discussed in the literature, if any grasped object slips, it causes high-frequency vibrations. The vibrations are

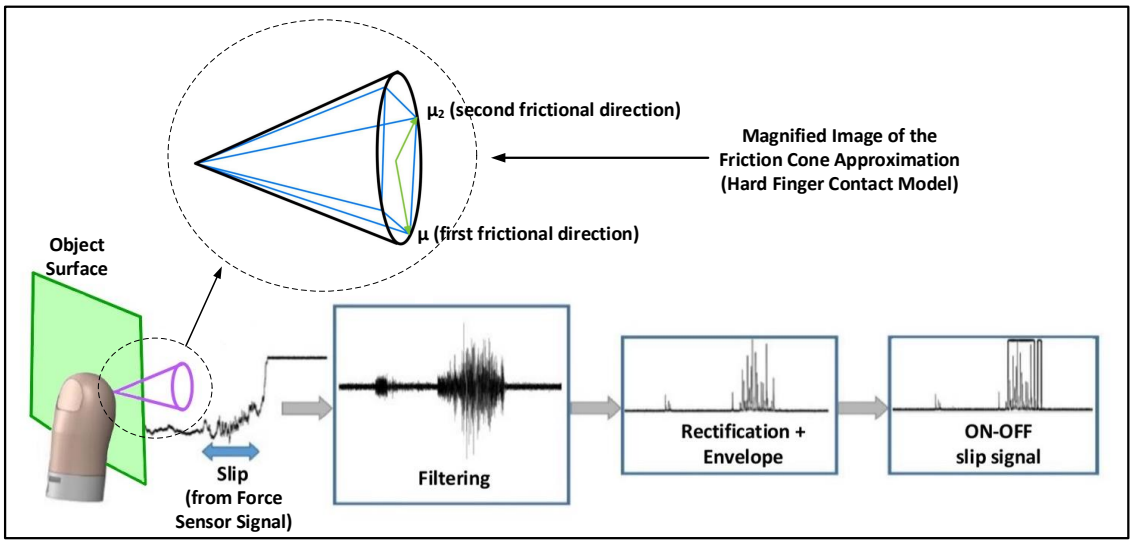


Figure 4.7: Typical slip detector algorithm block diagram. The anthropomorphic hand's fingertip force sensor signal, generated from touching the object's surface, is used to capture the raw voltage signal. The tactile signal is filtered, rectified, and enveloped after that. Finally, a binary ON-OFF signal indicating the existence of slip is produced [35]. The hard finger contact model-based friction cone approximation is depicted.

extracted from the force sensor signal in the first phase. To simultaneously estimate force and identify slippage, the filter network can be used to extract the voltage fluctuations caused by the slippage phenomena. A half-wave rectifier can be used to help obfuscate the negative variations because the filter network is bipolar. The rectified signal is then smoothed by the envelope, which also helps the subsequent thresholding method by reducing too-fast changes. The binary ON-OFF signal is then calculated using the previously mentioned thresholding, which operates on an empirically selected threshold on preset windows of the wrapped signal. The output of the force sensor signal can be transformed in the frequency domain to reveal hidden slippage information. Hence, the above conventional slip detector needs supervision and human input to identify the thresholds [35], which in turn deteriorates the generalization capability for unknown objects.

To get over the drawback of thresholding, I draw my inspiration from the work in [34], which utilized the analysis of the property of the trend of pairwise details in a

discrete wavelet transform (DWT)-based methodology to determine the occurrence of slip. Specifically, two subsequent DWT components had the same absolute value but a different sign due to the characteristics of the DWT that is used. Because the sign of paired components shifts from negative to positive in the load phase, as opposed to the slip phase, hence it is possible to discriminate between the two stages [34]. Wavelet analysis, using the DWT, is a standard method for efficiently examining localized power variations in time series data. In a one-dimensional time series, the force signal can be represented as:

$$f(t; F_N, \alpha) = \sum_{k \in Z} c_{j_0,k} \varphi_{j_0,k}(t) + \sum_{k \in Z} \sum_{j=j_0}^{\infty} d_{j,k} \psi_{j,k}(t) \tag{4.2}$$

where $f(t; F_N, \alpha)$ is influenced by the normal force F_N and the sensor parameter α. The other terms involved include φ (scaling function), ψ (wavelet function), k (shift parameter), j (scale parameter), $\{c_{j_0,k}\}_{k \in Z}$ (approximate coefficient sequence), and $\{d_{j,k}\}_{k \in Z}$ (detail coefficient sequence). $\sum_{k \in Z} c_{j_0,k} \varphi_{j_0,k}(t)$ provides a low resolution approximation of $f(t; F_N, \alpha)$, while $\sum_{k \in Z} \sum_{j=j_0}^{\infty} d_{j,k} \psi_{j,k}(t)$ depicts the details of $f(t; F_N, \alpha)$ in high resolution [85]. Since slip is a transient but rapid change in signal, hence in this work I focus on high-resolution details of the sensor signals. Next, I need to choose an appropriate wavelet function.

Haar Wavelet is one of the common techniques to detect slippage while grasping an object [86]. It is usually chosen to differentiate slippages because slip signals are transient changes in original signals, and the Haar wavelet is itself discontinuous, non-differentiable, and asymmetric. Hence, slip signals are well reflected and captured by the high-frequency components of the Haar wavelet [85]. Its advantages include multi-resolution analysis, the ability to detect localized changes, real-time processing, the ability to threshold, simplicity and speed, directional information, sparse representation, robustness to noise, low memory requirements, etc. In the current work, the Haar wavelet is hence utilized to detect slippage while grasping an object.

After I selected Haar as the mother wavelet, I decomposed the raw force signal at different levels and chose the 5th-level Haar decomposition. The selection of the 5th-level decomposition of the Haar wavelet transformed signal was based on careful consideration of several factors, including the trade-off between decomposition level and signal representation, as well as the specific requirements of my slip detection

task. I found that using a higher decomposition level, such as the 5th level, yielded satisfactory results in terms of extracting relevant features related to slip events while preserving the essential information in the tactile sensor signals. The 5th-level decomposition provided a balance between capturing fine-grained details of the signal and avoiding excessive noise or artifacts introduced by deeper levels of decomposition. Hence, the choice of the 5th-level decomposition was thus influenced by empirical observations and experimentation with various decomposition levels and finally sticking to the mentioned level as shown in Fig. 4.8. Fig. 4.8 shows Haar wavelet decomposition-based slippage detection used in this chapter to detect slip. After decomposing the force signal using the Haar wavelet, I looked at the gradient trend of the inverse Haar using the approximation and detailed coefficients of the 5th-level Haar transformed signal. If the gradient is negative then there is a slip occurring, while vice versa if there is no slip. The RL algorithm uses this trend as slip state observation after detecting it with slip detector logic.

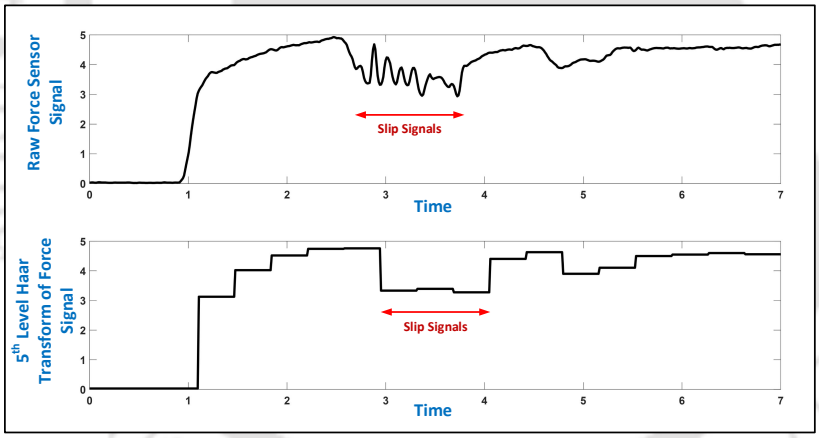


Figure 4.8: Force sensor signal while grasping and lifting an object. Fifth level Haar decomposition of raw force sensor signal is used to detect slip. The positive gradient is a reflection of the load being applied. The opposing variation trend is a representation of slip.

4.4.4 Reinforcement Learning Formulation

The MDP formulation of the slip prevention problem is given below:

1. **States:** Joint Angles, Joint Velocities, Fingertip Forces, Slip States, Wavelet Coefficients, Joint Torques

2. **Actions:** Joint Torques

3. **Rewards:**

$$Total\ \ Reward = 10(sign(\Delta d)) - 20\psi + 100\gamma - d^2$$

where, d is Haar wavelet detailed coefficients, Δ is the gradient,
ψ(boolean for "if there is drop") = $(sign(\Delta d) == 0\ \ \&\&\ \ time \geq 3)$,
γ (boolean for "if grasp is successful") = $(sign(\Delta d) == 1\ \ \&\&\ \ time == 7)$,

d^2 is the wavelet coefficient energy [8] (it is used as a measure to quantify the magnitude of slip)

4. **Policy (learned):** Learn the optimum torques to be applied to the joints based on the maximum accumulated reward function

4.4.5 Randomization of the Physics Parameters

As previously mentioned, DR is an adequate way to induce domain adaptation by introducing parameter randomization. For my current experiment, I randomize the physics parameters in Gazebo. I sample random trajectories for each episode from a set limit of randomized parameters and train my RL model. The parameters are randomized along a set of nominal values and range within limits.

DR **of the Mass**: Firstly, the object's mass is randomized, and the effect is seen. The nominal mass of the grasped object is 1 kg, and an additive variance of 50% to 150% of its nominal mass is considered. A custom gazebo model plugin is written to change the mass using the inertial class reference under the physics namespace of Gazebo.

DR **of the Friction Properties**: Secondly, randomization has been initiated on the surface parameters of contact properties. The friction properties are changed by inducing randomization of the grasped object's coefficient of friction (COF). Since COF is a surface property and not an intrinsic property, i.e., it exists only when two

surfaces are in contact; hence, randomization was induced in the COF to initiate slippage. Since the object gets lifted in the z-direction, the primary friction direction was kept in the z-axis in the object sdf file. The randomization is done along the first friction direction, which has the COF, μ, and along the second friction direction, which has a COF, μ_2. Whenever a contact forms, a contact coordinate frame gets constructed with the third axis as the secondary frictional direction, which has a COF as μ_2, which is the cross product of unit vector along the normal vector and first frictional direction (shown in Fig. 4.7). The COF for both μ and μ_2 are randomized within 1 - 0.04. The lower limit is bounded in that range, considering the lowest COF known, i.e., approximately 0.04 for polytetrafluoroethylene (PTFE). Though there are materials with high COF, the upper limit is kept at 1, keeping in mind that the thesis is more concentrated on slip-induced; hence, lower COF is more of interest.

4.5 Results and Discussions

4.5.1 Reward Plots of the Nominal and *DR* Agents

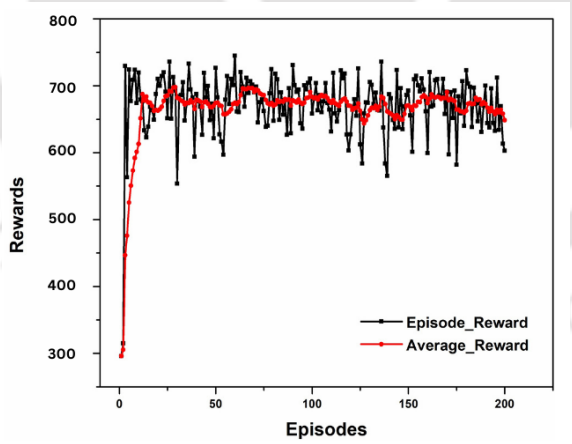


Figure 4.9: Reward plots of the Nominal *RL* agent

The reward plots, i.e., the episode reward and the average rewards, for the *RL* agents trained in different scenarios are plotted. Fig. 4.9 is the reward plot of the *RL* agent trained in the nominal environment, and Fig. 4.10 is the reward plot of the *RL* agent where both the mass and friction coefficients are randomized. The

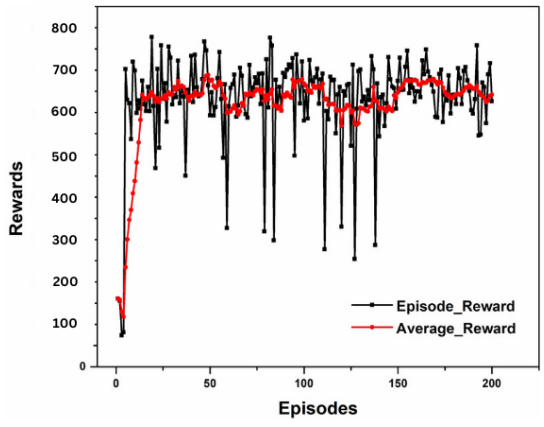
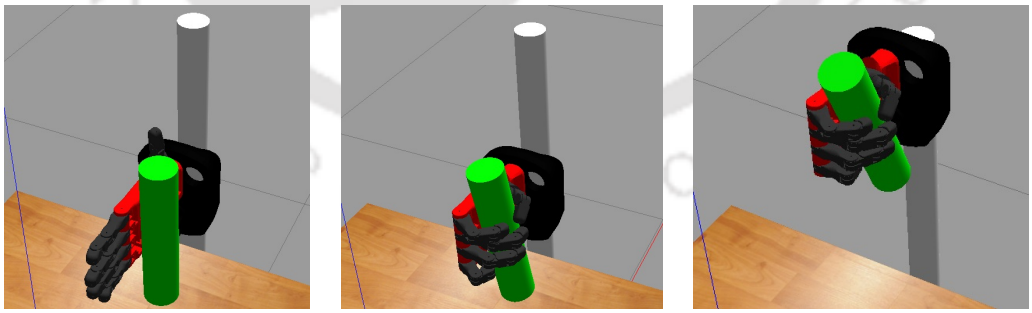


Figure 4.10: Reward plots of the *DR RL* agent (both mass and friction randomized)

convergence times of the reward plots are nearly similar. However, the rewards of the *DR RL* agent are randomized in the early stage than the agents trained in the nominal environment, as there are disturbances due to the randomized parameters, i.e., both weight and friction, which creates more slippages. However, the *DR RL* agent performs better at untrained and unseen object tasks (discussed in the next section). The trained agent's grasp simulation is shown in Fig. 4.11, and it is seen that the agent has been able to lift the object successfully without any slippage (Fig. 5.4(c)).



(a) (b) (c)

Figure 4.11: Learned Grasp simulation: (a) Initial Grasp Pose, (b) Grasping the object (c) Object Lift without Slippage

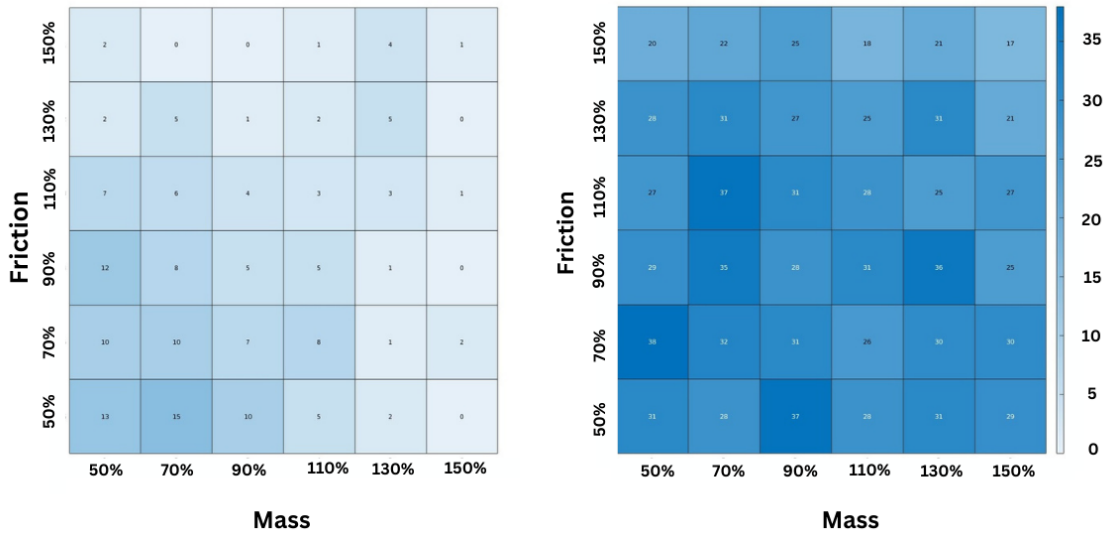


Figure 4.12: Heat Map of the success rates for the Nominal agent and the DR agent

4.5.2 Success Rates for Nominal and DR Agents

Fig. 4.12 shows the heat map of the success rates of grasping unknown objects for nominal and DR agents. The success rate is measured by the number of times the RL agent has been able to prevent slip. Each plot is a heat map demonstrating each agent's performance over different environment dynamics. Every grid point corresponds to a particular set of environment dynamics, and darker colors correspond to higher scores. The X-axis shows the percentage of randomization of the mass for the testing conditions from the nominal value. The Y-axis represents the percentage of randomization of friction coefficients for the test environment from the nominal value. From the success rates plot, it is evident that the agent trained in DR environment has been able to prevent more slips than the nominal agent, concluding that the learned DR agent is well equipped to better generalize to varying environmental conditions.

Also, as a measure to show that the amount of slip is minimized, I look into the sum of the wavelet coefficients in a trial and evaluate the amplitude of the wavelet coefficient energy as shown in Fig. 4.13. The results from the ten experimental trials show that the amplitude of wavelet coefficient energy is greater in the case of nominal agent as compared to DR agent, signifying higher transient disturbances and higher slippage occurrences. This validates the idea that the magnitude of slip decreases if trained on DR environment.

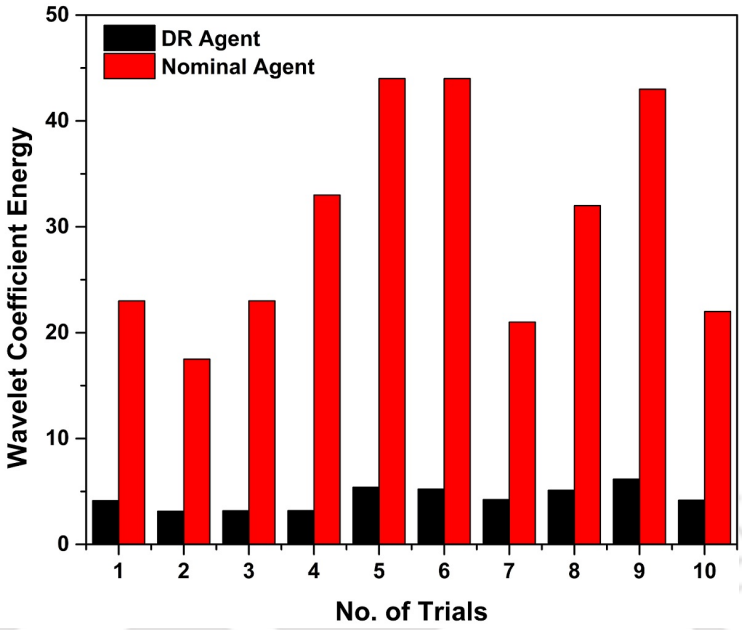


Figure 4.13: Wavelet coefficient energy across different iterations

4.5.3 Validation based on Grasp Quality Metric

Analytical measurement of the grasp quality, called largest minimum resisted wrench, is used to validate the bionic reflex controller. Together with the localization of the contact points, the feature "largest minimum resisted wrench", Q_{LMRW} adds further consideration about the amplitude of external forces/disturbances that the grasp can handle and reject. This feature resembles the force closure condition in that it is based on the concept of a convex envelope in the space of forces and wrenches that can be used to manipulate the grasped item [29].

To analyze the grasp quality, a friction cone can be considered at the contact point p_i by a pyramid with m edges, and the force f_i applied by the finger can be expressed as a positive linear combination of unitary forces $f_{ij}, j = 1, ..., m$ along the pyramid edges (often referred to as primitive forces), and the wrench produced by f_i at p_i can be expressed as a positive linear combination of the wrenches ω_{ij} produced by f_{ij}. On the item provided by: n fingers, a resulting wrench ω_0 is given by [29]:

$$\omega_0 = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_{ij} \omega_{ij} \tag{4.3}$$

where, $\alpha_{ij} \geq 0, \quad \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_{ij} \leq 1$

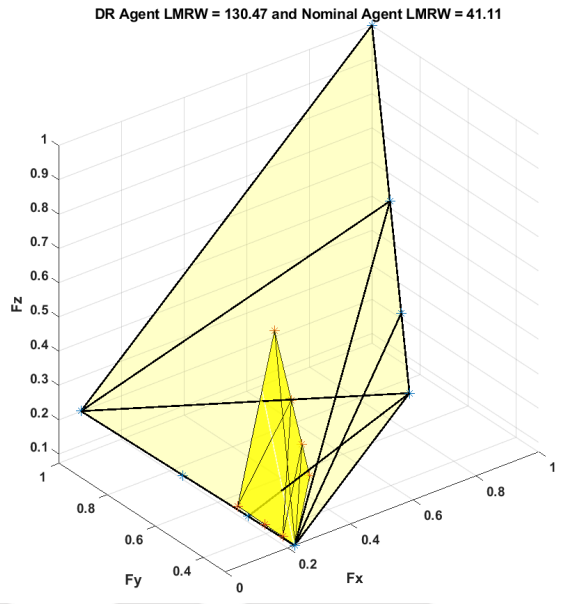


Figure 4.14: Grasp Wrench Space analysis of the Nominal Agent (dark yellow) and the DR agent (transparent yellow). The convex hull of the nominal agent is completely inscribed within the convex hull of the DR agent, indicating better external force/disturbance handling capacity of the DR agent.

The set P of potential consequent wrenches on the object is the convex hull of the primitive wrenches ω_{ij} when taking into account the possible changes of α_{ij}:

$$P = CH\left(\bigcup_{i=1}^{n} \{\omega_{i1}, ..., \omega_{im}\}\right) \tag{4.4}$$

The shortest distance between the origin and the nearest facet inside the envelope P can be used to establish the largest disturbance that a grasp can withstand in all directions. This geometrically equates to the diameter of the greatest sphere that is completely contained inside the convex hull, P, and is completely centered at the origin of the wrench space. To assess how much a grasp will withstand external force and wrench in all directions, I shall continue to use a grasp quality metric called the distance Q_{LMRW} :

$$Q_{LMRW} = \min_{\omega \in \delta P} \|\omega\| \tag{4.5}$$

where δP is the boundary of P.

The results of the analysis show that the largest minimum resisted wrench during lift just before slippage for the nominal agent is 41.11 and 130.47 for the DR agent. The convex hull of the grasp wrench space of the nominal agent is fully inscribed inside the convex hull of the grasp wrench space of DR agent (shown in Fig. 4.14). The grasp wrench space analysis shows that the DR agent has better disturbance (external force) handling capacity than that of the nominal agent.

4.6 Chapter Summary

The chapter discusses the design of a RL-based slip-prevention controller developed to enhance the functionality of an anthropomorphic hand. Experimental results demonstrate the viability of the RL controller in preventing slip by learning through interaction without prior knowledge or human supervision. Also, the RL agent is trained under DR to improve the efficiency of policy generalization capability of varying testing scenarios. Though the framework has been used to test the slippage control algorithm for an anthropomorphic hand, the approach can be applied to develop and test the slippage control algorithms with any robotic hand. The work undertaken in this chapter may provide a pathway for designing further robust autonomous slippage controllers without human intervention.

Chapter 5

Deformation Prevention with Reinforcement Learning Based Control

5.1 Introduction

While dealing with the challenges of slippage of the grasped object due to external disturbances, sub-optimal grasp forces may lead to object deformation leading to significant obstacles affecting the efficiency and safety of robotic grasping applications [18]. Addressing object deformation together with slippage is paramount to the evolution of robotic grasping. Tackling both together completes the emulation of human grasp reflex control, i.e. bionic reflex. Providing the sense of deformation is one of the distinctive features of human tactile perception for grasping and manipulation. Having a sense of the amount of deformation aids the controller in preventing undesirable grasping force overshoot meanwhile preventing deformation or damage. When there's a perception of slippage, the human brain automatically adjusts the hand to increase its grip strength, ensuring there's enough safety margin to prevent further slippage. However, this reflexive action is absent in prosthetic hands, leading to insufficient grasping force and potential object slippage. Additionally, unlike the human brain's ability to quickly adapt grip strength based on factors like object mass, stiffness, and friction, prosthetic hands lack the capabil-

ity to gather such essential information. Consequently, deformable objects may be damaged if the grip force is too high, while heavy objects may slip if the force isn't adequate. To address these limitations and create a more adaptable prosthetic hand, it's imperative to implement a controller capable of dynamically adjusting feedback gain based on detected slippage, while also regulating the desired grip force based on an estimation of object deformation to prevent crushing the object [5].

Although integrating a slippage sensor can enhance the grasping capabilities of a prosthetic hand by detecting slippage, including a deformation controller to create optimal grasping forces together with the slippage controller (discussed in Chapter 4), would increase the grasping efficiency preventing object damage. Without the information about the object's properties which are being grasped, there's a risk of unintended crushing or dropping of the objects, which can be frustrating. Indeed, achieving an optimal grasping force is paramount for prosthetic hands, and extensive research has been dedicated to this aspect. Therefore, in this chapter, I include a deformation prevention control paradigm along with slippage prevention. The main contribution of this chapter includes (1) a novel RL-based control pipeline for deformation prevention of grasped objects, (2) eliminating the need for human supervision in setting deformation thresholds or knowing object properties beforehand during control design, and (3) leveraging DR to enhance adaptability to object properties and environmental shifts, improving the generalization capability to unknown scenarios. The rest of the chapter is as follows: Section. 5.2 provides an overview of the background and related work to deformation prevention and deformable object handling literature. Section. 5.3 provides an overview of problem formulation. In section. 5.4, I formulate the design of the proposed deformation prevention controller and RL grasping policy. In section. 5.5, I discuss the results obtained from my experimental trials. In the end, in section 5.6, I provide the summary of the current chapter.

5.2 Background and Related Work

5.2.1 Deformation Detection

Deformation prevention necessitates stiffness detection or deformation measurement. Control design of this bionic reflex feature is quite challenging and is still an open

research problem [18]. Existing methodologies for stiffness detection and control in the literature include: (1) intrinsic vibration frequency-based signal processing, (2) time-domain analysis methods, (3) the integration of measuring devices [38] (4) Hooke's Law [39]. The first approach involves gathering the vibrational response elicited by excitation and then determining the stiffness of an object through frequency domain decomposition. This particular signal processing technique is known for its exceptional precision in measurement; however, this method is typically for offline analysis. A second approach involves, applicability for real-time usage, and it involves monitoring various parameters such as equivalent force, deflection, and velocity during contact. By analyzing these characteristics in the time domain, the stiffness of an object can be deduced. Continuous excitation is necessary for precise stiffness estimation, a requirement that can be enhanced to achieve real-time grasping motions, particularly beneficial for prosthetic hand applications. The third approach is a more intricate and costly approach that involves integrating a specialized measuring apparatus comprising both mechanical and electrical components at the terminus of the robot gripper. This method demands establishing a correlation between the stiffness of various sample materials and their corresponding gripping forces post-contact. Nevertheless, incorporating an additional device at the fingertips is unsuitable for prosthetic hands due to stringent size and weight constraints. Additionally, it's essential to note that this method solely discerns the contact stiffness between the gripper and the grasped object, a distinction from the inherent stiffness of the object itself. To detect stiffness using Hooke's law, I first calculate the forces and deformation then calculate the ratio of those to find the stiffness coefficient, ($K = F/d$), where F represents contact force due to grasping, d represents the deformation of the object. It should however be noted that formulating the deformation prevention control paradigm based on instantaneous calculation of Hooke's law (instantaneous deformation calculation) is not easy to use in underactuated prosthetic hands [38].

Some studies employ vision-based techniques for deformation detection [18, 87, 88]. Tactile sensors like the Gelsight sensor offer high-resolution tactile data by measuring elastomer deformation [18]. However, such sensors can be costly and less accessible. [88] relies on image processing techniques like pixel analysis to generate a stiffness map using the 'pix2stiffness' method. This map guides grasp pose detection for preventing damage to unknown deformable objects. However, stiffness map generation is not automated, and no force control based on contact dynamics to

minimize deformation was considered. There are also end-to-end frameworks that predict the geometry and dynamics of the objects for volumetric deformable object manipulation [89]. However, it requires training and labeling, introducing human intervention in the design stage.

5.2.2 Deformation Control

Impedance control is instrumental in preventing the deformation of grasped objects within robotic manipulation systems [90–92]. By leveraging real-time force sensing capabilities and adaptive control algorithms, robots can dynamically adjust their compliance and stiffness properties to match the characteristics of the object being handled. This enables the robot to exert the necessary force for a secure grip while minimizing the risk of deformation, whether the object is soft and compliant or rigid and brittle. Through a feedback control loop, impedance control ensures that the robot can react swiftly to changes in the environment, executing smooth trajectories and optimizing its interaction forces to prevent damage. Overall, impedance control enhances the precision and reliability of robotic manipulation, making it well-suited for applications where delicate object handling is essential. In [5], stiffness was identified and controlled using a PVDF sensor, but it relied on human supervision to establish voltage thresholds for categorizing deformable objects. Slippage detection was based on the empirical mode decomposition (EMD) of the force sensor signal. Deformation control utilizing Hooke's law and impedance-based control requires exact stiffness knowledge and desired model references, respectively, for their effective implementation.

Other non-conventional approaches addressing deformation control include utilizing kinematics for stiffness detection in underactuated mechanisms [38]. This approach relies on knowledge of the specific hand structure and does not consider factors like friction and weight. In [93] slippage was determined based on the tangential force-to-normal force ratio detected by tactile sensors and deformation was controlled by altering object weight through reorientation, requiring deformation and force thresholding. In [94], a bistable compliant underactuated gripper is designed to enhance the grasp of deformable objects. Conventional hard grippers struggle with grasping deformable and unpredictable items due to equilibrium constraints and computational complexity. In contrast, soft grippers leverage material compliance and flexibility, offering superior adaptability to objects of varying shapes,

including deformable ones [95]. However, existing soft robotic end-effectors face limitations such as limited variability, adaptability, miniaturization potential, and high complexity, weight, sensory, and control requirements, leading to higher costs [96]. The inherent rigid configurations of typical prosthetic and robotic hands pose challenges in effectively grasping deformable objects, primarily because of their inability to dynamically adjust stiffness. The preferred approach to address this limitation and achieve adaptable grasping akin to human hands is through the implementation of an active control system, which constitutes the primary objective of this chapter.

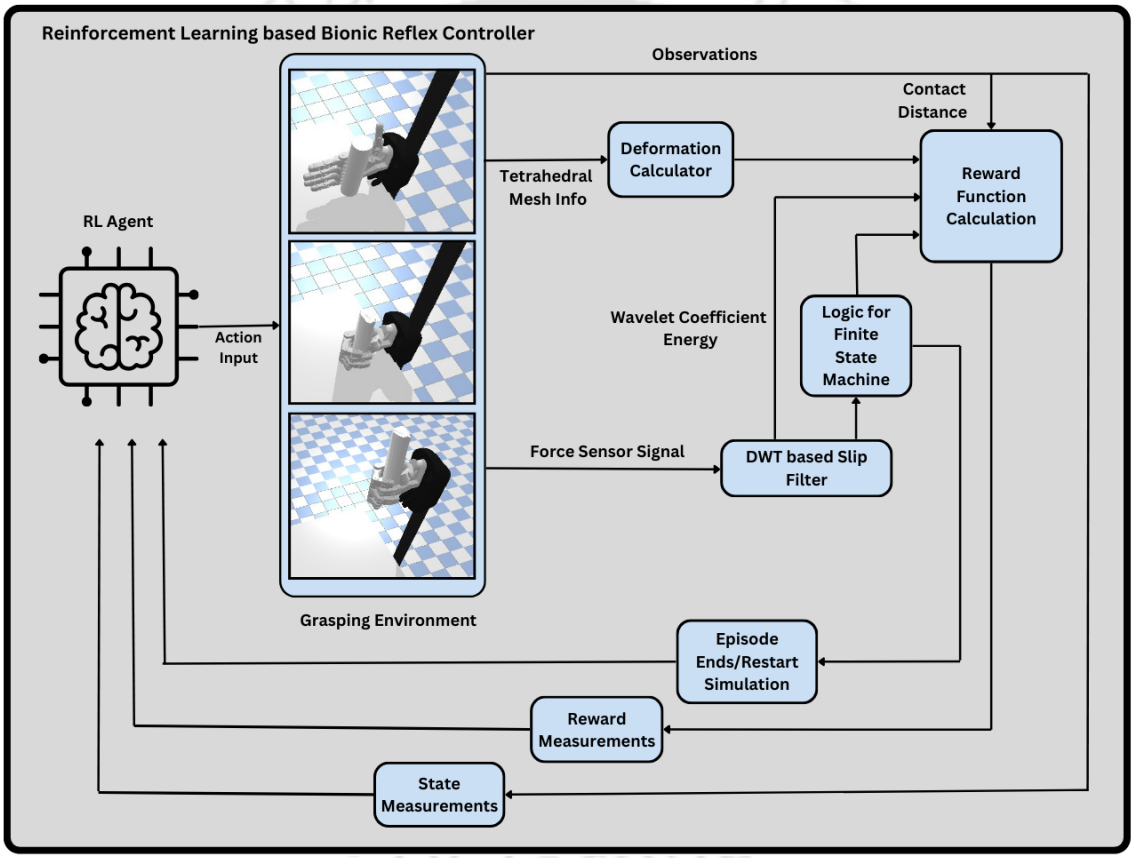


Figure 5.1: RL-based Bionic Reflex Control Pipeline

5.3 Problem Formulation

The goal of this chapter is to design an intelligent bionic reflex controller to minimize deformation while preventing incipient slips during grasping and lifting. As

mentioned in 4.1, if I can replicate the spinal reflex loop, then I would be able to achieve a bionic reflex controller. In this chapter, on top of slip prevention control I include deformation prevention control also in the control pipeline. I aim to achieve it by designing an RL-based slip prevention and deformation minimization controller. The problem of grasping and lifting the object is posed as an MDP. The detailed formulation of the grasping MDP is given in the next section. Slippage occurring during the lifting of the grasped object has been detected using DWT based slip filter as discussed in the previous chapter, and a deformation filter detects the deformation (explained in Section. 5.4.4). The advantage of the above methodology for slip discrimination and deformation detection lies in the ability to exclude thresholding of signals or knowing the object properties apriori, thus minimizing human intervention in the control design stage. Compared to the previous chapter, here, I consider a deformable body for conducting my experiments. The RL-based slippage and deformation prevention controller schematic is shown in Fig. 5.1. The reward function calculator receives the contact distance info, wavelet coefficient energy, deformation info, and slip state info from the finite state machine. The details of the reward function calculation are given in Section. 5.4.5. The observation state measurements, the reward measurements, and the boolean of the state machine informing episode end/restart simulation are sent to the RL agent which calculates the optimal action input (policy). The optimized policy will produce the most optimal torques, generating robust fingertip forces for slip prevention and minimizing deformation.

5.4 Design Methodology

5.4.1 Deformation Prevention based Bionic Reflex Grasping Policy

The deformation problem is formed as an MDP, and the entire grasping problem is formed as a $MFRL$ to make it learn policies by directly interacting with the environment; mapping from states to actions. The deformation of the grasping state is a continuous state representation. The other observation states include joint angles, joint velocities, fingertip forces, slip states, wavelet coefficients, and joint torques, which are continuous too. Moreover, the action space are the continuous

joint torques. Hence, I choose actor-critic as the RL algorithm for my problem as they are suitable for continuous state and action spaces. The pseudo-code of the algorithm of the RL-based bionic reflex controller is shown in Algorithm 1. This algorithm is designed to determine the optimal joint torques τ required for lifting a grasped object without experiencing slippage or deformation within a PyBullet [97], grasping environment (explained in the next section). It operates through a series of episodes, each consisting of steps aimed at achieving successful grasping and lifting actions. Initially, the algorithm initializes the observation of states (s) from the environment. During each step of the episode, actions are generated using the Soft Actor-Critic algorithm to manipulate the robot's joints and grasp the object. Following the grasp, the algorithm lifts the object while continuously monitoring for signs of slippage or deformation. If the object drops or slips, the episode ends, and the simulation is reset. If slippage or deformation occurs during lifting, the joint torques (τ) are adjusted accordingly to either increase or decrease the force applied. First, the algo checks for slip, and if there is one, then it corrects the joint torque using $\delta\tau$. After this action, if there is deformation occurring then, the joint torques are corrected by a value λ. If the grasp is successful without slippage or deformation, the episode ends, and the simulation is reset. This process iterates until the terminal time of the episode is reached, ensuring comprehensive exploration and refinement of the grasping strategy over multiple episodes.

5.4.2 Grasping Environment in PyBullet

The current study aims to train an anthropomorphic robotic hand to grasp an object and lift it while minimizing slippage and deformation, leveraging RL in PyBullet [98] (Fig. 5.1). The custom unified robot description format (URDF) model of the anthropomorphic hand from [99] is used. The RL agent sends joint torques as actions to the robotic hand grasping environment which helps it to grasp and lift the object. The slip filter and the deformation checker (explained in detail in the next section) are utilized to calculate the rewards and generate optimal actions. Torque control is chosen within the PyBullet simulation since it allows precise control over the joint motor by adjusting the voltage induced; a preferred mode for replicating real-world bionic device behavior. A custom OpenAI Gym environment is then developed, leveraging PyBullet's control functions within the class functions. In line with the standard structure of a Gym environment, the defining class consists of five essential

functions: initializer, step, reset, render, and close.

Algorithm 1: Bionic Reflex Control

Input: Grasping environment in PyBullet

Output: Optimal joint torques (τ) to lift the grasped object without slippage and deformation

begin
- **foreach** *episode* **do**
  - Initialize S (observations of states from the grasping environment)
  - **foreach** *step of episode* **do**
    - Generate actions based on **Soft Actor-Critic** algorithm
    - Grasp the object
    - Lift the object
    - Slip discrimination during object lifting
    - **if** *Object Drops* **then**
      - episode ends
      - reset simulation
    - **else**
      - **if** *Slip Occurs* **then**
        - $\tau \leftarrow \tau + \Delta\tau$
        - **if** *Deformation Occurs* **then**
          - $\tau \leftarrow \tau + \Delta\tau - \lambda$; `// (-λ is a small decrement in force)`
        - **else**
          - $\tau \leftarrow \tau + \Delta\tau + \lambda$
        - **end if**
        - continue
      - **end if**
    - **end if**
    - End lifting motion (Grasp Successful)
    - **if** *Grasp Successful* **then**
      - episode ends
      - reset simulation
    - **end if**
  - **end foreach**
  - Until Episode Time is terminal
- **end foreach**

end

In the initializer function, I initiate the PyBullet connection via 'pybullet.connect,' set gravity to $-9.8m/s^2$, and create observation and action spaces. My Gym environment's observation space includes joint positions, velocities, finger contact forces, slip and deformation states. The action space consists of torque values. The 'step' function executes commands based on the simulation's running time. When the

environment is active for less than 5 seconds, the RL agent executes the action of grasping. Beyond 5 seconds, it combines grasping and lifting, capping each episode at a maximum of 10 seconds. The 'step' function primarily executes actions in PyBullet, retrieves current rewards, and returns the environment's state. This setup ensures smooth interaction between the RL agent and the simulation, enabling the agent to adapt its actions based on the environment state and enhancing grasping and lifting capabilities. The render and close functions are straightforward as the former sets the PyBullet connect mode to GUI, and the latter simply disconnects it. Once the environment class is defined, I can save it as a Gym environment and call it while training and testing my RL agent.

5.4.3 Deformable Object

The deformable object is first generated in SolidWorks as an .stl (Standard Triangle Language/Standard Tessellation Language) file. Then the tetrahedral meshing was done in fTetWild and .msh file was generated [100]. The deformable object is generated by Gmsh as a .vtk file and simulated by the built-in Finite Element Method (FEM) in PyBullet. The FEM serves as a powerful tool to model deformable objects. Through discretizing the objects into small elements, the deformation can be derived by solving a set of partial differential equations. FEM can accurately represent the dynamics of deformable objects with fine tessellation. However, though it is computationally expensive, it efficiently approximates the true physical behavior of a deformable object [101]. The deformable object created is shown in Fig. 5.2. The parameters of the deformable object taken for my physics-based simulation are shown in Table. 5.1.

Table 5.1: Simulation Parameters for Deformable Object

Parameters	**Values**
Mass	1 Kg
Friction Coefficient	0.5
NeoHookeanMu (Shear Modulus)	33. 78 MPa
NeoHookean Lambda (First Lame Parameter)	810 MPa
NeoHookean Damping	10

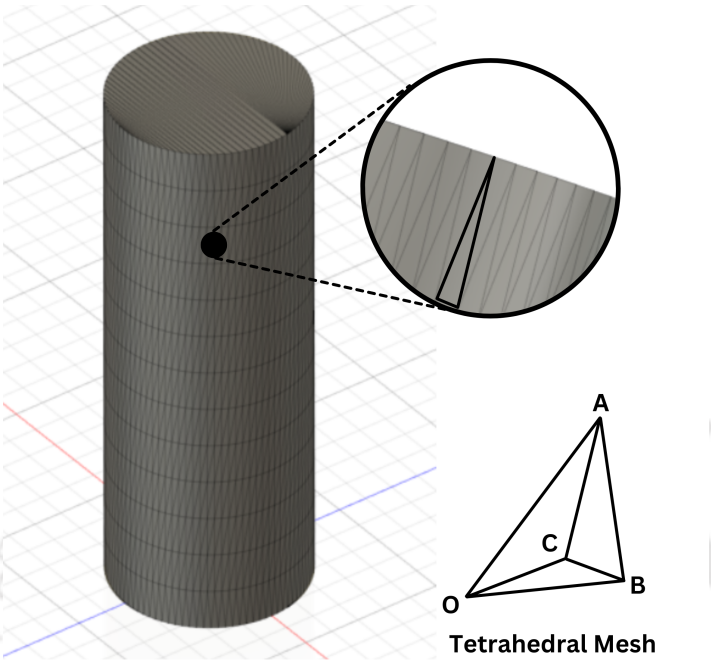


Figure 5.2: Triangular mesh description of the deformable cylinder. Point **O** represents reference origin. The tetrahedral mesh, ABCO is shown in the figure as well.

5.4.4 Deformation Calculation

PyBullet represents surface mesh data using vertex and triangle lists (as shown in Fig.5.2). The PyBullet function "getMeshData" returns the mesh information (vertice indices) of the triangular meshes that form the 3D object. The reference frame is also changed from the ground to the grasped object (shifted origin) dynamically during grasping and lifting. This step enables us to account for any translations or rotations experienced by the object during manipulation. To estimate volume changes for deformable objects without direct mathematical formulas, I define a general point (shifted origin) to create a tetrahedron (Fig. 5.2) and calculate the signed volume using the following formula (as given in [102]) :

$$Signed\ Volume = \mathbf{AO}.(\mathbf{AB} \times \mathbf{AC})/6 \tag{5.1}$$

where points A, B, and C are concurrently selected from the mesh, and O is a reference mesh point taken arbitrarily. The "signed volume" indicates the orientation or direction in which the volume is calculated; here, the surface normal of **ABC** determines the sign and weight of each tetrahedron, and the summation of all such

solids captures the object's overall volume. This volume is then used to calculate deformation (change in volume) at every time step. My approach provides a dynamic and comprehensive assessment of deformation and allows real-time monitoring for reward calculation.

5.4.5 Reward Formulation

To achieve my goal of a secure grasp with minimal object deformation, I define the rewards as follows:

$$\sum_{i=1}^{5}\left(\frac{1}{\ln(x_i+1.1051)}\right)+\sum_{i=1}^{5}(\delta_i.10)-\sum_{i=1}^{5}(\theta_i.10)-C\times\Delta-d^2 \tag{5.2}$$

The goal here is to slowly but accurately make the hand grasp the object so that it does not slip during lifting while also ensuring that this grasp does not damage the object through deformation. The first term uses an inverse logarithmic relation to define the reward for each step. x_i is the distance between the fingertip and the object; as the finger approaches the object, x_i approaches zero, and the reward term is maximized. The following term further signifies firm contact. The distance term can give similar rewards even when x_i is close to 0. To ensure that my agent can differentiate between close contact and contact, this discrete term, δ further increments a +10 reward on contact of each finger. δ is 1 if there is contact and 0 otherwise. If all goes well, by now, my agent should be able to entrap the object between the bionic fingers and ensure contact with the object. My next step involves lifting this object with minimal slipping. Keeping the scaling similar to other parts of the reward, a value of 10 is penalized from the total reward for each finger whenever slipping is detected. The slip value is formulated as a boolean (θ), taking values 1 if slip is detected and 0 otherwise. Although this seems accurate enough, a drawback of my previous three terms is the promotion of firm grasp, i.e., the grasp is tight enough to damage the object. To counter this, a penalty for deformation is also added. To calculate this, I quantify the deformation value from the volume gradient(Δ), that is, change in volume, and scale it accordingly ($C = 50/(initial\ volume)$). This value is subtracted from my reward function directly, penalizing the agent whenever the grasp causes excessive deformation. The final term is the wavelet coefficient energy term [8], which is used as a measure to quantify the amount of slippage.

5.4.6 Randomization of the Physics Parameters

As previously mentioned, *DR* is an adequate way to induce domain adaptation by introducing parameter randomization. For my current experiment, I randomize the physics parameters in PyBullet. I sample random trajectories for each episode from a set limit of randomized parameters and train my RL model. The parameters are randomized along a set of nominal values and range within limits.

DR **of the Mass**: Firstly, the object's mass is randomized, and the effect is seen. The nominal mass of the grasped object is 1 kg, and an additive variance of 50% to 150% of its nominal mass is considered. The mass properties are changed using the mass flag using ChangeDynamics option in PyBullet.

DR **of the Friction Properties**: Secondly, randomization has been initiated on the surface parameters of contact properties. The friction properties are changed by inducing randomization of the grasped object's coefficient of friction (COF). Since COF is a surface property and not an intrinsic property, i.e., it exists only when two surfaces are in contact; hence, randomization was induced in the COF to initiate slippage. Since I am considering the material as natural rubber, so I consider the nominal contact friction coefficient as 0.5, as it is the usual value during contact between rubber and plastic material [103]. The friction parameters are randomized with an additive variance of 50% to 50%. I changed the friction coefficient using the "lateralFriction" flag using changeDynamics function in PyBullet.

DR **of Stiffness Properties**: Finally, to randomize the deformation of the object, and help the agent get exposed to varying stiffness conditions, I utilize the "contactStiffness" flag changeDynamics function in PyBullet. The value of the contact stiffness is randomized within the values of 20 N/cm to 500 N/cm [103].

5.5 Results and Discussions

5.5.1 Reward Plots of the Nominal and Domain Randomized Agents

The reward plots, i.e., the average reward for the *RL* agents trained in different scenarios (one nominal and other *DR*), are plotted (Fig. 5.3). The black-colored graph is the reward plot of the *RL* agent trained in a nominal environment, and the

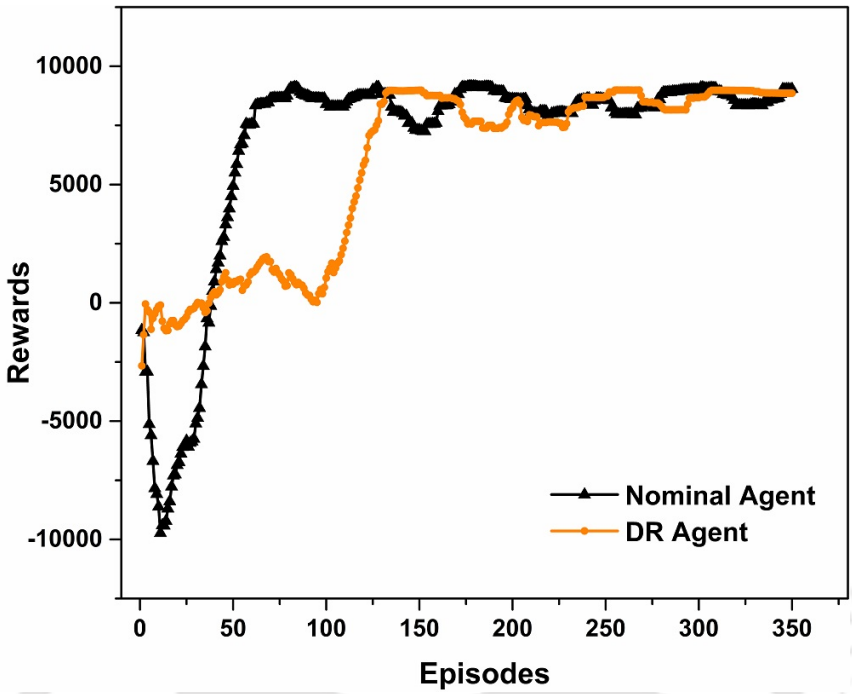
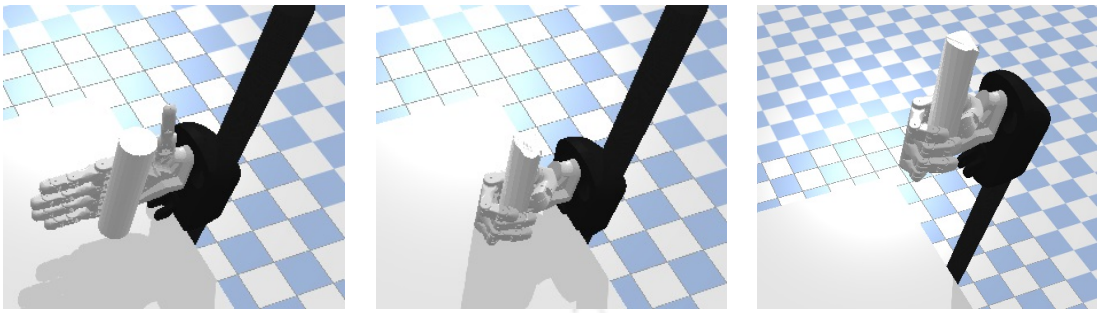


Figure 5.3: Average Reward Plots of Nominal and DR Agent

orange-colored graph is the reward plot of the *RL* agent where the mass, friction coefficient, and object stiffness are randomized. The convergence times of the reward plots are nearly similar. However, the rewards of the DR agent reach convergence a bit later than the agents trained in a nominal environment, as there are disturbances due to the randomized parameters, i.e., weight, friction, and stiffness, which creates more slippages and deformations; hence the cumulative reward falls at the beginning. However, the domain-randomized *RL* agent performs better at untrained and unseen object tasks (discussed in the next section). The trained agent's grasp simulation is shown in Fig. 5.4, and I can see that the agent has been able to grasp and lift the object successfully without any slippage with minimum deformation (Fig. 5.4(c)).

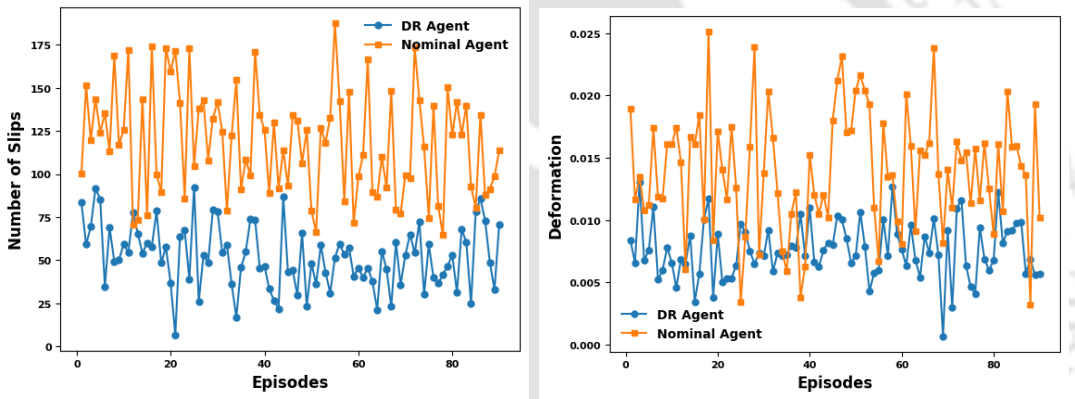
5.5.2 Success Rates for Nominal and Domain Randomized Agents

Fig. 5.5 shows performance tests of both the agents trained on the nominal environment as well as on the *DR* environment on the unknown object grasping task. The unknown objects are randomized to have different weights, stiffness, and friction



(a) (b) (c)

Figure 5.4: Learned Grasp Simulation: (a) Initial Grasp Pose, (b) Grasping the Object (c) Object Lift without Slippage



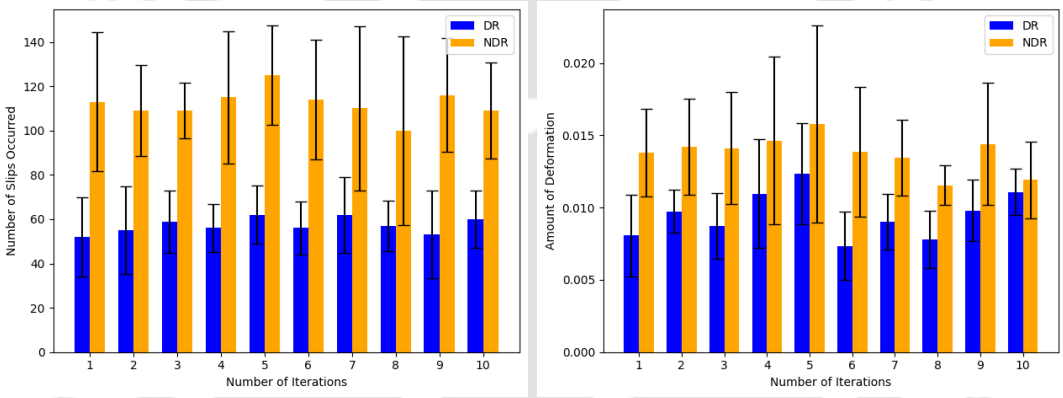
(a) (b)

Figure 5.5: Performance Tests of success rates on Unknown Objects. Parameters randomized were the object weights, stiffness, and friction coefficients while grasping: (a) Slips prevented by Nominal and DR Agent at unseen objects task. (b) Amount of deformation (in mm) prevented by Nominal and DR Agent on unknown objects task

properties. Fig. 5.5(a) shows the success rates of preventing slippage while grasping unknown objects for nominal and *DR* agents. The x-axis shows the number of episodes for which the learned agent was tested. The y-axis represents the frequency of slips occurring. From the success rates plot, it is evident that the agent trained in a *DR* environment has been able to prevent more slips than the nominal agent. Fig. 5.5(b) shows the deformation occurring on the deformable object when grasped by the nominal and the *DR*-trained agent on unknown objects tasks with different

object properties. The x-axis shows the number of episodes for which the learned agent was tested. The y-axis represents the amount of deformation undergone while the object is being grasped. It is seen that while grasping objects with unknown properties, the *DR* agent is able to grasp the object with lesser deformation than the nominal agent, together with preventing slip and droppage.

To validate the efficacy of the results obtained, I repeat the above experiments over 10 trials. I then present the error bar plots for each iteration to provide the visual representation of variability and illustrate the statistical significance of the experimental outcomes (shown in Fig. 5.6). From both performance tests and statistical variability tests, I can conclude that the learned *DR* agent is better equipped to improve the generalization capability than the nominal agent.



(a) (b)

Figure 5.6: Error bar plots for (a) Slips prevented by Nominal and DR Agent at unseen objects task. (b) Amount of deformation (in mm) prevented by Nominal and DR Agent on unknown objects task. Here DR represents domain randomized agent and NDR represents non-domain randomized agent, i.e., nominal agent.

Also, as a measure to show that the amount of slip is decreasing, I look into the sum of the wavelet coefficients in a trial and evaluate the amplitude of the wavelet coefficient energy as shown in Fig. 5.7. The results of the error bar plots from the ten experimental trials show that the amplitude of wavelet coefficient energy is more in the case of nominal agent as compared to *DR* agent, signifying higher transient disturbances and higher slippage occurrences. This validates the idea that the magnitude of slip decreases if trained on *DR* environment.

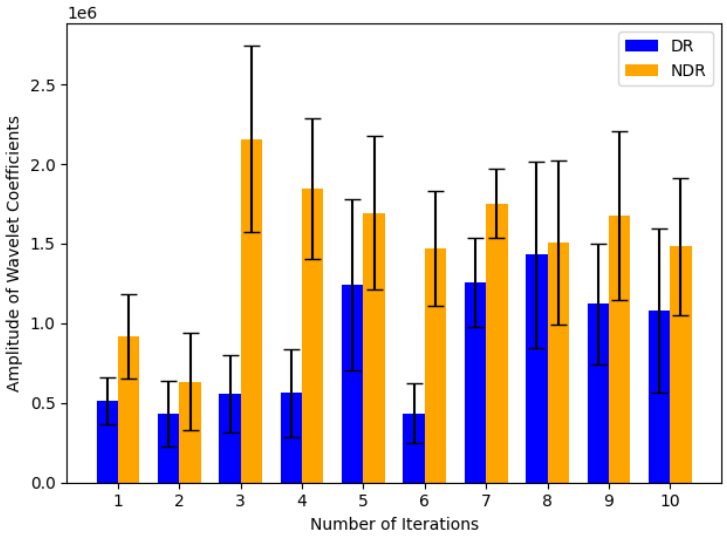


Figure 5.7: Amount of slip based on the amplitude of wavelet coefficients. Here DR represents domain randomized agent and NDR represents non-domain randomized agent, i.e., nominal agent.

5.5.3 Validation based on Grasp Quality Metric

Selecting an appropriate grasp quality metric is crucial when dealing with deformable objects, as traditional metrics developed for rigid bodies often fail to capture the continuously evolving contact dynamics. In this work, the linear instability metric derived from the shake task was chosen based on its robustness, computational feasibility, and direct relevance to the challenges of deformable grasping.

Unlike classical metrics such as the Ferrari-Canny epsilon metric or the volume quality metric, which assume static, point-contact interactions; the linear instability metric quantifies the grasp's ability to resist disturbances by averaging the loss-of-contact accelerations as the grasp is subjected to gradually increasing accelerations. This dynamic approach inherently accounts for the deformation-induced changes in contact configuration, which is a critical aspect when the object under manipulation is non-rigid [104].

Moreover, while full dynamic simulation of the grasping process can be computationally intensive, the shake task provides a practical compromise by focusing on the critical point at which the grasp fails. This results in a metric that is both sensitive to variations in grasping force and computationally more efficient, mak-

ing it well-suited for iterative grasp evaluation in both simulated and real-world scenarios [105].

In summary, the linear instability metric was selected for its ability to robustly assess the grasp quality of deformable objects by directly measuring the resistance to external disturbances, while maintaining computational efficiency. This makes it a suitable and validated choice for the control strategies developed in this thesis, particularly in applications where both slip prevention and deformation minimization are of paramount importance.

In the context of the shake task, I subjected the grasped object to linear acceleration until it dropped, at which point I computed the linear instability metric by averaging the loss-of-contact accelerations. A higher value of this metric indicates a better grasp, signifying the ability to endure more significant accelerations and maintain efficient object manipulation during abrupt jerks and external disturbances. In my experiments, I incrementally increased the acceleration applied to the grasp while elevating the object until the anthropomorphic robotic hand lost contact with the grasped object. Subsequently, I quantified the acceleration and calculated the linear instability grasp quality metric.

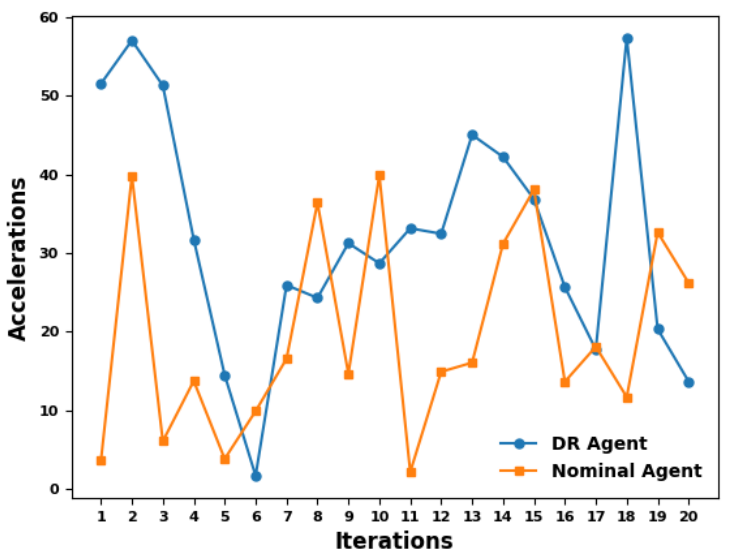


Figure 5.8: Accelerations of the grasped object during the shake task for the Nominal and DR agent

I ran the shake test for 20 iterations and plotted the results in Fig. 5.8. The average mean value of the grasp quality metric for the nominal agent is 17.19 m/s^2,

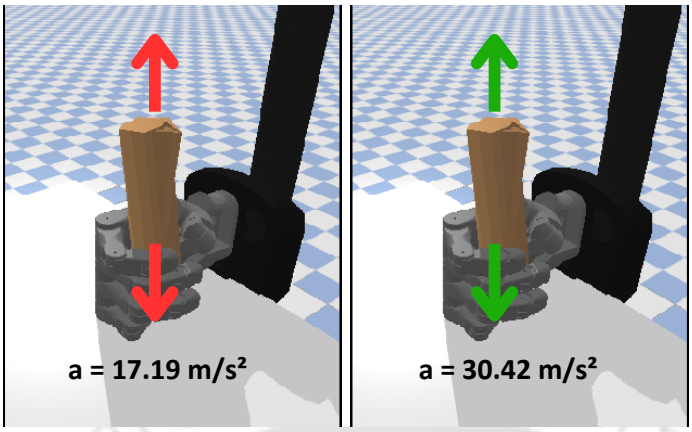


Figure 5.9: Linear instability metric calculation from shake task [105]. The average acceleration before loss of contact for the nominal agent is 17.19 m/s^2 (left) and 30.42 m/s^2 for the DR agent (right).

while that of the DR agent was 30.42 m/s^2. Percentage by which DR acceleration exceeds NDR acceleration: 63.04 %. Fig. 5.9 shows the illustration of the shake test, revealing the average acceleration before the loss of contact. These results imply that the object was more susceptible to displacement under external forces and disturbances in the case of the nominal agent, in contrast to the DR agent, suggesting a higher bionic reflex capacity for the latter.

5.6 Chapter Summary

Enabling bionic reflex control for robotic hands is essential in the pursuit of optimizing the efficiency of grasping. This chapter introduces a novel approach for enhancing the grasping and manipulation capabilities of robotic hands by employing a bionic reflex controller with both slippage and deformation prevention developed using DRL. The empirical findings, coupled with the success rates assessment and validation via the shake task, demonstrate the efficacy of the DRL controller in preventing slippage and deformation. The DRL controller autonomously learns preventive skills through interaction without human input, promising advancements in prosthetic and robotic grasping via robust, self-reliant bionic reflex controllers.

Chapter 6

Robustification of Reinforcement Learning-based Bionic Reflex Controller

6.1 Introduction

Maintaining a stable grasp on objects is a critical challenge in the realm of robotic hands and upper-limb prostheses. *RL*-trained policies are susceptible to failure in perturbed environments due to dynamic variations. To address the above problem, in this chapter, we propose an approach of augmenting a pre-trained *RL* policy with an Adaptive Sliding Mode Control (*ASMC*), particularly suited for systems with continuous state and action spaces. Leveraging the invariance property of the sliding mode algorithm against uncertainties, my approach enhances the robustness of *RL* policies trained without accounting for diverse dynamic variations in the input actions in the trained simulated settings. The main contribution of this chapter is an add-on control strategy to the trained *RL* agent, to grasp an object and lift it by rejecting disturbances, preventing both slip and deformation due to matched uncertainties. I robustify the *RL*-learned policy (control) of slip prevention during grasping trained under the nominal domain randomized environment, using a *ASMC* to take care of matched uncertainties, thus enabling improved performance under a perturbed environment.

6.2 Background and Related Work

The generalization ability of *RL* policy can be addressed using techniques like *DR*, adversarial reinforcement learning, learning to adapt (meta-learning), transfer learning, post-training augmentation, and knowledge distillation [46–48]. *DR* is a technique to bridge the sim-to-real reality gap, akin to robust control in control theory with uncertainties present. It focuses on designing controllers that maintain specific properties despite tolerable parameter variations and noise [47]. *DR*, even when trained sub-optimally in a simulator, can be effective in real-world scenarios due to its convergence properties, as discussed in the literature [106]. In adversarial reinforcement learning (ARL) the idea is to enhance robustness and controller transferability by training the controller in various environment models, not just one. ARL achieves diversity by training an adversarial sub-agent to generate challenging models that minimize cumulative rewards [47]. In robust reinforcement learning, it's framed as a two-player zero-sum differential game, where a 'disturbing' agent tries to create the worst disturbance, while a 'control' agent aims for the best control input [107, 108]. However, both domain randomization and ARL often lead to fixed policies that can result in overfitting [109]. Meta-learning, also known as Meta-RL or learning to learn, is another subfield of machine learning, which, focuses on building models that can adapt and improve performance across new tasks without extensive retraining. It offers fast adaptation for pre-trained policies in the presence of dynamic variations, improving the generalization of learned policies [110]. However, a drawback is that learning optimal policies for all possible problem scenarios can unnecessarily increase complexity, particularly for simple tasks [48]. In transfer learning, various techniques like zero-shot learning, few-shot learning, and domain adaptation are commonly applied [48], however, they may end up learning deterministic policies, which may not be efficient for sim-to-real transfer. When implementing a controller on a high-frequency physical device, meeting strict time constraints is crucial. *DNN* policies, especially ensembles, can pose challenges in this regard. Knowledge distillation, the process of transferring the expertise of a large, complex neural network to a smaller, more efficient one, reduces evaluation time. It's a technique where a reinforcement learning agent's policy, trained in a large network, is distilled into a smaller network performing at an expert level [111]. Policy distillation has proven to be efficient in various studies, surpassing domain randomization techniques [112, 113].

In the post-training augmentation-based methodology of increasing generalization capability, an augmented robust controller is added in a hierarchical approach to the reinforcement learning policy to compensate for any disturbances. Applications of $\mathcal{L}_1$ adaptive controller to robustify the pre-trained RL policy for a pendubot and quadrotor is presented in [46], which shows its efficiency in mitigating the effect of matched uncertainties. A disturbance-based observer was used to augment an RL policy to eliminate the mismatch between the simulated and the testing environment [114]. A similar approach was taken in [115], which estimated and addressed parametric uncertainties using a model reference adaptive control (MRAC) system. The distinct advantage of post-training augmentation, when compared to other policy generalization techniques discussed in the previous paragraph, lies in its ability to adapt dynamically to real-time disturbances. Unlike fixed policies, this approach accounts for variable disturbances that were not encountered during simulator training. As a result, the controller can effectively enhance the learned policy's performance in the real environment, even in the presence of unanticipated and untrained disturbances.

6.3 Problem Formulation

My approach, illustrated in Figure 6.1, starts with standard RL training in a nominal domain randomized environment, where a robotic hand grasps an object and lifts it without slippage and deformation. The nominal RL agent is trained in a domain-randomized setting, where the weights, coefficient of friction, and contact stiffness of the object are randomized, to enhance its adaptability to real-world scenarios. However, when matched uncertainties and control variations arise, the RL agent trained in the domain-randomized environment may undergo failure. To address this, I employ a hierarchical approach to bolster RL policy deployment. After obtaining a well-performing nominal policy in the nominal environment U_{RL}, I design an $ASMC$, U_{ASMC}, to complement and enhance the policy during execution, U. The $ASMC$ that I employ is an Adaptive Fast Non-singular Integral Terminal Sliding Mode ($AFNITSM$). The RL policy serves as a reference force trajectory for the $AFNITSM$ controller [116]. The $AFNITSM$ control enhances the RL policy by achieving ideal dynamics, avoiding the reaching phase, and swiftly transitioning to the sliding phase from the initial time instant. With the RL policy well-trained

in nominal domain-randomized dynamics, it acts as the nominal controller of the $AFNITSM$. The nominal control helps to achieve the ideal dynamics as if the perturbations don't exist, thus helping to avoid the reaching phase and bringing the system to the sliding phase from the initial time instant. After getting a nominal policy that functions well in the nominal environment, for policy execution, $AFNITSM$ is designed to augment and work together with the nominal policy. The $AFNITSM$ controller uses the dynamics of the nominal environment as an internal nominal model, estimates the discrepancy between the nominal model and the actual dynamics, and compensates for this discrepancy. Thus the discontinuous switching function of the $AFNITSM$ surface takes care of the matched disturbances and keeps the dynamics in the sliding surface. Since the RL policy is well-trained using the nominal dynamics, it is expected to function well in the presence of dynamic variations and disturbances represented as matched uncertainties with the $AFNITSM$ augmentation.

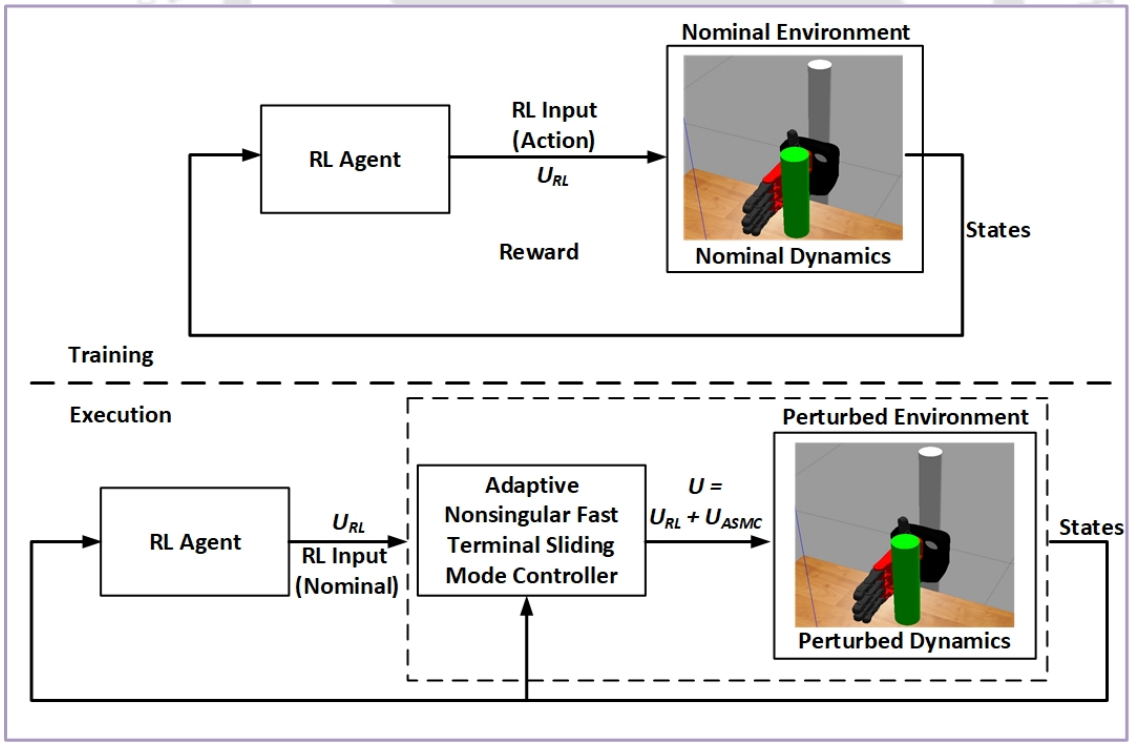


Figure 6.1: Proposed approach for RL control policy robustness improvement based on Adaptive Integral Sliding Mode Controller

6.4 Design Methodology

6.4.1 Adaptive Sliding Mode Control based Augmentation for RL Policy Robustification

6.4.1.1 *AFNITSM* controller for the robotic gripper

The dynamics of the anthropomorphic hand is given as:

$$B(\theta)\ddot{\theta} + C(\theta,\dot{\theta})\dot{\theta} + g(\theta) = \tau + \tau_{ext} \tag{6.1}$$

where $\theta, \dot{\theta}, \ddot{\theta} \in R^n$ stand for the vectors of position, velocity, and acceleration of the joints, $B(\theta) = B_0(\theta) + \Delta B(\theta)$ represents the inertia matrix, $C(\theta,\dot{\theta}) = C_0(\theta,\dot{\theta}) + \Delta C(\theta,\dot{\theta})$ represents the centripetal Coriolis matrix, $g(\theta) = g_0(\theta) + \Delta g(\theta)$ represents the gravitational matrix, τ represents the joint torques, τ_{ext} represents the external disturbance on the joint torque input. $B_0(\theta), C_0(\theta,\dot{\theta})$, and $g_0(\theta)$ are the nominal terms of the dynamic model, and $\Delta B_0(\theta), \Delta C_0(\theta,\dot{\theta})$, and $\Delta g_0(\theta)$ are the uncertainties of the dynamic model

The dynamical equation in Eq. (1) can be written as

$$B_0(\theta)\ddot{\theta} + C_0(\theta,\dot{\theta})\dot{\theta} + g_0(\theta) = \tau + \tau_d \tag{6.2}$$

where, $\tau_d = \tau_{ext} - \Delta B(\theta)\ddot{\theta} - \Delta C(\theta,\dot{\theta})\dot{\theta} - \Delta g(\theta)$ represents the lumped disturbances of the robotic gripper. I define the desired position vector by θ_d, and the tracking error is then defined as $e_1 = \theta - \theta_d$.

The derivatives of the tracking error are

$$\begin{aligned} \dot{e} &= \dot{\theta} - \dot{\theta}_d \\ \ddot{e} &= \ddot{\theta} - \ddot{\theta}_d \\ \dddot{e} &= \dddot{\theta} - \dddot{\theta}_d \end{aligned} \tag{6.3}$$

Using Equation 6.2, I have

$$\ddot{e} = {B_0}^{-1}(\theta)[\tau - f(\theta,\dot{\theta}) + \tau_d] - \ddot{\theta} \tag{6.4}$$

where, $f(\theta,\dot{\theta}) = C(\theta,\dot{\theta})\dot{\theta} + G(\theta)$. Differentiation of the above equation gives

$$\begin{aligned}\dddot{e} &= \dot{B}_0^{-1}(\theta)[\tau - f(\theta,\dot{\theta}) + \tau_d] + {B_0}^{-1}(\theta)[\dot{\tau} - f(\theta,\dot{\theta}) + \dot{\tau}_d(\theta,\dot{\theta},\ddot{\theta})] - \dddot{\theta}_d \\ &= \dot{B}_0^{-1}(\theta)[\tau - f(\theta,\dot{\theta})] + {B_0}^{-1}(\theta)[\dot{\tau} - f(\theta,\dot{\theta})] - \dddot{\theta}_d + F(\theta,\dot{\theta},\ddot{\theta})\end{aligned} \tag{6.5}$$

where $F(\theta,\dot{\theta},\ddot{\theta}) = \dot{B}_0^{-1}(\theta)\tau_d + B_0^{-1}\dot{\tau}_d(\theta,\dot{\theta},\ddot{\theta})$

Using Assumptions 1-2, I obtain $\left\|F(\theta,\dot{\theta},\ddot{\theta}\right\| < b_0 + b_1\left\|\theta\right\| + b_2\left\|\dot{\theta}\right\|^2$, where b_0, b_1 and b_2 are unknown positive constants.

6.4.2 Second-order Adaptive Integral Terminal Sliding Mode Controller Design:

The integral sliding mode surface (ISMC) surface can be given as [117]:

$$s = \dot{e} + c_1 e + c_2 \int e d\tau - \dot{e}(0) - c_1 e(0) \tag{6.6}$$

where $c_1 = diag(c_{11}, c_{12}, ..., c_{1n})$ and $c_2 = diag(c_{21}, c_{22}, ..., c_{2n})$ are positive-definite matrices. The first and second-time derivatives give

$$\dot{s} = \ddot{e} + c_1\dot{e} + c_2 e, \tag{6.7}$$

$$\ddot{s} = \dddot{e} + c_1\ddot{e} + c_2\dot{e}, \tag{6.8}$$

For faster finite time convergence and to avoid singularity problem, I utilize a fast nonsingular integral terminal sliding manifold (FNITSM). The fast nonsingular integral terminal sliding manifold is used to guarantee that s converges rapidly to equilibrium in finite time both at far and near to equilibrium, without a singularity problem occurring. The second order fast nonsingular integral terminal sliding mode controller (SOFNITSMC) surface is designed as [116, 118, 119]:

$$\sigma = \dot{s} + \int_0^t [\beta_1\lambda_1(\gamma_1,\rho_1,s,\varepsilon_1) + \beta_2\lambda_2(\gamma_2,\rho_2,\dot{s},\varepsilon_2)]d\tau \tag{6.9}$$

where,

$$\lambda_1(\gamma_1,\rho_1,s,\varepsilon_1) = \begin{cases} \mathrm{sgn}s^{\gamma_1}, & |s| \le \varepsilon_1 \\ {\varepsilon_1}^{\gamma_1-\rho_1}\mathrm{sgn}s^{\gamma_2}, & |s| > \varepsilon_1 \end{cases} \tag{6.10}$$

$$\lambda_2(\gamma_2, \rho_2, \dot{s}, \varepsilon_2) = \begin{cases} \text{sgn}\dot{s}^{\gamma_2}, & |\dot{s}| \leq \varepsilon_2 \\ {\varepsilon_2}^{\gamma_2-\rho_2}\text{sgn}\dot{s}^{\gamma_2}, & |\dot{s}| > \varepsilon_2 \end{cases} \tag{6.11}$$

$\beta_1 = diag(\beta_{11}, \beta_{12}, ..., \beta_{1n}$ and $\beta_2 = diag(\beta_{21}, \beta_{22}, ..., \beta_{2n}$ are positive-definite matrices, and γ_i, ε_i, and ρ_i are constants that satisfy $0 < \gamma_2 < 1, \gamma_1 = \gamma_2/(2-\gamma_2), \rho_i \geq 1$ and $\varepsilon_i > 0 \quad (i = 1, 2)$

The time derivative of (6.9) is

$$\dot{\sigma} = \ddot{s} + \beta_1\lambda_1(\gamma_1, \rho_1, s, \varepsilon_1) + \beta_2\lambda_2(\gamma_2, \rho_2, \dot{s}, \varepsilon_2) \tag{6.12}$$

The control law of an SMC consists of an equivalent control law τ_{eq} and a switching control law τ_{sw}. The SOITSMC is chosen as

$$\tau = \tau_{eq} + \tau_{sw} = \int_0^t (\dot{\tau}_{eq} + \dot{\tau}_{sw})d\tau \tag{6.13}$$

The τ is designed to guarantee the σ converges to zero. I can get $\dot{\tau}_{eq}$ by making $F(x, \dot{x}, \ddot{x}) = 0$, and then I design $\dot{\tau}_{sw}$ i.e., the discontinuous control action, to deal with the disturbances.

When $F(x, \dot{x}, \ddot{x}) = 0$, and let $\sigma = 0$, this enables us to obtain $\dot{\tau}_{eq}$ and $\dot{\tau}_{sw}$ as follows

$$\begin{aligned} \dot{\tau}_{eq} = &-B_0(x)\dot{B}_0^{-1}[\tau - f(x, \dot{x})] + B_0(x)(\ddot{x}_d - c_1\ddot{e} - c_2\dot{e}) - B_0[\beta_1\lambda_1(\gamma_1, \rho_1, s, \varepsilon_1) \\ &+ \beta_2\lambda_2(\gamma_2, \rho_2, s, \varepsilon_2)] + \dot{f}(x, \dot{x}), \end{aligned} \tag{6.14}$$

$$\dot{\tau}_{sw} = -B_0(x)[k\sigma + (b_0 + b_1 \|x\| + b_2 \|\dot{x}\|^2)] \tag{6.15}$$

where $k = diag(k_1, k_2, ..., k_n)$

6.4.2.0.1 Adaptive Law Design: In practice, I cannot obtain the values of b_0, b_1, and b_2 in Equation 6.15. Therefore, I will use the adaptive parameter tuning scheme to estimate them:

$$\dot{\tau}_{sw} = -B_0(x)[k\sigma + (\hat{b}_0 + \hat{b}_1 \|x\| + \hat{b}_2 \|\dot{x}\|^2)] \tag{6.16}$$

where $\hat{b}_0, \hat{b}_1$ and $\hat{b}_2$ are the respective estimates. The adaptive laws for $\hat{b}_i (i = 1, 2, 3)$ are as follows:

$$\begin{aligned}\dot{\hat{b}_0} &= \|\sigma\| \\ \dot{\hat{b}_1} &= \|\sigma\| \|x\| \\ \dot{\hat{b}_2} &= \|\sigma\| \|\dot{x}\|^2\end{aligned} \tag{6.17}$$

I define the adaptation error as $\tilde{b}_i = b_i - \hat{b}_i (i = 0, 1, 2)$. The Second-order Adaptive Integral Terminal Sliding Mode Controller ($SOAITSMC$) is then designed as follows:

$$\tau = \tau_{eq} + \tau_{asw} = \int_0^t (\dot{\tau}_{eq} + \dot{\tau}_{asw}) d\tau \tag{6.18}$$

where,

$$\begin{aligned}\dot{\tau}_{eq} = &-B_0(x)\dot{B}_0^{-1}[\tau - f(x, \dot{x})] + B_0(x)(\ddot{x}_d - c_1\ddot{e} - c_2\dot{e}) - B_0[\beta_1\lambda_1(\gamma_1, \rho_1, s, \varepsilon_1) \\ &+ \beta_2\lambda_2(\gamma_2, \rho_2, s, \varepsilon_2)] + \dot{f}(x, \dot{x}),\end{aligned} \tag{6.19}$$

$$\dot{\tau}_{asw} = -B_0(x)[k\sigma + (\hat{b}_0 + \hat{b}_1 \|x\| + \hat{b}_2 \|\dot{x}\|^2)] \tag{6.20}$$

6.4.2.1 Stability Analysis

In this section, I present the stability analysis of the SOAITSMC. The Lyapunov function candidate is considered as follows:

$$V = \frac{1}{2}(\sigma^T\sigma + \mu_0\tilde{b}_0^2 + \mu_1\tilde{b}_1^2 + \mu_2\tilde{b}_2^2) \tag{6.21}$$

The time derivative of the above equation gives:

$$\begin{aligned}\dot{V} &= \sigma^T\dot{\sigma} + \mu_0\tilde{b}_0\dot{\tilde{b}_0} + \mu_1\tilde{b}_1\dot{\tilde{b}_1} + \mu_2\tilde{b}_2\dot{\tilde{b}_2} \\ &= \sigma^T[\ddot{s} + \beta_1\lambda_1(\gamma_1, \rho_1, s, \varepsilon_1) + \beta_2\lambda_2(\gamma_2, \rho_2, \dot{s}, \varepsilon_2)] + \mu_0\tilde{b}_0\dot{\tilde{b}_0} + \mu_1\tilde{b}_1\dot{\tilde{b}_1} + \mu_2\tilde{b}_2\dot{\tilde{b}_2}\end{aligned} \tag{6.22}$$

Substituting Equations 6.5 and 6.7 into Equation. 6.22

$$\begin{aligned}\dot{V} = \sigma^T[\dot{B}_0^{-1}(x)[\tau - f(x, \dot{x})] + {B_0}^{-1}(x)[\dot{\tau} - f(x, \dot{x})] - \dddot{x}_d + F(x, \dot{x}, \ddot{x}) + c_1\ddot{e} + c_2\dot{e} \\ \beta_1\lambda_1(\gamma_1, \rho_1, s, \varepsilon_1) + \beta_2\lambda_2(\gamma_2, \rho_2, \dot{s}, \varepsilon_2)] + \mu_0\tilde{b}_0\dot{\tilde{b}_0} + \mu_1\tilde{b}_1\dot{\tilde{b}_1} + \mu_2\tilde{b}_2\dot{\tilde{b}_2}\end{aligned} \tag{6.23}$$

Using Equations (6.19) and (6.20), I have

$$\begin{aligned}
\dot{V} &= \sigma^T[-k\sigma - (\hat{b}_0 + \hat{b}_1 \|x\| + \hat{b}_2 \|\dot{x}\|)sign(\sigma) + F(x,\dot{x},\ddot{x})] + \mu_0\tilde{b}_0\dot{\hat{b}_0} + \mu_1\tilde{b}_1\dot{\hat{b}_1} + \mu_2\tilde{b}_2\dot{\hat{b}_2} \\
&\leq \|F(x,\dot{x},\ddot{x})\|\|\sigma\| - \left\|\hat{b}_0 + \hat{b}_1\|x\| + \hat{b}_2\|\dot{x}\|^2\right\| + \left\|b_0 + b_1\|x\| + b_2\|\dot{x}\|^2\right\|\|\sigma\| - \\
&\left\|b_0 + b_1\|x\| + b_2\|\dot{x}\|^2\right\|\|\sigma\| + \quad \mu_0\tilde{b}_0\dot{\hat{b}_0} + \mu_1\tilde{b}_1\dot{\hat{b}_1} + \mu_2\tilde{b}_2\dot{\hat{b}_2} \\
&\leq -[\left\|(b_0 + b_1\|x\| + b_2\|\dot{x}\|^2)\right\| - \|F(x,\dot{x},\ddot{x}\|\|\sigma\|] - \left\|\tilde{b}_0\right\|(\|\sigma\| - \mu_0\sigma) - \left\|\tilde{b}_1\right\|(\|\sigma\|\|x\| - \\
&\mu_1\|\sigma\|\|x\|) - \left\|\tilde{b}_2\right\|(\|\sigma\|\|\dot{x}\|^2 - \quad \mu_2\|\sigma\|\|\dot{x}\|^2) \\
&\leq -\varrho\sqrt{2}\frac{\|\sigma\|}{\sqrt{2}} - \xi_0\sqrt{2\mu_0}\frac{\left\|\tilde{b}_0\right\|}{\sqrt{2\mu_0}} - \xi_1\sqrt{2\mu_1}\frac{\left\|\tilde{b}_1\right\|}{\sqrt{2\mu_1}} - \xi_2\sqrt{2\mu_2}\frac{\left\|\tilde{b}_2\right\|}{\sqrt{2\mu_2}}
\end{aligned}$$

where

$$\begin{aligned}
\varrho &= \left\|(b_0 + b_1\|x\| + b_2\|\dot{x}\|^2\right\| - \|F(x,\dot{x},\ddot{x}\|, \\
\xi_0 &= \|\sigma\| - \mu_0\|\sigma\| = (1-\mu_0)\|\sigma\|, \\
\xi_1 &= (\|\sigma\|\|x\| - \mu_1\|\sigma\|\|x\|) = (1-\mu_1)\|\sigma\|\|x\|, \\
\xi_2 &= (\|\sigma\|\|\dot{x}\|^2 - \mu_2\|\sigma\|\|\dot{x}\|^2) = (1-\mu_2)\|\sigma\|\|\dot{x}\|^2
\end{aligned} \tag{6.24}$$

I then obtain

$$\begin{aligned}
\dot{V} &\leq -\min\{\sqrt{2}\varrho, \sqrt{2\mu_0^{-1}}\xi_0, \sqrt{2\mu_1^{-1}}\xi_1, \sqrt{2\mu_2^{-1}}\xi_2\} \times \left(\frac{\|\sigma\|}{\sqrt{2}} + \frac{\sqrt{\mu_0}}{\sqrt{2}}\left\|\tilde{b}_0\right\|\right. \\
&\left. + \frac{\sqrt{\mu_1}}{\sqrt{2}}\left\|\tilde{b}_1\right\| + \frac{\sqrt{\mu_2}}{\sqrt{2}}\left\|\tilde{b}_2\right\|\right) \\
&\leq -\alpha V^{1/2}
\end{aligned} \tag{6.25}$$

where,

$$\alpha = \min\{\sqrt{2}\varrho, \sqrt{2{\mu_0}^{-1}}\xi_0, \sqrt{2{\mu_1}^{-1}}\xi_1, \sqrt{2{\mu_2}^{-1}}\xi_2\}$$

and $\alpha > 0$. The above inequality holds if $\mu_0 < 1, \mu_1 < 1$ and $\mu_2 < 1$.

If $t_0 = 0$, then σ in Equation 6.9 can converge to zero in a finite time $t_1 = V^{1-1/2}(0)/\alpha(1-1/2) = 2V^{1/2}(0)/\alpha$. Thus, the sliding variable s will converge to zero in a finite time. The tracking error e can asymptotically converge to zero. This completes the proof of stability.

The control architecture of the post-training augmented robust RL policy is shown in Fig. 6.2. The architecture is almost similar to indirect force control with

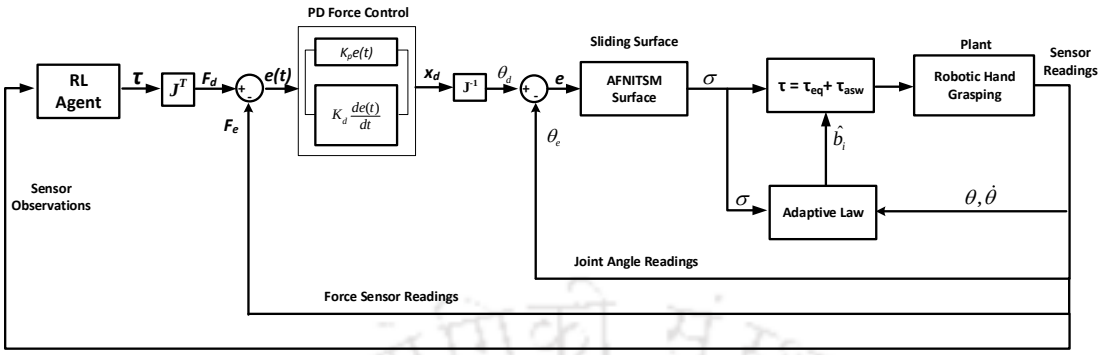


Figure 6.2: Post-training augmented robust RL control diagram for perturbed environment

inner position loop [77, 78], where the reference force is received from the trained RL policy. After the adaptive control input is applied to the anthropomorphic hand, the force error signals are generated between the experimental force signals F_e and the desired force signals F_d. F_d is related to the generated pre-trained RL policy (torques) through the Jacobian transpose. The error value, $e(t)$, becomes the input value for a proportional derivative (PD) force control algorithm made by a proportional and a derivative regulator, whose output x_d is given by:

$$x_d = K_p e(t) + K_d \frac{de(t)}{dt} \tag{6.26}$$

where, K_p and K_d are the proportional and derivative gains, respectively. The values of K_p and K_d can be identified by means of manual tuning through experimental tests to adjust the reaction time and limit the overshooting of the measured F_e with respect to F_d.

The output x_d is added to the measured value of the finger position θ_e to give the position error signals, e_p. This error signal serves as the input for the sliding surface as given in Eq. 6.3.

The nominal parameters for the simulation and the control parameters taken are given in Table. 6.1 (details of the dynamical equation given in Appendix. C) and Table. 6.2.

Table 6.1: Physical parameters of the anthropomorphic hand used for simulation

Symbol	**Definition**	**Values**
m_1	Nominal mass of proximal phalange	0.01 kg
m_2	Nominal mass of medial phalange	0.005 kg
m_3	Nominal mass of distal phalange	0.004 kg
l_1	Length of proximal phalange	5.5 cm
l_2	Length of medial phalange	3.5 cm
l_3	Length of distal phalange	2.5 cm
J_1	Moment of inertia of proximal phalange	0.1 kg m^2
J_2	Moment of inertia of medial phalange	0.85 kg m^2
J_3	Moment of inertia of distal phalange	0.75 kg m^2
g	Acceleration due to gravity	9.81 m/s^2

Table 6.2: Parameters of the proposed controller and their values

Parameters	**Values**
c_1	10
c_2	85
β_1	3
β_2	5
ρ_1	1.3
ρ_2	1.1
γ_1	0.1
γ_2	0.1
k_1	150
k_2	150
ε_1	0.02
ε_2	0.02
ϵ	0.1
$b_i(0)$	0

6.5 Results and Discussions

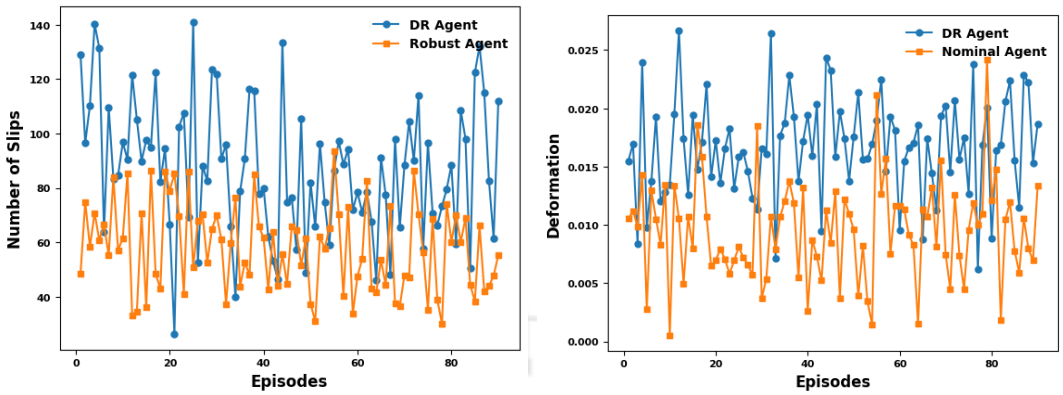
6.5.1 Success Rates

To assess the effectiveness of the robustified post-augmented RL controller compared to the nominal domain-randomized agent, I conducted a success rate test. This test involved counting the number of slips and amount of deformation during object grasping and lifting under conditions with randomized parameters (object weights, friction, and contact stiffness) and mismatched uncertainties introduced by sinusoidal disturbances as below:

$$\tau_d = \left[\; 2\sin(t) + 0.5\sin(200\pi t) \; \right] \tag{6.27}$$

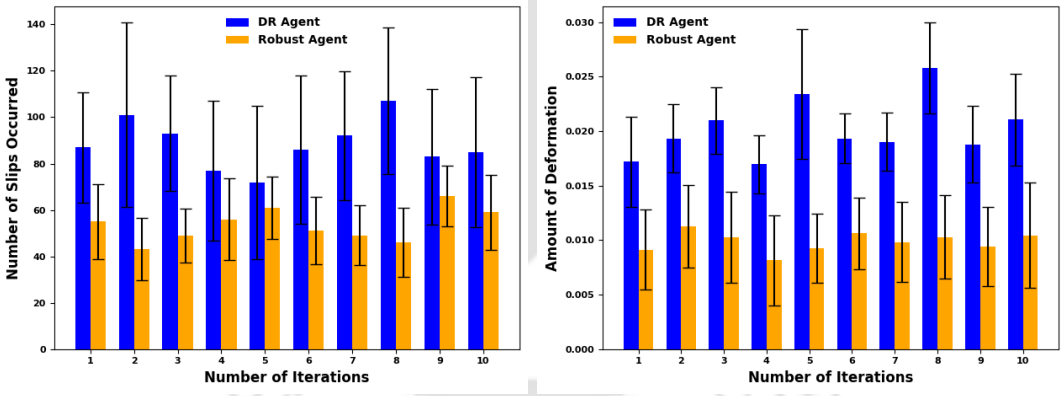
The above time-varying external input disturbances is a standard benchmark disturbance that is taken up for robotic manipulator control problems [120, 121]. The success rate plot is displayed in Fig. 6.3, where the number of slips and amount of deformation occurring is calculated for the DR and the robustified agent. Fig. 6.3(a) shows the success rates of preventing slippage while grasping unknown objects for nominal and DR agents. The x-axis shows the number of episodes for which the learned agent was tested. The y-axis represents the frequency of slips occurring. From the success rates plot, it is evident that the agent trained in a DR environment has been able to prevent more slips than the nominal agent. Fig. 6.3(b) shows the deformation occurring on the deformable object when grasped by the nominal and the DR-trained agent on unknown objects tasks with different object properties. It is to be noted that the success rate plots demonstrate the superior performance of the post-training augmented robust controller, which resulted in fewer slip occurrences compared to the nominal domain-randomized agent.

To validate the efficacy of the results obtained, I repeat the above experiments of success rate test over ten trials. I then present the error bar plots for each iteration to provide the visual representation of variability and illustrate the statistical significance of the experimental outcomes (shown in Fig. 6.4). To compare the two groups of error bar plots of slippage in Fig. 6.4(a), i.e., the DR agent group and the robust agent, I calculated the difference in means between the two groups and compared it to their combined standard deviations. This will give you a sense of the magnitude of the difference relative to the variability within each group. The percentage difference came as 103.35%. Further, the Cohen's D test gave a value



(a) (b)

Figure 6.3: Performance Tests of success rates on Unseen Objects. Parameters randomized were the object weights, stiffness, and friction coefficients while grasping: (a) Slips prevented by DR Agent and Robust Agent at unseen objects task. (b) Amount of deformation (in mm) prevented by DR Agent and Robust Agent on unknown objects task



(a) (b)

Figure 6.4: Error bar plots of statistical tests for (a) Slips prevented by DR Agent and Robust Agent at unseen objects task. (b) Amount of deformation (in mm) prevented by DR Agent and Robust Agent on unknown objects task

of 4.08. Cohen's D test considers the effect size, which quantifies the magnitude of the difference between groups. Similarly, for the deformation error bar plots Fig. 6.4(b), the magnitude of the difference relative to the variability percentage is 197%,

and Cohen's D test value is 5.35. From both performance tests and statistical variability tests, I can conclude that the post-training based robustified agent is better equipped to improve the generalization capability than the nominal DR agent.

6.5.2 Control parameters and comparison

The proposed control architecture integrates a reinforcement learning (RL) policy with a second order adaptive terminal sliding mode controller (ATSMC) to robustify the control action against matched disturbances. In this design, the RL policy directly determines the control action based on learned optimal strategies, while the ATSMC layer compensates for uncertainties and disturbances, effectively reducing the simulation-to-real (Sim2Real) gap.

Reinforcement Learning for Optimal Control

The RL policy learns a mapping from system states to control actions that maximizes a long-term performance criterion, providing a model-free solution to complex, nonlinear control problems [122]. This approach overcomes limitations of traditional model-based controllers, which require precise system models and can be sensitive to modeling errors [123].

Adaptive Terminal Sliding Mode Control (ATSMC)

The ATSMC layer robustifies the RL-generated control actions by ensuring finite-time convergence to a desired sliding manifold despite matched disturbances. Key technical features include:

- Finite-Time Convergence: Terminal sliding mode control guarantees convergence of system states in finite time, which is crucial for systems requiring rapid stabilization [124].
- Chattering Reduction: The second order formulation and adaptive gain tuning mitigate the chattering effects commonly associated with conventional sliding mode control.
- Adaptive Gain Tuning: Gains are automatically adjusted in real time, ensuring robust performance without the need for extensive manual tuning [120].

Comparison with Other Control Methodologies

- PID Control: Traditional PID controllers, while simple and widely used, rely on manually tuned fixed gains that are often inadequate for handling the nonlinearities and disturbances present in underactuated systems [125].

- Conventional Adaptive Control: Although adaptive control techniques adjust gains based on error feedback, they generally assume a known system structure and may not react swiftly enough to rapid dynamic changes or unexpected disturbances [126].

- Model-Based Control (e.g., LQR, MPC): Model-based controllers depend heavily on the accuracy of the system model. Inaccuracies can lead to suboptimal performance or instability, whereas the RL-ATSMC framework is more resilient to modeling errors due to its inherent robustness and adaptive tuning [123, 127].

- Robust H-Infinity Control: H-infinity control methods focus on worst-case disturbance rejection, which often results in conservative controller designs. In contrast, the RL-ATSMC approach adapts in real time to achieve both optimal performance and robustness without excessive conservatism [128].

- Fuzzy and Neural Network-Based Controllers: While these controllers can handle nonlinearities, they typically require extensive tuning and large datasets for training, and they generally lack the finite-time stability and rigorous robustness guarantees provided by sliding mode control [129].

Automatic Parameter Tuning via Reinforcement Learning

A significant advantage of the proposed framework is that the RL policy not only selects optimal control actions but also indirectly tunes the Adaptive Terminal Sliding Mode Control (ATSMC) parameters. By formulating the control problem as an optimal control task, the RL agent adjusts controller gains to minimize tracking errors, control efforts, and robustness margins, effectively automating what is traditionally a manual tuning process.

After the RL policy converges to optimal control parameters, further adaptability is provided by the ATSMC layer, which continuously refines the control response by adapting sliding mode parameters in real time. This hierarchical adaptation ensures:

- Fine-tuned robustness against disturbances and model uncertainties.
- Finite-time convergence, preventing slow asymptotic behavior common in traditional adaptive controllers.
- Reduced chattering, as the second-order sliding mode formulation dynamically smooths the control inputs.

This dynamic adaptation is particularly beneficial during the transition from simulation to real-world applications, where environmental variations and model mismatches often degrade control performance. By leveraging RL for optimal tuning and ATSMC for real-time robustness, the system achieves both learning-based adaptability and model-free disturbance rejection, making it highly suitable for complex, nonlinear, and uncertain robotic systems.

6.6 Chapter Summary

In this chapter, I propose a method for enhancing the robustness of a pre-trained reinforcement learning (RL) policy through post-training augmented adaptive control, with a focus on improving the learned policy generalization capabilities. I adopt a hierarchical approach by integrating an adaptive sliding mode controller into an existing RL policy pipeline, thereby improving the pre-trained agent's robustness and enhancing its overall efficiency. My results, evaluated through success rate tests, measuring slip occurrences and amount of deformation during object grasping and lifting under uncertain random disturbances, demonstrate the effectiveness of the robustified RL policy.

Chapter 7

Conclusion and Future Work

This thesis focuses on the bionic reflex control for an underactuated anthropomorphic hand optimized based on grasped quality measure. The thesis proposes an optimization framework to improve the design of the underactuated anthropomorphic hand. Thereafter, an autonomous bionic reflex control built upon the optimized hand to prevent slippage and deformation of grasped objects is explored. In this chapter, I first highlight the overall contribution of the thesis to the body of knowledge of grasping and manipulation research and then describe the individual contributions made by the thesis in separate later sections. I also identify the limitations and possible extensions of my contributions. Finally, in Section 7.7, I highlight a few promising research directions for future research in the field.

7.1 Overall Contribution

The thesis contributes to the area of grasping and manipulation in robotics by examining state-of-the-art design optimization and bionic reflex control techniques. It adds to the body of knowledge by introducing a novel optimization framework for the design of an underactuated hand. Autonomous control methodologies and adaptive robust generalization capabilities creating efficient grasping by the underactuated anthropomorphic hand are the other contributions to the body of knowledge. Together these methods form the essence of the bionic reflex control framework proposed in this thesis.

7.2 Design Optimization based on Grasp Quality Metric

7.2.1 Summary of Contribution

Several current works in the optimization of the design of anthropomorphic hands are based on maximizing finger workspace intersection, graspable object size, human trajectory tracking, or grasped force optimization. For an underactuated anthropomorphic hand, apart from replicating the geometry, biomimicry of the human counterpart requires design based on functional biomimesis. Optimization to generate a stable grasp for underactuated anthropomorphic hands has been a recent focus. I analyzed the state-of-the-art optimization strategies in the field of anthropomorphic hands and proposed a design methodology based on the grasp quality metric. To account for the optimization of the underactuated anthropomorphic hand based on the grasp quality metric, I considered the potential contact robustness and potential grasp robustness as part of the objective function. Experimental evaluation of the proposed optimization framework proved the efficacy of the framework in arriving at a solution that produces significantly good grasps.

7.2.2 Limitations and Possible Extensions

1. The current framework focuses exclusively on five specific grasping configurations relevant to daily living activities. An extension of this research could involve the incorporation of additional grasp types and quality metrics for comparative assessment. Furthermore, future investigations may explore the inclusion of adduction-abduction grasping actions to broaden the scope of analysis.

2. Presently, the optimization framework lacks consideration for lower-level grasp control factors such as slippage and deformation prevention. A potential enhancement to this framework could involve integrating mechanisms within the optimization objective function to autonomously address lower-level manipulation and control behaviors during the design phase.

3. An additional avenue for extending this research would be the optimization

of linkage-based underactuated mechanisms based on grasp quality metrics. This approach would facilitate comparative analysis with tendon-driven underactuated hands proposed in this thesis, thereby providing insights into the effectiveness and efficiency of different actuation methods for robotic grasping systems.

7.3 Slip Prevention with Reinforcement Learning Based Control

7.3.1 Summary of Contribution

Incorporating bionic reflex as a lower-level controller to adapt the grasping forces, enabling slippage prevention, could improve the efficiency of robotic hands. Slippage prevention techniques in the current state-of-the-art research have been applied for both grippers and anthropomorphic hands. However, existing slippage controllers required prior knowledge of the grasped object properties like coefficient of friction or thresholding/labeling the generated raw grasping force or slippage signals. The chapter proposes the design of an autonomous slippage prevention control framework for anthropomorphic hands. Slip is detected using the variations of grasping forces through a wavelet transformation technique, eliminating the need for thresholding. The detected slip signal informs an RL agent and enables learning of the optimal grasping forces to be applied at the fingertips, preventing slippage and dropping of the grasped object autonomously. Experimental evaluation demonstrates the viability of the RL-based controller in preventing slip by learning through interaction without prior knowledge or human supervision.

7.3.2 Limitations and Possible Extensions

1. Slip detection currently relies on a conventional Discrete Wavelet Transform (DWT) utilizing Haar wavelet decomposition. Future extensions of this research may explore the efficacy of alternative adaptive signal processing methods and feature engineering techniques to enhance slippage discrimination within grasping signals, thereby potentially improving overall system performance.

2. The Reinforcement Learning (RL) algorithm currently employed is a conventional policy gradient algorithm chosen to establish the effectiveness of the framework discussed in this thesis. Future investigations could consider the incorporation of novel RL algorithms for comparative analysis, with a focus on achieving reduced convergence times, enhanced adaptability, and improved generalization capabilities.

3. There is further scope for improving the domain randomization techniques presented in this work, particularly through the introduction of additional tasks alongside parameter variations or the integration of adversarial RL approaches. These enhancements aim to bolster the robustness of the agent within diverse environmental conditions, thereby advancing the efficacy of the overall learning process.

7.4 Deformation Prevention with Reinforcement Learning Based Control

7.4.1 Summary of Contribution

Robotic hands frequently encounter slippage and deformation during object manipulation, issues rarely encountered by humans due to their sensory receptors, experiential learning, and motor memory. Although remarkable advances have been made, minimizing deformation while preventing slippage is still not a trivial issue. I analyzed the state-of-the-art deformation prevention techniques in robotic grasping research. Past endeavors in the realm of deformation control predominantly relied on knowing the constitutive model parameters and supervised learning approaches, necessitating human intervention during thresholding and labeling tasks. I proposed an innovative bionic reflex control pipeline based on perceived slippage and deformation measurement, with slippage and deformation prevention, leveraging reinforcement learning (RL); thereby eliminating the need for human intervention during control design. The empirical findings, coupled with the success rates assessment and validation tests, demonstrate the efficacy of the DRL controller in preventing slippage and deformation. I anticipate that this autonomous, RL-based bionic reflex controller will catalyze the development of dependable and highly ef-

ficient robotic and prosthetic hands, revolutionizing human-robot interaction and assistive technologies.

7.4.2 Limitations and Possible Extensions

1. The simulation framework utilizes a Neo-Hookean volumetric Finite Element Method (FEM) to characterize the behavior of deformable objects. Future investigations may explore alternative simulators employing co-rotational FEM formulations or incremental potential contact models to enhance the precision and accuracy of deformable object modeling.

2. The proposed deformation calculation methodology involves the computation of signed volumes within the mesh structure, providing an estimation of the overall deformation akin to real-world volume calculation scenarios. Subsequent research endeavors may focus on refining deformation detection techniques by integrating visuo-tactile-based sensors for improved accuracy and sensitivity in capturing deformations.

3. The current research implements a bionic reflex mechanism by incorporating deformation calculation into the framework, which proves effective. Future advancements could involve the integration of stiffness detection along with deformation calculation within the bionic reflex framework to augment grasping success rates and overall efficiency. Furthermore, comparative analyses between sole deformation calculation versus stiffness detection methodologies for deformation control within the bionic reflex framework present avenues for further investigation.

7.5 Robustification of Reinforcement Learning-based Bionic Reflex Controller

7.5.1 Summary of Contribution

Maintaining a stable grasp on objects is a critical challenge in the realm of robotic hands and upper-limb prostheses. Despite its importance, external disturbances often disrupt grasp stability, leading to slippage and, if the forces are not optimal, may

deform the object. RL agents trained over different domain randomized scenarios can be efficient in autonomously controlling slippage and deformation. However, RL-trained policies are susceptible to failure in perturbed environments due to dynamic variations. To address this limitation, in this chapter, I propose an approach of augmenting a pre-trained RL policy with an adaptive sliding mode controller, particularly suited for systems with continuous state and action spaces. Leveraging the invariance property of the sliding mode algorithm against uncertainties, my approach enhances the robustness of RL policies to tackle real-time diverse dynamic variations in either simulated or real-world settings. Numerical experiments validate the effectiveness of my method in fortifying RL policies trained in idealized environments.

7.5.2 Limitations and Possible Extensions

1. The adaptive sliding mode controller employed to enhance the robustness of the Reinforcement Learning (RL) agent operates under finite-time convergence criteria. Subsequent research avenues may involve the exploration of enhanced adaptive control methodologies incorporating predefined convergence criteria. These methodologies could undergo rigorous testing under diverse conditions encompassing randomized physics parameters and dynamic variations within the environment.

2. The present study addresses matched disturbances within the plant dynamics. Future enhancements to the controller design could encompass strategies for mitigating the influence of both matched and unmatched disturbances. Additionally, research efforts may focus on the integration of model learning techniques to facilitate autonomous acquisition of unknown system dynamics. This direction of inquiry seeks to fortify the adaptability and robustness of the controller against a wider array of perturbations and uncertainties.

7.6 Comparison with state-of-the-art

Table 7.1: Comparison of my method with the state-of-the-art

Method	Hand Type	Fingertip Sensor	Simulation/ Real	Slip Response Time	Object Deformation Minimized
Multi-threshold-based control [36]	Five Fingered Hand	FSR	Real	<160 ms	No
ON-OFF Control [35]	Five Fingered Hand	FSR	Real	≈ 200 ms	No
Switched Adaptive Regrasping [130]	Two Fingered Gripper	FSR Laser Sensor	Real	≈ 500 ms	Yes
Hybrid Position/ Force Control [131]	Three-link soft finger	NA	Simulation	≈ 300 ms	No
Feedback Linearization [132]	Two Fingered Gripper	NA	Simulation	≈ 750 ms	No
Variable Impedance Control [133]	Two Fingered Gripper	NA	Simulation	≈ 200 ms	No
Ours	Five Fingered Hand	Force sensor	Simulation	≈ 200 ms	Yes

Although this thesis represents a novel effort in implementing bionic reflex control using reinforcement learning, a comparative analysis of our bionic reflex control system against state-of-the-art slippage and deformation prevention methods is reported. Table. 7.1 lists the values for various elements of comparison, including

well-recognized indicators of the performance of our method and several recent slippage and deformation prevention methods. The comparison criteria include the type of robotic hand, the sensor technology employed, whether simulation study or real-world implementation, the slip response time, and object deformation minimization. Though our main baselines are five-fingered hands, I also list two and three-fingered grippers. Compared to both grippers and five-fingered hands, our proposed methodology performs competitively, validating the efficacy of reinforcement learning-based control.

This research has led to the development of a real-time adaptive bionic reflex controller trained in a physics-based simulator that is deployed within a Sim-to-Sim testing environment. Future investigations may be taken up toward Sim-to-Real implementation by integrating the controller into a 3D-printed anthropomorphic hand equipped with embedded control.

7.7 Future Directions

During the time of this Ph.D. thesis, the research on bionic reflex control has gained great momentum. While RL-based bionic reflex controller have been explored, technologies like generative AI, large language models (LLMs), and explainable AI is taking center stage. At the end of the dissertation, I highlight a few research directions that may be worthwhile to investigate in the near future.

7.7.1 Generative AI for Bionic Reflex Control

Generative AI technology, particularly generative models like Generative Adversarial Networks (GANs) or Variational Autoencoders (VAEs), can play a significant role in slippage and deformation prevention in grasping objects with anthropomorphic robotic hands. **Augmenting real-world data:** Generative models can be used to augment existing datasets/ grasping scenarios. By generating synthetic images or simulations of objects with different shapes, textures, material properties, or sensor noise, these models can help train robotic grasping systems to handle a wider range of objects and grasp configurations, thereby reducing the likelihood of slippage or deformation during real-world manipulation. **Generating synthetic training data:** Generative models can be employed to simulate the behavior of objects under

different grasping conditions. By generating realistic simulations of object deformation and slippage, these models can help robotic systems anticipate potential issues and adjust their grasp strategy accordingly to prevent slippage or deformation during grasping and manipulation, thus improving the generalization and robustness of control policies. **Grasp Planning and Optimization:** Generative models can assist in optimizing grasp planning algorithms by generating diverse candidate grasps for a given object. By exploring the grasp space using generative techniques, robotic systems can identify more robust grasp configurations that minimize the risk of slippage or deformation while ensuring stable manipulation. **Adaptive Control and Feedback:** Generative models can be used to generate feedback signals for adaptive control algorithms. By comparing the predicted outcome of a grasp with the actual outcome, these models can provide feedback to the robotic system, enabling it to adjust its grasp in real time to prevent slippage or deformation during manipulation.

7.7.2 Natural Language Processing (NLP) for Bionic Reflex Control

Natural Language Processing (NLP) and Large Language Models (LLMs) can play a fascinating role in future advancements of robotic grasping, particularly in preventing slippage and deformation. **Understanding user intent:** NLP can be used to interpret user instructions describing the desired object and its intended manipulation. This allows robots to grasp objects based on high-level human commands, reducing the need for complex programming of specific grasp configurations. NLP can facilitate natural language communication between humans and robotic systems, allowing users to provide instructions, feedback, and preferences regarding grasping tasks. By understanding human intent and preferences, robotic systems can adapt their grasp strategy to minimize slippage and deformation while achieving the desired task objectives. **Reasoning about object properties and manipulation:** By understanding the language describing the object (e.g., "fragile vase") and the desired action (e.g., "pick up carefully"), NLP can guide the robot towards safer grasp strategies minimizing the risk of slippage or deformation. Large language models can process contextual information from various sources, including visual data, sensor readings, and historical grasp experiences. By understanding the context of the grasping task, robotic systems can adjust their grasp strategy to account

for factors such as object shape, surface texture, and environmental conditions, thereby reducing the risk of slippage and deformation. **Generating instructions for human assistance:** If the robot encounters an object it cannot grasp safely on its own, NLP can be used to generate clear and concise instructions for a human to assist in the manipulation process, fostering effective collaboration. NLP can be used to generate natural language descriptions justifying the robot's chosen grasp for an object. This explanation can improve human-robot collaboration by promoting trust and understanding of the robot's decision-making process. **Grasp Planning and Optimization:** NLP techniques can be used to generate natural language descriptions of grasping scenarios and constraints. Large language models can then assist in generating and evaluating candidate grasp configurations that minimize the likelihood of slippage and deformation based on the provided descriptions. Additionally, NLP can aid in optimizing grasp planning algorithms by processing textual feedback and refining grasp strategies iteratively.

7.7.3 Neurosymbolic AI for Bionic Reflex Control

Neurosymbolic AI, which combines neural networks with symbolic reasoning, can offer innovative approaches to slippage and deformation prevention in grasping objects with anthropomorphic robotic hands. **Knowledge Representation and Reasoning for Grasp Planning:** Neurosymbolic systems can represent knowledge about object properties (e.g., fragility, texture) and grasping constraints (e.g., force limitations, stability requirements) in a symbolic format, enabling the robot to reason about safe and effective grasps. Symbolic representations can be integrated with deep learning models trained on vast grasping data. This allows the robot to leverage the strengths of both approaches: symbolic reasoning for logical reasoning and decision-making, and deep learning for capturing complex relationships and patterns from data. **Planning and Adaptation in Uncertain Environments:** Neurosymbolic AI can handle situations with incomplete or uncertain information. By reasoning about possible scenarios and their outcomes, the robot can select grasps that are robust to different environmental conditions or unexpected object properties, minimizing the risk of slippage or deformation. Neurosymbolic systems can continuously learn from new experiences and adapt their grasping strategies over time. This is crucial for robots operating in dynamic environments where objects and situations might change frequently. **Explainable AI for Grasp Planning:**

Neurosymbolic AI techniques can provide interpretable explanations for grasp planning decisions by combining neural network-based models with symbolic reasoning mechanisms for generating explanations. By explaining the rationale behind grasp selections, robotic systems can improve transparency and trustworthiness, enabling human operators to understand and further refine the chosen grasp strategies for slippage and deformation prevention. **Semantic Grasping:** Neurosymbolic AI can enable robotic systems to understand the semantic context of grasping tasks by combining deep learning with symbolic representations of object properties, grasp affordances, and task objectives. By reasoning about the semantics of objects and their interactions with the environment, robotic systems can select appropriate grasp configurations that minimize the risk of slippage and deformation.

Publications

Journals

1. Basumatary, Hirakjyoti, and Shyamanta M. Hazarika. "State of the art in bionic hands." *IEEE Transactions on Human-Machine Systems* 50.2 (2020): 116-130.

2. Basumatary, Hirakjyoti, and Shyamanta M. Hazarika. "Design optimization of an underactuated tendon-driven anthropomorphic hand based on grasp quality measures." *Robotica* 40.11 (2022): 4056-4075.

3. Dey, Abhijit, Hirakjyoti Basumatary, and Shyamanta M. Hazarika. "A Decade of Haptic Feedback for Upper Limb Prostheses." *IEEE Transactions on Medical Robotics and Bionics* (2023).

4. Basumatary, Hirakjyoti, Adhar, Daksh, and Hazarika, Shymanta Moni, "Robustifying a Reinforcement Learning based Bionic Reflex Controller through Adaptive Sliding Mode Control", *Robotica (2024).*

Conferences

1. Basumatary, Hirakjyoti, et al. "Non-Assembly 3-D Printed Anthropomorphic Hand Prosthesis for Low-Income Countries." *Advances in Robotics-5th International Conference of The Robotics Society.* 2021.

2. Basumatary, Hirakjyoti, et al. "Low Level Grasp Controller for Slippage and Deformation Prevention exploiting Deep Reinforcement Learning." *Proceedings of the 2023 6th International Conference on Advances in Robotics.* 2023.

3. Basumatary, Hirakjyoti, Adhar, Daksh, Shaw, Riddhiman and Hazarika, Shymanta M. "Grasp Force Optimization as a Bilinear Matrix Inequality Problem: A Deep Learning Approach." *6th National Conference on Multidisciplinary Design, Analysis and Optimization (NCMDAO 2023)* 2023.

4. Basumatary, Hirakjyoti, Sarkar, Tamen Thapa,and Hazarika, Shymanta M. "Robustifying a Reinforcement Learning Based Grasped Object Slippage Prevention Controller Through Adaptive Sliding Mode Control" *28th International Conference on Methods and Models in Automation and Robotics, (MMAR 2024)* 2024.

Book Chapters

1. Basumatary, Hirakjyoti, Adhar, Daksh, Shaw, Riddhiman and Hazarika, Shymanta M. "Grasp Force Optimization as a Bilinear Matrix Inequality Problem: A Deep Learning Approach." *Advances in Multidisciplinary Design, Analysis and Optimization, Springer,* 2025.

References

[1] Marzke M.W. and Marzke R.F. (2000) 'Evolution of the human hand: approaches to acquiring, analysing and interpreting the anatomical evidence', *The Journal of Anatomy*, vol. 197(1), pp. 121–140.

[2] Cordella F., Ciancio A.L., Sacchetti R., Davalli A., Cutti A.G., Guglielmelli E., and Zollo L. (2016) 'Literature review on needs of upper limb prosthesis users', *Frontiers in Neuroscience*, vol. 10, p. 209.

[3] Murray C.D. (2009) 'Being like everybody else: the personal meanings of being a prosthesis user', *Disability and Rehabilitation*, vol. 31(7), pp. 573–581.

[4] Dimante D., Logina I., Sinisi M., and Krūmiņa A., 'Sensory feedback in upper limb prostheses', in 'Proceedings of the Latvian Academy of Sciences. Section B. Natural, Exact, and Applied Sciences.', vol. 74 (2020), pp. 308–317.

[5] Deng H., Zhong G., Li X., and Nie W. (2016) 'Slippage and deformation preventive control of bionic prosthetic hands', *IEEE/ASME Transactions On Mechatronics*, vol. 22(2), pp. 888–897.

[6] Svensson P., Wijk U., Björkman A., and Antfolk C. (2017) 'A review of invasive and non-invasive sensory feedback in upper limb prostheses', *Expert Review of Medical Devices*, vol. 14(6), pp. 439–447.

[7] Ghafoor U., Kim S., and Hong K.S. (2017) 'Selectivity and longevity of peripheral-nerve and machine interfaces: a review', *Frontiers in Neurorobotics*, vol. 11, p. 59.

[8] Deng H., Zhang Y., and Duan X.G. (2016) 'Wavelet transformation-based fuzzy reflex control for prosthetic hands to prevent slip', *IEEE Transactions on Industrial Electronics*, vol. 64(5), pp. 3718–3726.

[9] Tucker M.R., Olivier J., Pagel A., Bleuler H., Bouri M., Lambercy O., Millán J.d.R., Riener R., Vallery H., and Gassert R. (2015) 'Control strategies for active lower extremity prosthetics and orthotics: a review', *Journal of Neuroengineering and Rehabilitation*, vol. 12(1), pp. 1–30.

[10] Michalec R. (2011) *Modeling and control of multifingered dextrous manipulation for humanoid robot hands*, Ph.D. thesis, Université Pierre et Marie Curie-Paris VI.

[11] Noce E., Bellingegni A.D., Ciancio A.L., Sacchetti R., Davalli A., Guglielmelli E., and Zollo L. (2019) 'Emg and eng-envelope pattern recognition for prosthetic hand control', *Journal of Neuroscience Methods*, vol. 311, pp. 38–46.

[12] Stephens-Fripp B., Alici G., and Mutlu R. (2018) 'A review of non-invasive sensory feedback methods for transradial prosthetic hands', *IEEE Access*, vol. 6, pp. 6878–6899.

[13] Schweitzer W., Thali M.J., and Egger D. (2018) 'Case-study of a user-driven prosthetic arm design: bionic hand versus customized body-powered technology in a highly demanding work environment', *Journal of Neuroengineering and Rehabilitation*, vol. 15, pp. 1–27.

[14] Jang C.H., Yang H.S., Yang H.E., Lee S.Y., Kwon J.W., Yun B.D., Choi J.Y., Kim S.N., and Jeong H.W. (2011) 'A survey on activities of daily living and occupations of upper extremity amputees', *Annals of rehabilitation medicine*, vol. 35(6), pp. 907–921.

[15] Basumatary H. and Hazarika S.M. (2020) 'State of the art in bionic hands', *IEEE Transactions on Human-Machine Systems*, vol. 50(2), pp. 116–130.

[16] Deng H., Zhong G., Li X., and Nie W. (2017) 'Slippage and deformation preventive control of bionic prosthetic hands', *IEEE/ASME Transactions on Mechatronics*, vol. 22(2), pp. 888–897.

[17] Nakagawa-Silva A., Thakor N.V., Cabibihan J.J., and Soares A.B. (2019) 'A bio-inspired slip detection and reflex-like suppression method for robotic manipulators', *IEEE Sensors Journal*, vol. 19(24), pp. 12443–12453.

[18] Zhu J., Cherubini A., Dune C., Navarro-Alarcon D., Alambeigi F., Berenson D., Ficuciello F., Harada K., Kober J., Li X., et al. (2022) 'Challenges and outlook in robotic manipulation of deformable objects', *IEEE Robotics & Automation Magazine*, vol. 29(3), pp. 67–77.

[19] Wang Y., Li W., Togo S., Yokoi H., and Jiang Y. (2021) 'Survey on main drive methods used in humanoid robotic upper limbs', *Cyborg and Bionic Systems*.

[20] Ozawa R., Kobayashi H., and Hashirii K. (2013) 'Analysis, classification, and design of tendon-driven mechanisms', *IEEE transactions on robotics*, vol. 30(2), pp. 396–410.

[21] Mottard A., Laliberté T., and Gosselin C., 'Underactuated tendon-driven robotic/prosthetic hands: Design issues.', in 'Robotics: Science and Systems', vol. 7 (2017).

[22] Wen L., Li Y., Cong M., Lang H., and Du Y., 'Design and optimization of a tendon-driven robotic hand', in '2017 IEEE International Conference on Industrial Technology (ICIT)', (IEEE, 2017), pp. 767–772.

[23] Datta R., Pradhan S., and Bhattacharya B. (2015) 'Analysis and design optimization of a robotic gripper using multiobjective genetic algorithm', *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 46(1), pp. 16–26.

[24] Dong H., Asadi E., Qiu C., Dai J., and Chen I.M. (2018) 'Geometric design optimization of an under-actuated tendon-driven robotic gripper', *Robotics and Computer-Integrated Manufacturing*, vol. 50, pp. 80–89.

[25] Salman H.D., Bakhy S.H., and Hamzah M.N., 'Design and optimization of coupled and self-adaptive of an underactuated robotic hand using particle swarm optimization', in 'IOP Conference Series: Materials Science and Engineering', vol. 928 (IOP Publishing, 2020), p. 022153.

[26] Ko T. (2020) 'A tendon-driven robot gripper with passively switchable underactuated surface and its physics simulation based parameter optimization', *IEEE Robotics and Automation Letters*, vol. 5(4), pp. 5002–5009.

[27] You W.S., Lee Y.H., Kang G., Oh H.S., Seo J.K., and Choi H.R. (2019) 'Kinematic design optimization for anthropomorphic robot hand based on interactivity of fingers', *Intelligent Service Robotics*, vol. 12, pp. 197–208.

[28] Votta A.M., Günay S.Y., Zylich B., Skorina E., Rameshwar R., Erdoğmuş D., and Onal C.D., 'Kinematic optimization of an underactuated anthropomorphic prosthetic hand', in '2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)', (IEEE, 2020), pp. 3397–3403.

[29] Roa M.A. and Suárez R. (2015) 'Grasp quality measures: review and performance', *Autonomous robots*, vol. 38(1), pp. 65–88.

[30] Mnyussiwalla H., Seguin P., Vulliez P., and Gazeau J.P. (2022) 'Evaluation and selection of grasp quality criteria for dexterous manipulation', *Journal of Intelligent & Robotic Systems*, vol. 104(2), p. 20.

[31] Pozzi M., Malvezzi M., and Prattichizzo D. (2016) 'On grasp quality measures: Grasp robustness and contact force distribution in underactuated and compliant robotic hands', *IEEE Robotics and Automation Letters*, vol. 2(1), pp. 329–336.

[32] Prattichizzo D., Salisbury J.K., and Bicchi A., 'Contact and grasp robustness measures: Analysis and experiments', in 'Experimental Robotics IV: The 4th International Symposium, Stanford, California, June 30–July 2, 1995', (Springer, 1997), pp. 83–90.

[33] Suthar B., Abdulrahman Y., and Zweiri Y. (2024) 'Robotic fingers: Advancements, challenges, and future directions-a comprehensive review', *IEEE Access*.

[34] Romeo R.A. and Zollo L. (2020) 'Methods and sensors for slip detection in robotics: A survey', *IEEE Access*, vol. 8, pp. 73027–73050.

[35] Romeo R.A., Lauretti C., Gentile C., Guglielmelli E., and Zollo L. (2021) 'Method for automatic slippage detection with tactile sensors embedded in prosthetic hands', *IEEE Transactions on Medical Robotics and Bionics*, vol. 3(2), pp. 485–497.

[36] Yang D. and Wu G. (2021) 'A multi-threshold-based force regulation policy for prosthetic hand preventing slippage', *IEEE Access*, vol. 9, pp. 9600–9609.

[37] Engeberg E.D. and Meek S.G. (2011) 'Adaptive sliding mode control for prosthetic hands to simultaneously prevent slip and minimize deformation of grasped objects', *IEEE/ASME Transactions on Mechatronics*, vol. 18(1), pp. 376–385.

[38] Zhang Y., Xu X., Xia R., and Deng H. (2022) 'Stiffness-estimation-based grasping force fuzzy control for underactuated prosthetic hands', *IEEE/ASME Transactions on Mechatronics*, vol. 28(1), pp. 140–151.

[39] Andrecioli R. and Engeberg E.D. (2013) 'Adaptive sliding manifold slope via grasped object stiffness detection with a prosthetic hand', *Mechatronics*, vol. 23(8), pp. 1171–1179.

[40] Gillen S.M., *Improving Reinforcement Learning for Robotics with Control and Dynamical Systems Theory* (University of California, Santa Barbara, 2022).

[41] Tobin J., Biewald L., Duan R., Andrychowicz M., Handa A., Kumar V., McGrew B., Ray A., Schneider J., Welinder P., et al., 'Domain randomization and generative models for robotic grasping', in '2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)', (IEEE, 2018), pp. 3482–3489.

[42] Huber J., Hélénon F., Watrelot H., Amar F.B., and Doncieux S., 'Domain randomization for sim2real transfer of automatically generated grasping datasets', in '2024 IEEE International Conference on Robotics and Automation (ICRA)', (IEEE, 2024), pp. 4112–4118.

[43] Ren X., Luo J., Solowjow E., Ojea J.A., Gupta A., Tamar A., and Abbeel P., 'Domain randomization for active pose estimation', in '2019 International Conference on Robotics and Automation (ICRA)', (IEEE, 2019), pp. 7228–7234.

[44] Rabinovitz C., Grupen N., and Tamar A., 'Unsupervised feature learning for manipulation with contrastive domain randomization', in '2021 IEEE International Conference on Robotics and Automation (ICRA)', (IEEE, 2021), pp. 10153–10159.

[45] Leibovich G., Jacob G., Endrawis S., Novik G., and Tamar A., 'Validate on sim, detect on real-model selection for domain randomization', in '2022 International Conference on Robotics and Automation (ICRA)', (IEEE, 2022), pp. 7528–7535.

[46] Cheng Y., Zhao P., Wang F., Block D.J., and Hovakimyan N. (2022) 'Improving the robustness of reinforcement learning policies with l1 adaptive control', *IEEE Robotics and Automation Letters*, vol. 7(3), pp. 6574–6581.

[47] Salvato E., Fenu G., Medvet E., and Pellegrino F.A. (2021) 'Crossing the reality gap: A survey on sim-to-real transferability of robot controllers in reinforcement learning', *IEEE Access*, vol. 9, pp. 153171–153187.

[48] Güitta-López L., Boal J., and López-López Á.J. (2023) 'Learning more with the same effort: how randomization improves the robustness of a robotic deep reinforcement learning agent', *Applied Intelligence*, vol. 53(12), pp. 14903–14917.

[49] Xu Z., 'The functional capacity of the humanlike robotic hands', in 'Hand Function', (Springer, 2019), pp. 291–312.

[50] Balasubramanian R. and Dollar A.M., 'Performance of serial underactuated mechanisms: Number of degrees of freedom and actuators', in 'Redundancy in Robot Manipulators and Multi-Robot Systems', (Springer, 2013), pp. 1–13.

[51] Xu R. and Özgüner Ü. (2008) 'Sliding mode control of a class of underactuated systems', *Automatica*, vol. 44(1), pp. 233–241.

[52] Xu K., Liu H., Du Y., and Zhu X. (2014) 'Design of an underactuated anthropomorphic hand with mechanically implemented postural synergies', *Advanced Robotics*, vol. 28(21), pp. 1459–1474.

[53] Prattichizzo D., Malvezzi M., Gabiccini M., and Bicchi A. (2013) 'On motion and force controllability of precision grasps with hands actuated by soft synergies', *IEEE transactions on robotics*, vol. 29(6), pp. 1440–1456.

[54] Kakoty N.M. (2014) *A biomimetic hand with EMG based grasp emulation*, Ph.D. thesis, Computer Sc. & Engg., Tezpur University, PhD Thesis.

[55] Gosselin C., Pelletier F., and Laliberte T., 'An anthropomorphic underactuated robotic hand with 15 dofs and a single actuator', in '2008 IEEE International Conference on Robotics and Automation', (IEEE, 2008), pp. 749–754.

[56] Inouye J.M. and Valero-Cuevas F.J. (2014) 'Anthropomorphic tendon-driven robotic hands can exceed human grasping capabilities following optimization', *The International Journal of Robotics Research*, vol. 33(5), pp. 694–705.

[57] Chen T., Wang L., Haas-Heger M., and Ciocarlie M. (2020) 'Underactuation design for tendon-driven hands via optimization of mechanically realizable manifolds in posture and torque spaces', *IEEE Transactions on Robotics*, vol. 36(3), pp. 708–723.

[58] Malvezzi M. and Prattichizzo D., 'Evaluation of grasp stiffness in underactuated compliant hands', in '2013 IEEE International conference on robotics and automation', (IEEE, 2013), pp. 2074–2079.

[59] Prattichizzo D. and Trinkle J.C. (2016) 'Grasping', *Springer Handbook of Robotics*, pp. 955–988.

[60] Farnioli E., Gabiccini M., and Bicchi A., 'Quasi-static analysis of synergistically underactuated robotic hands in grasping and manipulation tasks', in 'Human and Robot Hands', (Springer, 2016), pp. 211–233.

[61] Malvezzi M., Salvietti G., and Prattichizzo D., 'Evaluation of grasp stiffness in underactuated compliant hands exploiting environment constraints', in 'ROMANSY 22–Robot Design, Dynamics and Control', (Springer, 2019), pp. 409–416.

[62] Almeida L. and Moreno P., 'Potential grasp robustness for underactuated hands: New heuristics and uncertainty considerations', in '2020 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)', (IEEE, 2020), pp. 233–238.

[63] Bicchi A. (1995) 'On the closure properties of robotic grasping', *The International Journal of Robotics Research*, vol. 14(4), pp. 319–334.

[64] Malvezzi M., Gioioso G., Salvietti G., and Prattichizzo D. (2015) 'Syngrasp: A matlab toolbox for underactuated and compliant hands', *IEEE Robotics & Automation Magazine*, vol. 22(4), pp. 52–68.

[65] Zangrandi A., D'Alonzo M., Cipriani C., and Di Pino G. (2021) 'Neurophysiology of slip sensation and grip reaction: insights for hand prosthesis control of slippage', *Journal of Neurophysiology*, vol. 126(2), pp. 477–492.

[66] Osborn L.E. (2014) *Reflex: A closed-loop tactile feedback system for use in upper limb prosthesis grip control*, Ph.D. thesis, Johns Hopkins University.

[67] Luo S., Bimbo J., Dahiya R., and Liu H. (2017) 'Robotic tactile perception of object properties: A review', *Mechatronics*, vol. 48, pp. 54–67.

[68] Gamal M., Mousa M.H., Eldawlatly S., and Elbasiouny S.M. (2021) 'In-silico development and assessment of a kalman filter motor decoder for prosthetic hand control', *Computers in Biology and Medicine*, vol. 132, p. 104353.

[69] Aboseria M., Clemente F., Engels L.F., and Cipriani C. (2018) 'Discrete vibrotactile feedback prevents object slippage in hand prostheses more intuitively than other modalities', *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26(8), pp. 1577–1584.

[70] Li Y., Wang P., Li R., Tao M., Liu Z., and Qiao H. (2022) 'A survey of multifingered robotic manipulation: Biological results, structural evolvements, and learning methods', *Frontiers in Neurorobotics*, vol. 16, p. 843267.

[71] Murray R.M., Li Z., and Sastry S.S., *A mathematical introduction to robotic manipulation* (CRC press, 2017).

[72] Cordella F., Gentile C., Zollo L., Barone R., Sacchetti R., Davalli A., Siciliano B., and Guglielmelli E., 'A force-and-slippage control strategy for a poliarticulated prosthetic hand', in '2016 IEEE International Conference on Robotics and Automation (ICRA)', (IEEE, 2016), pp. 3524–3529.

[73] Muratore F., Ramos F., Turk G., Yu W., Gienger M., and Peters J. (2022) 'Robot learning from randomized simulations: A review', *Frontiers in Robotics and AI*, vol. 9, p. 799893.

[74] Khin P.M., Low J.H., Ang Jr M.H., and Yeow C.H. (2021) 'Development and grasp stability estimation of sensorized soft robotic hand', *Frontiers in Robotics and AI*, vol. 8, p. 619390.

[75] Nazari K., Mandil W., et al. (2022) 'Proactive slip control by learned slip model and trajectory adaptation', *arXiv preprint arXiv:2209.06019*.

[76] James J.W. and Lepora N.F. (2020) 'Slip detection for grasp stabilization with a multifingered tactile robot hand', *IEEE Transactions on Robotics*, vol. 37(2), pp. 506–519.

[77] Siciliano B., Sciavicco L., Villani L., and Oriolo G., *Force control* (Springer, 2009).

[78] Carbone G., Iannone S., and Ceccarelli M. (2010) 'Regulation and control of larm hand iii', *Robotics and Computer-Integrated Manufacturing*, vol. 26(2), pp. 202–211.

[79] Sutton R.S. and Barto A.G., *Reinforcement learning: An introduction* (MIT press, 2018).

[80] Kober J., Bagnell J.A., and Peters J. (2013) 'Reinforcement learning in robotics: A survey', *The International Journal of Robotics Research*, vol. 32(11), pp. 1238–1274.

[81] Shahid A.A., Roveda L., Piga D., and Braghin F., 'Learning continuous control actions for robotic grasping with reinforcement learning', in '2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)', (IEEE, 2020), pp. 4066–4072.

[82] Wang X., Wang S., Liang X., Zhao D., Huang J., Xu X., Dai B., and Miao Q. (2022) 'Deep reinforcement learning: A survey', *IEEE Transactions on Neural Networks and Learning Systems*.

[83] Valencia D., Jia J., Li R., Hayashi A., Lecchi M., Terezakis R., Gee T., Liarokapis M., MacDonald B.A., and Williams H., 'Comparison of model-based and model-free reinforcement learning for real-world dexterous robotic manipulation tasks', in '2023 IEEE International Conference on Robotics and Automation (ICRA)', (IEEE, 2023), pp. 871–878.

[84] Lillicrap T.P., Hunt J.J., Pritzel A., Heess N., Erez T., Tassa Y., Silver D., and Wierstra D. (2015) 'Continuous control with deep reinforcement learning', *arXiv preprint arXiv:1509.02971*.

[85] Yang H., Hu X., Cao L., and Sun F., 'A new slip-detection method based on pairwise high frequency components of capacitive sensor signals', in '2015 5th International Conference on Information Science and Technology (ICIST)', (IEEE, 2015), pp. 56–61.

[86] Zhang T., Zhang N., Li Y., Zeng B., and Jiang L. (2022) 'Design and experimental evaluation of a sensorimotor-inspired grasping strategy for dexterous prosthetic hands', *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 738–748.

[87] Cretu A.M., Payeur P., and Petriu E.M. (2011) 'Soft object deformation monitoring and learning for model-based robotic hand manipulation', *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 42(3), pp. 740–753.

[88] Makihara K., Domae Y., Ramirez-Alpizar I.G., Ueshiba T., and Harada K. (2022) 'Grasp pose detection for deformable daily items by pix2stiffness estimation', *Advanced Robotics*, vol. 36(12), pp. 600–610.

[89] Shen B., Jiang Z., Choy C., Guibas L.J., Savarese S., Anandkumar A., and Zhu Y. (2022) 'Acid: Action-conditional implicit visual dynamics for deformable object manipulation', *arXiv preprint arXiv:2203.06856*.

[90] Ji W., Zhang J., Xu B., Tang C., and Zhao D. (2021) 'Grasping mode analysis and adaptive impedance control for apple harvesting robotic grippers', *Computers and Electronics in Agriculture*, vol. 186, p. 106210.

[91] Duan X.G., Zhang Y., and Deng H., 'A simple control method to avoid overshoot for prosthetic hand control', in '2014 IEEE International Conference on Information and Automation (ICIA)', (IEEE, 2014), pp. 736–739.

[92] Jiang L., Tian X., Zhan Q., Xu Q., and Zhang Y. (2023) 'Impedance control of an anthropomorphic hands without finger force sensors', *IEEE Transactions on Automation Science and Engineering*.

[93] Kaboli M., Yao K., and Cheng G., 'Tactile-based manipulation of deformable objects with dynamic center of mass', in '2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)', (IEEE, 2016), pp. 752–757.

[94] Mouaze N. and Birglen L. (2022) 'Bistable compliant underactuated gripper for the gentle grasp of soft objects', *Mechanism and Machine Theory*, vol. 170, p. 104676.

[95] Wang W. and Ahn S.H. (2017) 'Shape memory alloy-based soft gripper with variable stiffness for compliant and effective grasping', *Soft Robotics*, vol. 4(4), pp. 379–389.

[96] Milojević A., Linß S., Ćojbašić Ž., and Handroos H. (2021) 'A novel simple, adaptive, and versatile soft-robotic compliant two-finger gripper with an inherently gentle touch', *Journal of Mechanisms and Robotics*, vol. 13(1), p. 011015.

[97] Coumans E. and Bai Y. (2016), 'Pybullet, a python module for physics simulation for games, robotics and machine learning (https://pybullet.org/wordpress/)', .

[98] Collins J., Chand S., Vanderkop A., and Howard D. (2021) 'A review of physics simulators for robotic applications', *IEEE Access*, vol. 9, pp. 51416–51431.

[99] Basumatary H. and Hazarika S.M. (2022) 'Design optimization of an underactuated tendon-driven anthropomorphic hand based on grasp quality measures', *Robotica*, vol. 40(11), pp. 4056–4075.

[100] Hu Y., Schneider T., Wang B., Zorin D., and Panozzo D. (2020) 'Fast tetrahedral meshing in the wild', *ACM Transactions on Graphics (TOG)*, vol. 39(4), pp. 117–1.

[101] Arriola-Rios V.E., Guler P., Ficuciello F., Kragic D., Siciliano B., and Wyatt J.L. (2020) 'Modeling of deformable objects for robotic manipulation: A tutorial and review', *Frontiers in Robotics and AI*, vol. 7, p. 82.

[102] Zhang C. and Chen T., 'Efficient feature extraction for 2d/3d objects in mesh representation', in 'Proceedings 2001 International Conference on Image Processing (Cat. No. 01CH37205)', vol. 3 (IEEE, 2001), pp. 935–938.

[103] Ma X., Chen L., Gao Y., Liu D., and Wang B. (2023) 'Modeling contact stiffness of soft fingertips for grasping applications', *Biomimetics*, vol. 8(5), p. 398.

[104] Le T.N., Lundell J., Abu-Dakka F.J., and Kyrki V., 'A novel simulation-based quality metric for evaluating grasps on 3d deformable objects', in '2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)', (IEEE, 2022), pp. 3123–3129.

[105] Huang I., Narang Y., Eppner C., Sundaralingam B., Macklin M., Bajcsy R., Hermans T., and Fox D. (2022) 'Defgraspsim: Physics-based simulation of grasp outcomes for 3d deformable objects', *IEEE Robotics and Automation Letters*, vol. 7(3), pp. 6274–6281.

[106] Chen X., Hu J., Jin C., Li L., and Wang L. (2021) 'Understanding domain randomization for sim-to-real transfer', *arXiv preprint arXiv:2110.03239*.

[107] Pinto L., Davidson J., Sukthankar R., and Gupta A., 'Robust adversarial reinforcement learning', in 'International Conference on Machine Learning', (PMLR, 2017), pp. 2817–2826.

[108] Morimoto J. and Doya K. (2005) 'Robust reinforcement learning', *Neural Computation*, vol. 17(2), pp. 335–359.

[109] Rice L., Wong E., and Kolter Z., 'Overfitting in adversarially robust deep learning', in 'International Conference on Machine Learning', (PMLR, 2020), pp. 8093–8104.

[110] Nagabandi A., Clavera I., Liu S., Fearing R.S., Abbeel P., Levine S., and Finn C. (2018) 'Learning to adapt in dynamic, real-world environments through meta-reinforcement learning', *arXiv preprint arXiv:1803.11347*.

[111] Rusu A.A., Colmenarejo S.G., Gulcehre C., Desjardins G., Kirkpatrick J., Pascanu R., Mnih V., Kavukcuoglu K., and Hadsell R. (2015) 'Policy distillation', *arXiv preprint arXiv:1511.06295*.

[112] Kadokawa Y., Zhu L., Tsurumine Y., and Matsubara T. (2023) 'Cyclic policy distillation: Sample-efficient sim-to-real reinforcement learning with domain randomization', *Robotics and Autonomous Systems*, vol. 165, p. 104425.

[113] Niu Z., Yuan J., Ma X., Xu Y., Liu J., Chen Y.W., Tong R., and Lin L. (2023) 'Knowledge distillation-based domain-invariant representation learning for domain generalization', *IEEE Transactions on Multimedia*.

[114] Kim J.W., Shim H., and Yang I., 'On improving the robustness of reinforcement learning-based controllers using disturbance observer', in '2019 IEEE 58th Conference on Decision and Control (CDC)', (IEEE, 2019), pp. 847–852.

[115] Guha A. and Annaswamy A. (2020) 'Mrac-rl: A framework for on-line policy adaptation under parametric model uncertainty', *arXiv preprint arXiv:2011.10562*.

[116] Hao S., Hu L., and Liu P.X. (2021) 'Second-order adaptive integral terminal sliding mode approach to tracking control of robotic manipulators', *IET Control Theory & Applications*, vol. 15(17), pp. 2145–2157.

[117] Utkin V. and Shi J., 'Integral sliding mode in systems operating under uncertainty conditions', in 'Proceedings of 35th IEEE Conference on Decision and Control', vol. 4 (IEEE, 1996), pp. 4591–4596.

[118] Li P., Ma J., Zheng Z., and Geng L., 'Fast nonsingular integral terminal sliding mode control for nonlinear dynamical systems', in '53rd IEEE Conference on Decision and Control', (IEEE, 2014), pp. 4739–4746.

[119] Alattas K.A., Mobayen S., Din S.U., Asad J.H., Fekih A., Assawinchaichote W., and Vu M.T. (2021) 'Design of a non-singular adaptive integral-type finite time tracking control for nonlinear systems with external disturbances', *IEEE Access*, vol. 9, pp. 102091–102103.

[120] Mondal S. and Mahanta C. (2014) 'Adaptive second order terminal sliding mode controller for robotic manipulators', *Journal of the Franklin Institute*, vol. 351(4), pp. 2356–2377.

[121] Boukattaya M., Mezghani N., and Damak T. (2018) 'Adaptive nonsingular fast terminal sliding-mode control for the tracking problem of uncertain dynamical systems', *ISA transactions*, vol. 77, pp. 1–19.

[122] Sutton R.S., Barto A.G., et al., *Reinforcement learning: An introduction*, vol. 1 (MIT press Cambridge, 1998).

[123] Khalil I., Doyle J., and Glover K., *Robust and optimal control*, vol. 2 (Prentice hall New York, 1996).

[124] Yu X., Feng Y., and Man Z. (2020) 'Terminal sliding mode control–an overview', *IEEE Open Journal of the Industrial Electronics Society*, vol. 2, pp. 36–52.

[125] Åström K.J. and Hägglund T., *Advanced PID control* (ISA-The Instrumentation, Systems and Automation Society, 2006).

[126] Ioannou P.A. and Sun J., *Robust adaptive control*, vol. 1 (PTR Prentice-Hall Upper Saddle River, NJ, 1996).

[127] Camacho E.F., Bordons C., Camacho E.F., and Bordons C. (2007) 'Introduction to model predictive control', *Model Predictive Control*, pp. 1–11.

[128] Glover K., 'H-infinity control', in 'Encyclopedia of systems and control', (Springer, 2021), pp. 896–902.

[129] Jang J.S.R., Sun C.T., and Mizutani E. (1997) 'Neuro-fuzzy and soft computing-a computational approach to learning and machine intelligence [book review]', *IEEE Transactions on automatic control*, vol. 42(10), pp. 1482–1484.

[130] Al-Mohammed M., Adem R., and Behal A. (2022) 'A switched adaptive controller for robotic gripping of novel objects with minimal force', *IEEE Transactions on Control Systems Technology*, vol. 31(1), pp. 17–26.

[131] Fakhari A., Kao I., and Keshmiri M. (2019) 'Modeling and control of planar slippage in object manipulation using robotic soft fingers', *ROBOMECH Journal*, vol. 6(1), p. 15.

[132] Fakhari A., Keshmiri M., Kao I., and Hadian Jazi S. (2016) 'Slippage control in soft finger grasping and manipulation', *Advanced Robotics*, vol. 30(2), pp. 97–108.

[133] Logothetis M., Karras G.C., Alevizos K., and Kyriakopoulos K.J., 'A variable impedance control strategy for object manipulation considering non–rigid grasp', in '2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)', (IEEE, 2020), pp. 7411–7416.

[134] Bicchi A., 'Force distribution in multiple whole-limb manipulation', in '[1993] Proceedings IEEE International Conference on Robotics and Automation', (IEEE, 1993), pp. 196–201.

[135] Mavrogiannis C.I., Bechlioulis C.P., Liarokapis M.V., and Kyriakopoulos K.J., 'Task-specific grasp selection for underactuated hands', in '2014 IEEE International Conference on Robotics and Automation (ICRA)', (IEEE, 2014), pp. 3676–3681.

[136] Carbone G., *Grasping in robotics*, vol. 10 (Springer Science & Business Media, 2012).

[137] Chen C.H. and Naidu D.S., *Fusion of hard and soft control strategies for the robotic hand* (John Wiley & Sons, 2017).

Appendix A: Compliant Grasp Model

The classical grasp model assumes rigid bodies, point contacts, and is quasistatic, i.e., inertial forces are ignored. However, for more realistic modeling, the rigid body model can be extended to include compliance. Introducing compliance in the grasp model helps implement compliant behavior of the hand and the grasped object thus, increasing the contact area and thereby the robustness of the grasp. Details of the compliant grasp model are given in the next section. The nomenclatures describing the mathematical symbols have been included in the Table 3.1.

A.1. The Complaint Grasp Model

In static equilibrium conditions, by utilizing the virtual work principle, we can write the equation of object and hand equilibrium, respectively as [58]:

$$\boldsymbol{w} = -G\boldsymbol{\lambda} \tag{1}$$

$$\boldsymbol{\tau} = J^T\boldsymbol{\lambda} \tag{2}$$

However, if $\mathcal{N}(G) \cap \mathcal{N}(J^T) \neq 0$, then the solution of the two algebraic system of equations, i.e. Eq. 1 is not unique. [134] solved this problem by framing the problem into a linearized quasistatic system. The method extends the rigid-body model to include compliance and assumes that the contact constraint violation and contact force variation are related through a contact stiffness model. This problem usually occurs in a power grasp for an underactuated hand, when the number of contacts points is more than the controllable degrees of freedom.

In modeling the compliant underactuated grasp, a set of hypothetical springs are introduced at the contact locations between the object and the fingers. The contact force variation $\delta\boldsymbol{\lambda}$ from the initial equilibrium position is given as:

$$\delta\boldsymbol{\lambda} = K_c(\delta\boldsymbol{c^h} - \delta\boldsymbol{c^o}) = K_c(\mathbf{J}\delta\boldsymbol{q} - \mathbf{G}^T\delta\boldsymbol{u}) \tag{3}$$

The actual joint variables, $\boldsymbol{q}$, may differ from the reference one, $\boldsymbol{q_r}$, if the hand structural parameters are not perfectly stiff. This can be modeled as:

$$\delta\boldsymbol{\tau} = K_q(\delta\boldsymbol{q_r} - \delta\boldsymbol{q}) \tag{4}$$

The hand joint reference configuration $\boldsymbol{q_r}$ for an underactuated hand is defined with lower-dimensional inputs than the number of DoF (hand joints), i.e., we consider that the relative movements of the joints are coupled/constrained. A vector $\boldsymbol{z}$, named Lagrangian variables, is introduced, whose dimensionality equals the number of DoFs of the hand.

The Lagrangian variables can be split up as $\boldsymbol{z} = \left[\boldsymbol{z_a}^T \ \boldsymbol{z_p}^T\right]^T$, with dimension $n_z = n_{za} + n_{zp}$ with $\boldsymbol{z_a}$ representing the active/controllable input variables (e.g. actuators) while the variables, $\boldsymbol{z_p}$, representing the passive/uncontrollable variables (e.g. the spring system).

The joint references $\boldsymbol{q_r}$, as a function of the Lagrangian variables $\boldsymbol{z}$, is defined as:

$$\boldsymbol{q_r} = f_z(\boldsymbol{z}) \tag{5}$$

where, $f_z : R^{n_z} \rightarrow R^{n_q}$ represents non–linear kinematic relationship. We can express the above equation in a linearized form, by introducing a coupling matrix called synergy matrix, S [58], and express the variation of joint reference variable as a function of Langrangian variable variation.

$$\delta\boldsymbol{q_r} = S\delta\boldsymbol{z} \tag{6}$$

The postural synergy based underactuation is imposed upon the grasping system by means of the synergy matrix.

The external contact wrenches can be expressed in the Lagrangian co-ordinate as

$$\boldsymbol{\sigma} = S^T\boldsymbol{\tau} \tag{7}$$

$\boldsymbol{\sigma}$ can be splitted as

$$\boldsymbol{\sigma_a} = {S_a}^T \boldsymbol{\tau} \qquad \boldsymbol{\sigma_p} = {S_p}^T \boldsymbol{\tau} \tag{8}$$

The above equation represents the contribution to hand equilibrium of the forces generated by the actuators, $\boldsymbol{\sigma_a}$ (active Lagrangian forces) and of the passive elements, $\boldsymbol{\sigma_p}$ (passive Lagrangian forces).

A.2. Soft synergies

The postural synergy-based kinematic model is insufficient to explain a grasp by an underactuated multi-fingered hand accurately. In addition, the interaction with a grasped object must be represented and included in the study. This has recently been accomplished using "Soft Synergy," a modeling framework that allows synergies to control the hand's kinematics in such a way that the final posture matches the geometry of the gripped object. This is accomplished by taking into account the hand's structural compliance as well as the object's rigidity. Specifically, during a grip, forces appear as the fingers of the hand make contact with the object, and torques appear at the finger joints as a result of these forces. These torques alter hand posture based on contact and joint compliance, allowing the hand to conform to the shape of the object. [135].

In the Lagrangian co-ordinate space, assuming close loop control of the position of actuators, the actuation forces are expressed as

$$\delta\boldsymbol{\sigma_a} = K_{za}(\delta\boldsymbol{z_{ra}} - \delta\boldsymbol{z_a}) \tag{9}$$

$$\delta\boldsymbol{\sigma_p} = -K_{zp}\delta\boldsymbol{z_p} \tag{10}$$

Hence, we can write

$$\delta\boldsymbol{\sigma} = \begin{bmatrix} \delta\boldsymbol{\sigma_a} \\ \delta\boldsymbol{\sigma_p} \end{bmatrix} = K_z(\delta\boldsymbol{z_r} - \delta\boldsymbol{z}) \tag{11}$$

$$K_z = \begin{bmatrix} K_{za} & 0 \\ 0 & K_{zp} \end{bmatrix}, \quad \delta\boldsymbol{z_r} = \begin{bmatrix} \delta\boldsymbol{z_{ra}} \\ 0 \end{bmatrix}, \quad \delta\boldsymbol{z} = \begin{bmatrix} \delta\boldsymbol{z_a} \\ \delta\boldsymbol{z_p} \end{bmatrix} \tag{12}$$

A.3. Compliant grasp analysis in underactuated hands

The linearization of object equilibrium of Eq. 1, yields the following expression in quasi-static condition

$$\delta \boldsymbol{w} + G\delta \boldsymbol{\lambda} = 0 \tag{13}$$

Further, the linearization of the hand equilibrium of Eq. 2, yields the joint torque variation expression as

$$\delta \boldsymbol{\tau} = J^T \delta\lambda + K_{J,q}\delta \boldsymbol{q} + K_{J,u}\delta \boldsymbol{u} \tag{14}$$

where, $K_{J,q}\delta \boldsymbol{q} = \frac{\partial J^T \boldsymbol{\lambda_0}}{\partial \boldsymbol{q}}$ and $K_{J,u}\delta \boldsymbol{q} = \frac{\partial J^T \boldsymbol{\lambda_0}}{\partial \boldsymbol{u}}$, i.e., the derivatives of hand Jacobian matrix wrt $\boldsymbol{q}$ and $\boldsymbol{u}$.

Variation of the Lagrangian forces can be expressed as

$$\delta \boldsymbol{\sigma} = S^T \delta \boldsymbol{\tau} + K_{S,z}\delta \boldsymbol{z} \tag{15}$$

where, $K_{S,z} = \frac{\partial S^T \boldsymbol{\tau_0}}{\partial \boldsymbol{z}}$. Here, since S is constant, therefore $K_{S,z} = [\mathbf{0}]$ [58].

All the main equations 4, 6,11, 13, 14, 15 can be summarized in the equation below. For detailed analysis, the readers can refer to [136]:

$$A\delta \boldsymbol{x} = \delta \boldsymbol{y} \tag{16}$$

$$A = \begin{bmatrix} -G & 0 & 0 & 0 & 0 & 0 & 0 \\ J^T & K_{J,u} & -I & K_{J,q} & 0 & 0 & 0 \\ 0 & 0 & S^T & 0 & -I & K_{S,z} & 0 \\ C_c & G^T & 0 & -J & 0 & 0 & 0 \\ 0 & 0 & -I & -K_q & 0 & 0 & K_q \\ 0 & 0 & 0 & 0 & C_z & I & 0 \\ 0 & 0 & 0 & 0 & 0 & S & -I \end{bmatrix} \tag{17}$$

$$\delta \boldsymbol{x} = [\delta \boldsymbol{\lambda} \;\; \delta \boldsymbol{u} \;\; \delta \boldsymbol{\tau} \;\; \delta \boldsymbol{q} \;\; \delta \boldsymbol{\sigma} \;\; \delta \boldsymbol{z} \;\; \delta \boldsymbol{q_r}]^T \tag{18}$$

$$\delta \boldsymbol{y} = [\delta \boldsymbol{w} \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; \delta \boldsymbol{z_r} \;\; 0]^T \tag{19}$$

Appendix B: Grasp Quality Measure

Applying perturbations to initial equilibrium configuration produces new contact force distribution expressed as $\boldsymbol{\lambda} = \boldsymbol{\lambda_0} + \delta\boldsymbol{\lambda}$. Considering i-th contact, the contact forces $\boldsymbol{\lambda_i}$ requires to satisfy unilateral constraint and Coulomb friction constraint in Eq. 1 and Eq. 2 respectively. These constraints helps in avoiding slip and detachment of the contact during object grasps.

$$\boldsymbol{\lambda_{i,n}} \geq 0 \tag{1}$$

$$\sqrt{\boldsymbol{\lambda_{i,t}}^2 + \boldsymbol{\lambda_{i,o}}^2} \leq \mu_i \boldsymbol{\lambda_{i,n}} \tag{2}$$

Let us consider, $\boldsymbol{d}(\boldsymbol{\lambda}) = [\boldsymbol{d}_{1,c}, \boldsymbol{d}_{1,f}, \boldsymbol{d}_{1,\max}....., \boldsymbol{d}_{n_c,c}, \boldsymbol{d}_{n_c,f}, \boldsymbol{d}_{n_c,\max}]$ where $\boldsymbol{d}_{1,c}$ represents contact force in the normal direction, $\boldsymbol{d}_{1,f}$, distance of $\boldsymbol{\lambda_i}$ from the friction cone surface and $\boldsymbol{d}_{i,\max} = f_{i,\max} - \|\boldsymbol{\lambda_i}\|$, maximum force $f_{i,max}$ applied at the individual contact points.

In R^{3n}, the inequality $\|\delta\boldsymbol{\lambda}\| \leq \|\boldsymbol{d}(\boldsymbol{\lambda})\|_\infty$ expresses a sphere which is centered inside the equilibrium contact force. In order to prevent slip, the preceding inequality is a sufficient condition on the maximum euclidean norm of perturbations of the contact forces $\delta\boldsymbol{\lambda}$ in order to prevent slip at the contact points. This property is referred to as "contact robustness".

The limiting space of $\boldsymbol{d}(\boldsymbol{\lambda})$, described in the contact wrench space, must be mirrored in the space of external disturbances $\delta\boldsymbol{w}$ acting on the grasped object, to evaluate its contact robustness.

In the quasi-static condition, the grasping contact forces and the grasped object disturbance wrenches can be written as:

$$\delta\boldsymbol{\lambda} = -G_K^R \delta\boldsymbol{w} \tag{3}$$

The ellipsoid in the wrench space centered in zero and with major principal axes of length $2\left\|\boldsymbol{d}\right\|_\infty / \boldsymbol{\sigma_k}(G_K^R)$, is described by the relationship $\delta\boldsymbol{\lambda}^T\delta\boldsymbol{\lambda} = \delta\boldsymbol{w}^T {G_K^R}^T G_K^R \delta\boldsymbol{w} \leq \|d\boldsymbol{\lambda}\|_\infty^2$, which is derived from the preceding equation, Eq. 3. Despite the wrench disturbance, the inscribed sphere with radius, $\frac{\|\boldsymbol{d}(\lambda)\|}{\boldsymbol{\sigma_{max}}(G_K^R)}$, constitutes a restriction on the euclidean norm of $\delta\boldsymbol{w}$ ensuring that all contact requirements hold. As a result, in quasistatic condition a particular grasp may withstand any disturbance wrench $\delta\boldsymbol{w}$ without defying the constraints given by Eq. 1 and Eq. 2, i.e. to say without separation and slippage at any contact point provided that

$$\|\delta\boldsymbol{w}\| \leq \frac{\|\boldsymbol{d}(\boldsymbol{\lambda})\|}{\boldsymbol{\sigma_{max}}(G_K^R)} \tag{4}$$

The internal forces which are controllable belongs to the subspace $\mathcal{F}_h$ [31]

$$\mathcal{F}_h = \mathcal{R}(E) = \mathcal{N}(G) \cap (\mathcal{R}(KJS) + \mathcal{R}(KG^T)) \tag{5}$$

where matrix $E \in R^{n_\lambda \times h}$ represents a base of the subspace and $K = ({K_c}^{-1} + J{K_q}^{-1}J^T)^{-1}$. The controllable internal forces is expressed as $\delta\boldsymbol{\lambda_h} = E\boldsymbol{y}$, where $\boldsymbol{y} \in R^h$, generic vector.

Considering $\boldsymbol{d}_{\min}^{\mathcal{F}_h}$ is the minimum of the vector $\boldsymbol{d}(\boldsymbol{\lambda})$, which is calculated using controllable contact forces $\boldsymbol{\lambda} = G_K^R\boldsymbol{w} + E\boldsymbol{y}$, where $G_K^R = KG^T(GKG^T)^{-1}$. Sufficient condition for the maximum force constraint and friction constraints to be satisfied even during a contact force perturbation $\delta\boldsymbol{\lambda}$ is given as : $\|\delta\boldsymbol{\lambda}\| \leq d_{\min}^{\mathcal{F}_h}$

Considering the external wrench distribution we get

$$\|\delta\boldsymbol{w}\| \leq \frac{\boldsymbol{d}_{\min}^{\mathcal{F}_h}}{\boldsymbol{\sigma_{max}}(G_K^R)} \tag{6}$$

where, $\boldsymbol{\sigma_{max}}$: maximum singular value

Since $\boldsymbol{d}_{\min}^{\mathcal{F}_h}$ depends on $\boldsymbol{y}$, optimal contact force distribution is given by $\hat{\boldsymbol{\lambda}}_{opt} = G_K^R\boldsymbol{g} + E\hat{\boldsymbol{y}}_{opt}$ where $\hat{\boldsymbol{y}}_{\boldsymbol{opt}} = \arg\max(\boldsymbol{d}_{\min}^{\mathcal{F}_h}/\boldsymbol{\sigma_{max}}(G_K^R))$ [31].

The grasp quality measure, PCR (potential contact robustness) , which assumes friction constraints to be satisfied at all contact points is expressed as:

$$PCR = \max_{\boldsymbol{y}} \frac{\boldsymbol{d}_{\min}^{\mathcal{F}_h}}{\boldsymbol{\sigma_{max}}(G_K^R)} \tag{7}$$

On the other hand, grasp robustness measure called PGR (potential grasp robustness), shows that even if some contacts are lost a grasp can be stable

$$PGR = \max_{C_j} \max_{\boldsymbol{y}} \frac{\boldsymbol{d}^{\mathcal{F}_h}_{\min}(C_j)}{\boldsymbol{\sigma_{max}}(G^R_{K(C_j)})} \tag{8}$$

$$subject\ to \quad \mathcal{N}(K(C_j)G^T) = 0 \tag{9}$$

where C_j are 3 contact states (j=1,2,3) as described below:

- Contact State 1: Contact forces are exerted in all directions via the contact locations and therefore, $K_{ic} = diag(K_{itx}, K_{ity}, K_{itn})$, where K_{ic}: contact stiffness matrix; K_{itx} and K_{ity}: tangential stiffnesses; K_{itn}: normal stiffness.
- Contact State 2: Contact forces are exerted only in normal direction and $K_{ic} = K_{itn}$.
- Contact State 3: Contact is considered as detached and $K_{ic} = [\mathbf{0}]$.

Appendix C: Three Link Serial Manipulator

Dynamics of a finger are obtained as [137]:

$$B(\theta)\ddot{\theta} + C(\theta,\dot{\theta})\dot{\theta} + g(\theta) = \tau + \tau_{ext} \tag{1}$$

$$\begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix} \begin{bmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \\ \ddot{\theta}_3 \end{bmatrix} + \begin{bmatrix} C_1 \\ C_2 \\ C_3 \end{bmatrix} + \begin{bmatrix} G_1 \\ G_2 \\ G_3 \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix} + \begin{bmatrix} {\tau_{ext}}^1 \\ {\tau_{ext}}^2 \\ {\tau_{ext}}^3 \end{bmatrix}$$

$$\begin{aligned} B_{11} = {} & 2m_2L_1l_2\sin(\theta_1)\sin(\theta_1+\theta_2) + 2m_2L_1l_2\cos(\theta_1)\cos(\theta_1+\theta_2) \\ & + 2m_3L_1L_2\sin(\theta_1)\sin(\theta_1+\theta_2) + 2m_3L_1L_2\cos(\theta_1)\cos(\theta_1+\theta_2) \\ & + 2m_3L_1l_3\sin(\theta_1)\sin(\theta_1+\theta_2+\theta_3) + 2m_3L_1l_3\cos(\theta_1)\cos(\theta_1+\theta_2+\theta_3) \\ & + 2m_3L_2l_3\sin(\theta_1+\theta_2)\sin(\theta_1+\theta_2+\theta_3) + 2m_3L_2l_3\cos(\theta_1+\theta_2)\cos(\theta_1+\theta_2+\theta_3) \\ & + m_1{l_1}^2 + m_2{L_1}^2 + m_2{l_2}^2 + m_3{L_1}^2 + m_3{L_2}^2 + m_3{l_3}^2 + I_{zz1} + I_{zz2} + I_{zz3}, \end{aligned} \tag{2}$$

$$\begin{aligned} B_{12} = {} & m_2L_1l_2\sin(\theta_1)\sin(\theta_1+\theta_2) + m_2L_1l_2\cos(\theta_1)\cos(\theta_1+\theta_2) \\ & + 2m_3L_2l_3\sin(\theta_1+\theta_2)\sin(\theta_1+\theta_2+\theta_3) + 2m_3L_2l_3\cos(\theta_1+\theta_2)\cos(\theta_1+\theta_2+\theta_3) \\ & + m_3L_1L_2\sin(\theta_1)\sin(\theta_1+\theta_2) + m_3L_1L_2\cos(\theta_1)\cos(\theta_1+\theta_2) \\ & + m_3L_1l_3\sin(\theta_1)\sin(\theta_1+\theta_2+\theta_3) \\ & + m_3L_1l_3\cos(\theta_1)\cos(\theta_1+\theta_2+\theta_3) + m_2{l_2}^2 + m_3{L_2}^2 + m_3{l_3}^2 + I_{zz2} + I_{zz3} \end{aligned} \tag{3}$$

$$B_{13} = m_3 L_1 l_3 \sin(\theta_1)\sin(\theta_1+\theta_2+\theta_3) + m_3 L_1 l_3 \cos(\theta_1)\cos(\theta_1+\theta_2+\theta_3) + m_3 L_2 l_3 \sin(\theta_1+\theta_2)\sin(\theta_1+\theta_2+\theta_3) + m_3 L_2 l_3 \cos(\theta_1+\theta_2)\cos(\theta_1+\theta_2+\theta_3) + m_3 {l_3}^2 + I_{zz3} \tag{4}$$

$$B_{21} = B_{12} \tag{5}$$

$$B_{22} = 2m_3 L_2 l_3 \sin(\theta_1+\theta_2)\sin(\theta_1+\theta_2+\theta_3) + 2m_3 L_2 l_3 \cos(\theta_1+\theta_2)\cos(\theta_1+\theta_2+\theta_3) + m_2 {l_2}^2 + m_3 {L_2}^2 + m_3 {l_3}^2 + I_{zz2} + I_{zz3} \tag{6}$$

$$B_{23} = m_3 L_2 l_3 \sin(\theta_1+\theta_2)\sin(\theta_1+\theta_2+\theta_3) + m_3 L_2 l_3 \cos(\theta_1+\theta_2)\cos(\theta_1+\theta_2+\theta_3) + m_3 {l_3}^2 + I_{zz3} \tag{7}$$

$$B_{31} = B_{13}, B_{32} = B_{23} \tag{8}$$

$$B_{33} = m_3 {l_3}^2 + I_{zz3} \tag{9}$$

$$G_1 = g(m_1 l_1 \cos(\theta_1) + m_2 L_1 \cos(\theta_1) + m_3 L_1 \cos(\theta_1) + m_1 l_2 \cos(\theta_1+\theta_2) + m_3 L_2 \cos(\theta_1+\theta_2) + m_3 l_3 \cos(\theta_1+\theta_2+\theta_3)) \tag{10}$$

$$G_2 = g(m_2 l_2 \cos(\theta_1+\theta_2) + m_3 L_2 \cos(\theta_1+\theta_2) + m_3 l_3 \cos(\theta_1+\theta_2+\theta_3)) \tag{11}$$

$$G_3 = g(m_3 l_3 \cos(\theta_1+\theta_2+\theta_3)) \tag{12}$$

$$
\begin{aligned}
C_1 = {} & (2m_2L_1l_2\sin(\theta_1)\cos(\theta_1+\theta_2) - 2m_2L_1l_2\cos(\theta_1)\sin(\theta_1+\theta_2) \\
& + 2m_3L_1L_2\sin(\theta_1)\cos(\theta_1+\theta_2) - 2m_3L_1l_2\cos(\theta_1)\sin(\theta_1+\theta_2) \\
& + 2m_3L_1l_3\sin(\theta_1)\cos(\theta_1+\theta_2+\theta_3) - 2m_3L_1l_3\cos(\theta_1)\sin(\theta_1+\theta_2+\theta_3)) \\
& \times \left(\frac{\partial\theta_1}{\partial t}\right)\left(\frac{\partial\theta_2}{\partial t}\right) \\
& + 2m_3L_1l_3\sin(\theta_1)\cos(\theta_1+\theta_2+\theta_3) - 2m_3L_1l_3\cos(\theta_1)\sin(\theta_1+\theta_2+\theta_3) \\
& + 2m_3L_2l_3\sin(\theta_1+\theta_2)\cos(\theta_1+\theta_2+\theta_3) - 2m_3L_2l_3\cos(\theta_1+\theta_2)\sin(\theta_1+\theta_2+\theta_3) \\
& \times \left(\frac{\partial\theta_1}{\partial t}\right)\left(\frac{\partial\theta_3}{\partial t}\right) \\
& + 2m_3L_1l_3\sin(\theta_1)\cos(\theta_1+\theta_2+\theta_3) - 2m_3L_1l_3\cos(\theta_1)\sin(\theta_1+\theta_2+\theta_3) \\
& + 2m_3L_2l_3\sin(\theta_1+\theta_2)\cos(\theta_1+\theta_2+\theta_3) - 2m_3L_1l_3\cos(\theta_1+\theta_2)\sin(\theta_1+\theta_2+\theta_3) \\
& \times \left(\frac{\partial\theta_2}{\partial t}\right)\left(\frac{\partial\theta_3}{\partial t}\right) \\
& + (m_2L_1l_2\sin(\theta_1)\cos(\theta_1+\theta_2) - m_2L_1l_2\cos(\theta_1)\sin(\theta_1+\theta_2) \\
& + m_3L_1L_2\sin(\theta_1)\cos(\theta_1+\theta_2) - m_3L_1l_2\cos(\theta_1)\sin(\theta_1+\theta_2) \\
& + m_3L_1l_3\sin(\theta_1)\cos(\theta_1+\theta_2+\theta_3) - m_3L_1l_3\cos(\theta_1)\sin(\theta_1+\theta_2+\theta_3)) \\
& \times \left(\frac{\partial\theta_2}{\partial t}\right)\left(\frac{\partial\theta_2}{\partial t}\right) \\
& + (m_3L_1l_3\sin(\theta_1)\cos(\theta_1+\theta_2+\theta_3) - m_3L_1l_3\cos(\theta_1)\sin(\theta_1+\theta_2+\theta_3) \\
& + m_3L_2l_3\sin(\theta_1+\theta_2)\cos(\theta_1+\theta_2+\theta_3) - m_3L_1l_3\cos(\theta_1+\theta_2)\sin(\theta_1+\theta_2+\theta_3)) \\
& \times \left(\frac{\partial\theta_3}{\partial t}\right)\left(\frac{\partial\theta_3}{\partial t}\right)
\end{aligned}
\tag{13}
$$

$$
\begin{aligned}
C_2 = {} & (m_2 L_1 l_2 \sin(\theta_1)\cos(\theta_1+\theta_2) - m_2 L_1 l_2 \cos(\theta_1)\sin(\theta_1+\theta_2) \\
& + m_3 L_1 L_2 \sin(\theta_1)\cos(\theta_1+\theta_2) - m_3 L_1 l_2 \cos(\theta_1)\sin(\theta_1+\theta_2) \\
& + m_3 L_1 l_3 \sin(\theta_1)\cos(\theta_1+\theta_2+\theta_3) - m_3 L_1 l_3 \cos(\theta_1)\sin(\theta_1+\theta_2+\theta_3))\times \\
& \left(\frac{\partial\theta_1}{\partial t}\right)\left(\frac{\partial\theta_2}{\partial t}\right) \\
& + 2m_3 L_2 l_3 \sin(\theta_1+\theta_2)\cos(\theta_1+\theta_2+\theta_3) - 2m_3 L_2 l_3 \cos(\theta_1+\theta_2)\sin(\theta_1+\theta_2+\theta_3)\times \\
& \left(\frac{\partial\theta_1}{\partial t}\right)\left(\frac{\partial\theta_3}{\partial t}\right) \\
& + 2m_3 L_2 l_3 \sin(\theta_1+\theta_2)\cos(\theta_1+\theta_2+\theta_3) - 2m_3 L_2 l_3 \cos(\theta_1+\theta_2)\sin(\theta_1+\theta_2+\theta_3)\times \\
& \left(\frac{\partial\theta_2}{\partial t}\right)\left(\frac{\partial\theta_3}{\partial t}\right) \\
& + (-m_2 L_1 l_2 \sin(\theta_1)\cos(\theta_1+\theta_2) + m_2 L_1 l_2 \cos(\theta_1)\sin(\theta_1+\theta_2) \\
& - m_3 L_1 L_2 \sin(\theta_1)\cos(\theta_1+\theta_2) + m_3 L_1 l_2 \cos(\theta_1)\sin(\theta_1+\theta_2) \\
& - m_3 L_1 l_3 \sin(\theta_1)\cos(\theta_1+\theta_2+\theta_3) + m_3 L_1 l_3 \cos(\theta_1)\sin(\theta_1+\theta_2+\theta_3))\times \\
& \left(\frac{\partial\theta_1}{\partial t}\right)\left(\frac{\partial\theta_1}{\partial t}\right) \\
& + (m_3 L_2 l_3 \sin(\theta_1+\theta_2)\cos(\theta_1+\theta_2+\theta_3) - m_3 L_2 l_3 \cos(\theta_1+\theta_2)\sin(\theta_1+\theta_2+\theta_3))\times \\
& \left(\frac{\partial\theta_3}{\partial t}\right)\left(\frac{\partial\theta_3}{\partial t}\right)
\end{aligned}
\tag{14}
$$

$$
\begin{aligned}
C_3 = {} & (2m_3L_2l_3\cos(\theta_1+\theta_2)\sin(\theta_1+\theta_2+\theta_3) - 2m_3L_2l_3\sin(\theta_1+\theta_2)\cos(\theta_1+\theta_2+\theta_3)\times \\
& \left(\frac{\partial\theta_1}{\partial t}\right)\left(\frac{\partial\theta_2}{\partial t}\right) \\
& + (m_3L_1l_3\sin(\theta_1)\cos(\theta_1+\theta_2+\theta_3) - m_3L_1l_3\cos(\theta_1)\sin(\theta_1+\theta_2+\theta_3) \\
& + m_3L_2l_3\sin(\theta_1+\theta_2)\cos(\theta_1+\theta_2+\theta_3) - m_3L_2l_3\cos(\theta_1+\theta_2)\sin(\theta_1+\theta_2+\theta_3)\times \\
& \left(\frac{\partial\theta_1}{\partial t}\right)\left(\frac{\partial\theta_3}{\partial t}\right) \\
& + (m_3L_2l_3\sin(\theta_1+\theta_2)\cos(\theta_1+\theta_2+\theta_3) - m_3L_2l_3\cos(\theta_1+\theta_2)\sin(\theta_1+\theta_2+\theta_3)\times \\
& \left(\frac{\partial\theta_2}{\partial t}\right)\left(\frac{\partial\theta_3}{\partial t}\right) \\
& + (m_3L_1l_3\cos(\theta_1)\sin(\theta_1+\theta_2+\theta_3) - m_3L_1l_3\sin(\theta_1)\cos(\theta_1+\theta_2+\theta_3) \\
& + m_3L_2l_3\cos(\theta_1+\theta_2)\sin(\theta_1+\theta_2+\theta_3) - m_3L_2l_3\sin(\theta_1+\theta_2)\cos(\theta_1+\theta_2+\theta_3))\times \\
& \left(\frac{\partial\theta_1}{\partial t}\right)\left(\frac{\partial\theta_1}{\partial t}\right) \\
& + (m_3L_2l_3\cos(\theta_1+\theta_2)\sin(\theta_1+\theta_2+\theta_3) - m_3L_2l_3\sin(\theta_1+\theta_2)\cos(\theta_1+\theta_2+\theta_3))\times \\
& \left(\frac{\partial\theta_2}{\partial t}\right)\left(\frac{\partial\theta_2}{\partial t}\right)
\end{aligned}
\tag{15}
$$