

Humanoid Robotic Manipulation Benchmarking and Bimanual Manipulation Workspace Analysis

by

William Thibault

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Mechanical and Mechatronics Engineering

Waterloo, Ontario, Canada, 2022

© William Thibault 2022

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

This thesis contains content from two first-authored manuscripts that are currently under review at the 2022 IEEE-RAS International Conference on Humanoid Robots (Humanoids 2022) in Ginowan, Okinawa, Japan:

1. William Thibault, Francisco Javier Andrade Chavez, and Katja Mombaur. **A standardized benchmark for humanoid whole-body manipulation.** *Submitted to 2022 IEEE-RAS International Conference on Humanoid Robots (Humanoids 2022).*
2. William Thibault, Vidyasagar Rajendran, and Katja Mombaur. **Bimanual Manipulation Workspace Analysis of Humanoid Robots with Object Specific Coupling Constraints.** *Submitted to 2022 IEEE-RAS International Conference on Humanoid Robots (Humanoids 2022).*

For the first work, I was responsible for contributing to the design of the study, developing the code in simulation and on the robot to perform motions, carrying out the data analysis from the experiments, developing the final benchmarks (test bed, protocols and key performance indicators), and drafting and submitting manuscripts. My co-authors provided guidance throughout the research performed and feedback on the draft manuscript.

For the second work, I was responsible for equal contribution to the simulation design, code development, and drafting and submitting manuscripts with Vidyasagar Rajendran as we were equal co-first-authors on the paper. Dr. Katja Mombaur provided guidance throughout the research performed and feedback on the draft manuscript.

The content from these papers has been adapted and extended for this thesis.

Abstract

The growing adoption of robots for new applications has led to the use of robots in human environments for human-like tasks, applications well-suited to humanoid robots as they are designed to move like a human and operate in similar environments. However, a user must decide which robot and control algorithm is best suited to the task, motivating the need for standardized performance comparison through benchmarking. Typical humanoid robotic scenarios in many household and industrial tasks involve manipulation of objects with two hands, bimanual manipulation. Understanding how these can be performed in the humanoid’s workspace is especially challenging due to the highly constrained nature due to grasp and stability requirements, but very important for introducing humanoid robots into human environments for human-like tasks.

The first topic this thesis focuses on is benchmarking manipulation for humanoid robotics. The evaluation of humanoid manipulation can be considered for whole-body manipulation, manipulation while standing and remaining balanced, or loco-manipulation, taking steps during manipulation. As part of the EUROBENCH project, which aims to develop a unified benchmarking framework for robotic systems performing locomotion tasks, benchmarks for whole-body manipulation and loco-manipulation are proposed consisting of standardized test beds, comprehensive experimental protocols, and insightful key performance indicators. For each of these benchmarks, partial initial benchmarks are performed to begin evaluating the difference in performance of the University of Waterloo’s REEM-C, “Seven”, using two different motion generation and control strategies. These partial benchmarks showed trade-offs in speed and efficiency for placement accuracy.

The second topic of interest is bimanual manipulation workspace analysis of humanoid robots. To evaluate the ability of a humanoid robot to bimanually manipulate a box while remaining balanced, a new metric for combined manipulability-stability is developed based on the volume of the manipulability ellipsoid and the distance of the capture point from the edge of the support polygon. Using this metric, visualizations of the workspace are performed for the following scenarios: when the center of mass of the humanoid has a velocity, manipulating objects of different size and mass, and manipulating objects using various grips. To examine bimanual manipulation with different fixed grasps the manipulation of two different boxes, a broom and a rolling pin are visualized to see how grip affects the feasibility and manipulability-stability quality of a task. Visualizations of REEM-C and TALOS are also performed for a general workspace and a box manipulation task to compare their workspaces as they have different kinematic structures. These visualizations provide a better understanding of how manipulability and stability are impacted in a bimanual manipulation scenario.

Acknowledgements

I would like to thank Katja Mombaur for her support throughout my MASC degree and for providing me the mentoring and guidance to begin my research career in humanoid robotics. Her passion and enthusiasm for research has been truly inspiring and helped me develop my skills as a humanoids researcher as I never could have imagined.

Thank you to my readers, William Melek and Stewart McLachlin, for providing valuable feedback on this thesis and supporting my interest in research since my BASC degree.

I would like to thank Francisco Andrade Chavez, for his support throughout every stage of my research from the first paper we collaborated on to the continuation of my research career in humanoids. I look forward to more work together as we learn more about REEM-C and TALOS.

I would like to thank Jonathan Lin, for his endless knowledge about the workings of Waterloo as I navigated my MASC and his support with learning to use motion capture to enhance my humanoids research.

Thank you to Vidyasagar (Sagar) Rajendran for all the help you provided during my MASC, from debugging anything ROS or Docker related to our collaboration on course projects, the bimanual workspace analysis, and our REEM-C repair trip to Barcelona. I look forward to future collaborations with you.

Thank you to Kareem, Anas, Jan, Jiwon, and Menna for our many study sessions and course project collaborations. I had a great time working with you all and hope to continue our collaborations in the future.

Thank you Peter for your linux and machine vision insights along with your effort to pave the way for database use my work certainly needed.

I would also like to thank the entire Human-centred Robotics and Machine Intelligence lab of Katja Mombaur. It was such a pleasure to interact with and learn from everyone in the group and I look forward to more lab wide events to enjoy together.

Thank you to Robert Wagner and Alex Werner of the Robohub for all the PAL related assistance with both REEM-C and TALOS and the insight into the practical issues involved with humanoid systems.

I would like to extend a special thanks to UWaterloo and the NSERC CERC program for the funding and scholarships provided to make my research possible.

Finally, I would like to thank my family for all the encouragement and support. Hannah, my sister; Mary, my mother; Peter, my father; and Wallace, my sister's dog; thank you for believing in me and making all this possible.

Dedication

I would like to dedicate this thesis to my Chocolate Labrador Retriever, Pecan. Whether it was joining my video calls, participating in long coding sessions, or going on our daily walks where our best ideas came from, you are the perfect research partner. Your support and love every step of the way made this possible.



Table of Contents

List of Figures	x
List of Tables	xiii
1 Introduction	1
1.1 Problem Definition	2
1.2 Thesis Contributions	3
1.2.1 Benchmarking Loco-manipulation and Whole-body Manipulation	3
1.2.2 Bimanual Manipulation Workspace Analysis	3
1.3 Thesis Organization	4
2 Literature Review	7
2.1 Robotic Manipulation	7
2.1.1 Kinematics	7
2.1.2 Manipulability	9
2.1.3 Planning	10
2.2 Humanoid Control and Motion Generation	12
2.2.1 Stability	12
2.2.2 Walking	14
2.2.3 Whole-body Control	16
2.3 Benchmarking	18

2.3.1	Locomotion Benchmarking	19
2.3.2	Manipulation Benchmarking	19
2.3.3	Loco-manipulation and Whole-body Motion Benchmarking	20
2.3.4	EUROBENCH Benchmarking Framework	21
2.4	Bimanual Manipulation	21
2.5	Summary	25
3	Whole-body Manipulation Benchmarking	27
3.1	Whole-body Manipulation Benchmark	27
3.1.1	Test Bed	28
3.1.2	Protocols	29
3.1.3	Key Performance Indicators	32
3.2	Partial Benchmark Performed	36
3.2.1	Experimental Setup	36
3.2.2	Performance Metrics	37
3.2.3	The Humanoid Robot “Seven”	37
3.2.4	Motion Generation and Control Methods	39
3.2.5	Experimental Results & Discussion	43
4	Loco-manipulation Benchmarking	47
4.1	Pick and Carry Benchmark	48
4.1.1	Test Bed	48
4.1.2	Protocols	49
4.1.3	Key Performance Indicators	51
4.2	Partial Benchmark Performed	52
4.2.1	Experimental Setup	53
4.2.2	Performance Metrics	53
4.2.3	Motion Generation and Control Methods	54
4.2.4	Experimental Results & Discussion	56

5	Bimanual Manipulation Workspace Analysis	57
5.1	Workspace Metrics	57
5.1.1	Manipulability	58
5.1.2	Stability	58
5.1.3	Combined Manipulability-Stability Metric	59
5.2	Workspace Generation and Visualization	60
5.2.1	Whole-body Inverse Kinematics	60
5.2.2	General Workspace Generation	61
5.2.3	Workspace Generation with Objects	62
5.2.4	Workspace Visualization	64
5.3	Results	65
5.3.1	General Workspace Maps	65
5.3.2	Workspace Maps with Varying Center of Mass Velocities	66
5.3.3	Workspace Maps with Different Object Masses	69
5.3.4	Workspace Maps with Different Objects	70
5.3.5	Workspace Maps of TALOS	72
5.3.6	Applications of the Combined Manipulability Metric and Workspace Visualization	74
6	Thesis Conclusions	75
6.1	Summary of Findings	75
6.2	Limitations	76
6.3	Future Work	77
	References	78

List of Figures

1.1	EUROBENCH benchmarking work flow with blue highlighting the work in this thesis.	4
1.2	Objects with coordinated, tightly coupled bimanual manipulation scenarios: a) a small 1 <i>kg</i> box b) a large 7 <i>kg</i> box c) a sweeping broom (with a left-handed grasp) d) a rolling pin.	6
2.1	Two adjacent links of a robotic manipulator with revolute joints using D-H parameters. Adapted from Lynch and Park 2019 [28, app. C.1].	8
2.2	The 3 dimensional linear inverted pendulum model used to calculate the capture point. Note that the CoM remains at a constant height indicated by the light red plane.	14
2.3	A 3D LIPM for a CoM at a constrained height. Adapted from Kajita et al. 2014 [22, ch. 4.3.1].	15
2.4	A cart of mass M rolling on a mass-less table. Adapted from Kajita et al. 2014 [22, ch. 4.4.1].	16
2.5	ZMP tracking control with a servo controller. Adapted from Kajita et al. 2014 [22, ch. 4.4.3].	16
2.6	The Koroibot locomotion tasks. Adapted from [2].	19
2.7	The outcomes of the EUROBENCH benchmarking project. Adapted from [1].	22
2.8	The partners and parties involved in the EUROBENCH project. Adapted from [1].	23
2.9	The bimanual manipulation workspace of ARMAR-III using the extended manipulability metric. Adapted from Vahrenkamp et al. 2014 [43].	24

2.10	The stability metric visualized at feet for ARMAR-4 with extended manipulability metric visualized for left hand. Adapted from Kaiser et al. 2015 [20].	25
3.1	EUROBENCH benchmarking work flow with blue highlighting the whole-body manipulation work in this chapter.	28
3.2	Standard whole-body manipulation benchmarking box.	29
3.3	Whole-body manipulation test bed model.	30
3.4	A box foot print, in red, misplaced on the shelf with respect to the target.	34
3.5	EUROBENCH benchmarking performance metric work flow where the protocol is performed and data collected in orange. Green highlights the data converted to the EUROBENCH format and red indicates the performance indicator calculation workflow, with Docker use for the metric calculation is implemented (dark red). The output to the user is shown in blue. Adapted from Aller et al. 2022 [3].	36
3.6	Partial whole-body manipulation benchmarking set-up.	38
3.7	PAL Robotics walking controller. Adapted from Baños 2020 [5].	40
3.8	Separate upper and lower body control method.	41
3.9	Separate upper and lower body controllers method motion sequence.	41
3.10	Whole-body control method.	42
3.11	Whole-body controller method motion sequence.	43
3.12	REEM-C grasping milk crate handle.	46
4.1	EUROBENCH benchmarking work flow with blue highlighting the loco-manipulation benchmarking work in this chapter.	48
4.2	Pick and carry test bed model.	50
4.3	Partial pick and carry benchmarking set-up.	53
4.4	Separate upper and lower body controllers method pick and carry motion sequence ordered across then down.	55
4.5	ArUco markers on shelving unit as seen by the robot.	56
5.1	A simplified support polygon showing the capture point and the selected minimum distance for the stability metric, d_f	59

5.2	a) The REEM-C humanoid with <i>World</i> , left and right TCP frames with x in red, y in green and z in blue. b) A single point on the workspace grid indicates the center between the two end effectors of the robot for general workspace generation.	61
5.3	Workspace visualizations for the general generation method. Low quality voxels are blue and high quality voxels are yellow/orange. Voxels that are unreachable or unstable are translucent grey. The first row shows the entire workspace that was considered and the second row shows a horizontal cut through the workspace. a) and d) show the manipulability metric, b) and e) show the stability metric and c) and f) show the combined manipulability-stability metric as detailed in Equation 5.5.	66
5.4	Various CoM velocity cases visualized for a subset of the general workspace with the stability metric. a) $\dot{r}_{CoM} = [0.0, 0.0, 0.0]$ b) $\dot{r}_{CoM} = [0.0, 0.2, 0.0]$ box c) $\dot{r}_{CoM} = [0.0, 0.6, 0.0]$ d) $\dot{r}_{CoM} = [0.2, 0.6, 0.0]$	68
5.5	Visualization of the bimanual manipulability-stability metric for the REEM-C manipulating a 1 <i>kg</i> box. a) A slice at $z = 1.0$ m b) A slice at $z = 1.2$ m.	69
5.6	Visualization of the bimanual manipulability-stability metric for the REEM-C manipulating a 7 <i>kg</i> box. a) A slice at $z = 1.0$ m b) A slice at $z = 1.2$ m.	70
5.7	Visualization of the bimanual manipulability-stability distribution for REEM-C manipulating a rolling pin. a) A cut through at $z = 0.9$ m b) A cut through at $z = 1.1$ m	71
5.8	Visualization of the bimanual manipulability-stability distribution for REEM-C manipulating a broom. a) cut through at $z = 0.9$ m b) cut through at $z = 1.0$ m	72
5.9	Visualization of the bimanual manipulability-stability metric with TALOS, for the general workspace generation with spherical coordinates. a) A slice at $z = 0.9$ m c) A slice at $z = 1.1$ m.	73
5.10	Visualization of the bimanual manipulability-stability metric for the TALOS manipulating a 1 <i>kg</i> box. a) A slice at $z = 1.0$ m b) A slice at $z = 1.2$ m.	73

List of Tables

3.1	Key performance indicators for the whole-body manipulation benchmark. .	33
3.2	Key performance indicators for the partial whole-body manipulation benchmark.	39
3.3	Results from 10 trials of two motion generation approaches for the partial whole-body manipulation benchmark	44
4.1	Key performance indicators for the pick and carry benchmark.	52
4.2	Key performance indicators for the partial pick and carry benchmark. . . .	54
5.1	EXOTica task maps in whole body inverse kinematics formulation.	60

Chapter 1

Introduction

To extend the abilities of robots to perform human-like tasks they need to be integrated into human designed environments for everyday use. Humanoid robots, being designed to replicate a human's size and motions, offer great possibilities for operating in environments that only humans can typically use. Motivated by disasters like the Fukushima Daiichi nuclear disaster and the Deepwater Horizon oil spill, the DARPA Robotics Challenge offered a glimpse into the applications of humanoid robots to operate in disaster relief situations that require human intervention but are a danger to humans [10]. The tasks involved a variety of locomotion and manipulation like traversing rough terrain, using tools, closing a valve, and entering a door, which at their core involve skills that could be used in daily scenarios. While work exists outlining humanoid robot applications in simultaneous locomotion and manipulation tasks, also known as loco-manipulation tasks, operating in everyday human environments, and performing human tasks with similar performance and dexterity is still a challenging problem.

With the advent of humanoid robots being able to operate in human-like environments and perform tasks typically performed by humans, users need to decide which robot is suitable for their application. The topic of benchmarking is useful at evaluating the performance of a robot and helping users determine if the needs for their application will be met. This evaluation can also extend to motion generation and techniques used. Most work toward benchmarking robotics has involved international robotics competitions like the previously mentioned DARPA Robotics Challenge, Cybathlon, RoboCup, RoCKIn, the European Robotics League and the Amazon Robotics Challenge. While these competitions do compare the performance of the robots, the evaluation is focused largely on qualitative performance (described quality and completion) rather than rigorous evaluation of quantitative (measured and calculated quantities) metrics. Aller et al. [4] performed a literature

review on the state-of-the-art evaluation of robot performance in a number of scenarios with a focus on balancing and locomotion, which identified deficiencies in previous benchmarking efforts. There remains a lack of well-defined benchmarking scenarios with specific protocols and quantifiable key performance indicators. As a result, the European project, EUROBENCH, was developed to create a unified benchmarking framework for bipedal locomotion in robotic systems [1]. This project created two testing facilities hosting test beds for benchmarking wearable and humanoid robotics as well as a software system for computing experimental and simulated performance.

When bringing humanoid robots into the real world to perform daily, human-like tasks, the need for bimanual or dual-arm abilities becomes essential for complex human actions. Humans can easily manipulate objects in a variety of shapes, sizes and weights, something that bimanual manipulation with humanoids could offer. For example, consider the bimanual tasks identified by Krebs et al. [23] in the KIT Bimanual Manipulation Dataset that contains human motion capture for motions like sweeping with a broom, opening juice, mixing cooking ingredients and using a rolling pin. Other human bimanual tasks include lifting boxes, pushing carts, and carrying trays. Humanoid robots are well-suited for these tasks as they can adapt to working in human settings such as homes or factories and offer more motion possibilities given the ability to walk with legs and perform dual-arm tasks. However, there are several challenges for developing bimanual task motions due to the high degree of freedom, coordinated, multi-contact nature.

1.1 Problem Definition

This thesis identifies two main problems that are addressed.

Problem 1: Benchmarking loco-manipulation and whole-body manipulation

The first problem is to develop a benchmark for loco-manipulation and whole-body manipulation scenarios with a focus on humanoid robots. In this thesis, *loco-manipulation* will be used to reference motions that involve manipulation of an object with the arms of a humanoid and performing locomotion over large distances that require many steps. *Whole-body manipulation* will refer to motions involving manipulation of an object with the arms of a humanoid, but with little to no locomotion performed.

Problem 2: Bimanual manipulation workspace analysis

The second problem arises in relation to the first, due to the difficulties bimanual tasks present in many loco-manipulation and whole-body manipulation scenarios. This problem is to evaluate the bimanual workspace of humanoid robots to better understand the complex manipulation tasks involved in coordinated, tightly coupled bimanual manipulation motions for objects with symmetric and asymmetric grasps.

1.2 Thesis Contributions

1.2.1 Benchmarking Loco-manipulation and Whole-body Manipulation

This thesis focuses on benchmarking whole-body manipulation and loco-manipulation scenarios. The creation of these benchmarks is in collaboration with EUROBENCH, the European benchmarking project for bipedal locomotion in robots, that aims to develop a unified benchmarking framework for wearable robotics and humanoid robotics along with a test facility for both robotics fields available to academia and industry [1]. Benchmarks for whole-body manipulation and loco-manipulation are proposed for the EUROBENCH framework for scenarios that are representative of many warehouse and logistics scenarios with similar tasks. The whole-body manipulation benchmark considers the manipulation of a box on a set of shelving units from one shelf to another. The pick and carry benchmark considers the transport of a box between two shelving units. The benchmarks consist of the design of a test bed along with the definition of specific test protocols and meaningful key performance indicators (KPIs). To develop a better understanding of the benchmarks, partial whole-body manipulation and loco-manipulation benchmarks with the humanoid robot REEM-C are performed. These partial benchmarks evaluate a subset of KPIs through the comparison of two different motion generation approaches. [Figure 1.1](#) shows the EUROBENCH benchmarking flow in the context of the whole-body manipulation and loco-manipulation benchmarks.

1.2.2 Bimanual Manipulation Workspace Analysis

This thesis contributes to the existing work on bimanual manipulation workspace analysis and the existing gaps related to humanoid robots. The analysis considers the manipulability and stability of a humanoid robot in a fixed bimanual grasping scenario with object

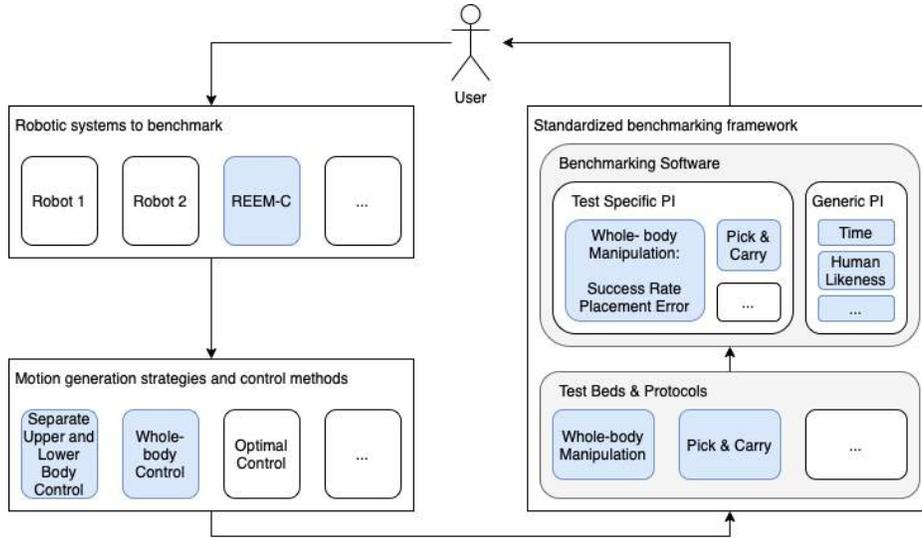


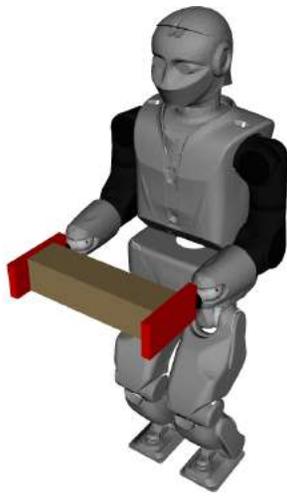
Figure 1.1: EUROBENCH benchmarking work flow with blue highlighting the work in this thesis.

specific considerations like constraints on the hands, object weight and collision avoidance. Stability considerations extend beyond static stability. Thus the contributions are a new bimanual manipulability-stability metric for bimanual tasks under static and dynamic stability cases; workspace analyses using object specific considerations like object mass, collisions and end effector constraints; the generation and visualizations of the new metric for workspaces using various objects (see [Figure 1.2](#)) and center of mass velocities; a workspace comparison for REEM-C and TALOS, two robots that vary significantly in kinematics and dynamics. This contribution is the result of an equal collaboration with Vidyasagar Rajendran.

1.3 Thesis Organization

[Chapter 2](#) covers the background of robotic manipulation and humanoid control and motion generation then presents related work on benchmarking in robotics and bimanual manipulation. [Chapter 3](#) proposes a whole-body manipulation benchmark consisting of test beds, protocols and KPIs then performs and evaluates a partial version for different motion generation and control strategies using the REEM-C. In a similar manner, [Chapter 4](#) proposes a loco-manipulation benchmark then performs a partial version of it for

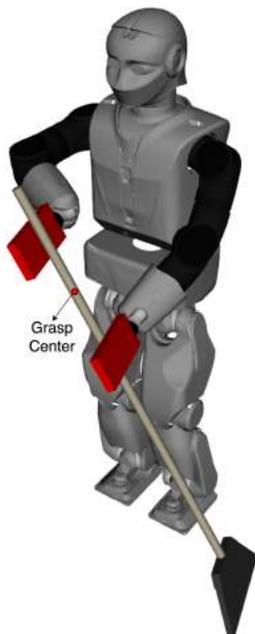
another set of motion generation and control strategies using the REEM-C. [Chapter 5](#) develops the bimanual manipulation workspace and generates visualizations for the various cases, which are analyzed and compared. [Chapter 6](#) presents conclusions and discusses future work.



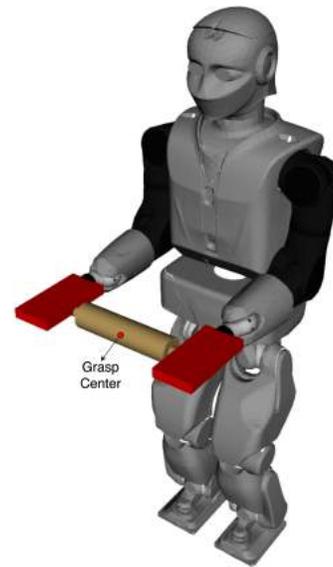
(a)



(b)



(c)



(d)

Figure 1.2: Objects with coordinated, tightly coupled bimanual manipulation scenarios: a) a small 1 *kg* box b) a large 7 *kg* box c) a sweeping broom (with a left-handed grasp) d) a rolling pin.

Chapter 2

Literature Review

In this literature review, it first reviews robotic manipulation then humanoid control and motion generation. Next, it covers previous work on benchmarking in robotics and bimanual manipulation.

2.1 Robotic Manipulation

This section provides the basic background for robotic manipulation. The coverage of the following topics will be limited to open-chain robotic manipulators with single degree of freedom (DoF) revolute joints [28, ch. 2.2]. Open-chain manipulators do not include closed loops. Revolute joints are joints that allow for rotational motion about a single axis.

2.1.1 Kinematics

From Lynch and Park [28, app. C], determining the position of the end effector based on the joint angles is a common problem for a robotic manipulator of n DoF. This problem is referred to as forward kinematics. Given the robotics frames in terms of Denavit-Hartenberg (D-H) convention, a popular convention used in robotics as it requires only four parameters to describe the frames at each joint of a robotics chain, the transformation matrix, $T_{i-1,i}$, from frame $i - 1$ to frame i can be written as follows for the robotic chain seen in Figure

2.1:

$$T_{i-1,i} = \begin{bmatrix} \cos\phi_i & -\sin\phi_i & 0 & a_{i-1} \\ \sin\phi_i\cos\alpha_{i-1} & \cos\phi_i\cos\alpha_{i-1} & -\sin\alpha_{i-1} & -d_i\sin\alpha_{i-1} \\ \sin\phi_i\sin\alpha_{i-1} & \cos\phi_i\sin\alpha_{i-1} & \cos\alpha_{i-1} & d_i\cos\alpha_{i-1} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.1)$$

$$= \begin{bmatrix} R_{i-1,i} & x_{i-1,i} \\ [0,0,0] & 1 \end{bmatrix} \quad (2.2)$$

where $R_{i-1,i}$ is the rotation matrix from frame $i-1$ to frame i and $x_{i-1,i}$ is the vector from frame $i-1$ to frame i . To compute the forward kinematics of a robotic manipulator with n joints, the transformation matrices for all the links are multiplied sequentially as follows:

$$T_{0,n}(\theta_1, \dots, \theta_n) = T_{0,1}(\theta_1)T_{1,2}(\theta_2)\dots T_{n-1,n}(\theta_n) \quad (2.3)$$

where θ_i is the joint angle of the i th joint.

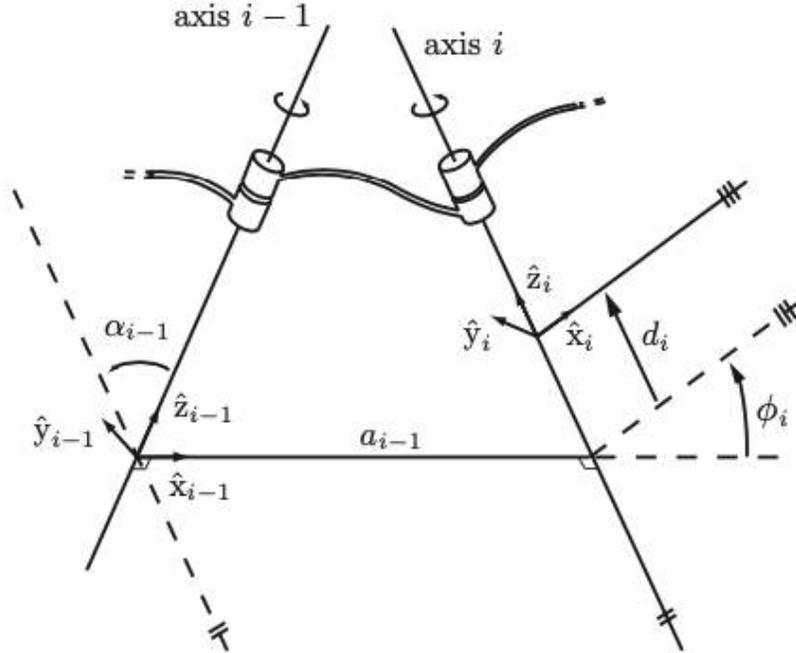


Figure 2.1: Two adjacent links of a robotic manipulator with revolute joints using D-H parameters. Adapted from Lynch and Park 2019 [28, app. C.1].

Another important set of parameters described by Lynch and Park is the Jacobian, which gives the sensitivity of the end effector velocity to the joint velocity of a robotic manipulator [28, ch. 5]. When considering forward kinematics for joint angles $\theta \in \mathbb{R}^n$ for an end effector in coordinates $x \in \mathbb{R}^m$, the forward kinematics are:

$$x(t) = f(\theta(t)) \quad (2.4)$$

In turn, the end effector velocity can be written as:

$$\dot{x} = \frac{\partial f(\theta)}{\partial \theta} \frac{\partial \theta(t)}{\partial t} \quad (2.5)$$

$$= J(\theta)\dot{\theta} \quad (2.6)$$

where $J(\theta) \in \mathbb{R}^{m \times n}$ is the Jacobian.

Another problem in robotic manipulator kinematics is when the joint angles are desired for a specific end effector position, called inverse kinematics. This problem can be solved a variety of ways as detailed by Lynch and Park [28, ch. 6]. In simple cases with few joints, the joint angles can be solved analytically using trigonometric methods, but in more complex, non-analytic cases it can be solved numerically. The numerical methods, such as the Newton-Raphson method, are iterative and in the case where multiple solutions exist it will find the solution nearest to the initial guess used.

2.1.2 Manipulability

On the topic of robotic manipulation, manipulability is another key concept with the development below following that of Lynch and Park [28, ch. 5.4]. Consider a robotic manipulator of n degrees of freedom in a task space of $q \in \mathbb{R}^m$ with $m \leq n$ and Jacobian J . The manipulability ellipsoid can be defined for the end effector velocities for a unit sphere in the joint velocity space with joint velocities $\dot{\theta}$:

$$\dot{\theta}^T \dot{\theta} = 1 \quad (2.7)$$

$$(J^{-1}\dot{q})^T (J^{-1}\dot{q}) = 1 \quad (2.8)$$

$$\dot{q}^T J^{-T} J^{-1} \dot{q} = 1 \quad (2.9)$$

$$\dot{q}^T (J J^T)^{-1} \dot{q} = 1 \quad (2.10)$$

$$\dot{q}^T (A)^{-1} \dot{q} = 1 \quad (2.11)$$

The square, symmetric and positive definite matrix, $A = J J^T \in \mathbb{R}^{m \times m}$ is a result of this development, where Equation 2.11 represents a m dimensional manipulability ellipsoid.

For matrix A the eigenvalues are λ_i , with the principal semi-axis lengths being $\sigma_i = \sqrt{\lambda_i}$. σ_i are also considered the singular values. With these values a number of manipulability metrics can be defined:

$$\mu_1(A) = \sqrt{\frac{\lambda_{max}(A)}{\lambda_{min}(A)}} = \frac{\sigma_{max}(A)}{\sigma_{min}(A)} \geq 1 \quad (2.12)$$

$$\mu_2(A) = \frac{\lambda_{max}(A)}{\lambda_{min}(A)} = \frac{\sigma_{max}^2(A)}{\sigma_{min}^2(A)} \geq 1 \quad (2.13)$$

$$\mu_3(A) = \sqrt{\lambda_1 \lambda_2 \dots \lambda_m} = \sigma_1 \sigma_2 \dots \sigma_m = \sqrt{\det A} = \sqrt{\det JJ^T} \quad (2.14)$$

[Equation 2.12](#) is the ratio of the longest and shortest semi-axis lengths and [Equation 2.13](#) is the condition number, both of which use values closer to 1 to indicate high manipulabilities. [Equation 2.14](#) is proportional to the volume of the manipulability ellipsoid, but indicates high manipulability with larger values.

2.1.3 Planning

Motion planning is a key part of developing feasible motions for a robot manipulator, such as moving a robot's end effector from one point in the workspace to another during a pick and place motion. For the motion to be feasible it must obey the limits of the robot, like joint and torque limits, and avoid collisions, both with the environment and the robot itself. A variety of motion planning techniques exist including sampling methods, grid methods and virtual potential fields [28, ch. 10]. The focus here will be on sampling based methods as they make up many of the basics for the Open Motion Planning Library [41], which includes sampling-based planning algorithms that are fast and useful for high DoF robots.

From Lynch and Park, a rapidly exploring random tree (RRT) develops a tree like structure for a single query planning method in the configuration space, C-space, or state space [28, ch. 10]. The RRT algorithm uses a uniform distribution to randomly sample from the state of the robot, \mathcal{X} , and connects the sample to the nearest collision free node in the tree. This allows the free space, \mathcal{X}_{free} , to be explored for an initial state of x_{start} to a goal set, \mathcal{X}_{goal} , where the growth starts from x_{start} . [Algorithm 1](#) shows the basic form of the RRT algorithm, though other variations exist such as the bidirectional RRT or RRT*. Note that the local planner mentioned in line 5 could be as simple as a straight line planner.

Algorithm 1 RRT algorithm adapted from Lynch and Park 2019 [28, ch. 10]

- 1: initialize search tree \mathcal{T} with x_{start}
- 2: **while** \mathcal{T} is less than the maximum tree size **do**
- 3: $x_{samp} \rightarrow$ sample from \mathcal{X}
- 4: $x_{nearest} \rightarrow$ nearest node in \mathcal{T} to x_{samp}
- 5: employ a local planner to find a motion from $x_{nearest}$ to x_{new} in the direction of x_{samp}
- 6: **if** the motion is collision free **then**
- 7: add x_{new} to \mathcal{T} with an edge from $x_{nearest}$ to x_{new}
- 8: $N \leftarrow \frac{N}{2}$
- 9: **if** x_{new} is in \mathcal{X}_{goal} **then**
- 10: Return SUCCESS and the motion to x_{new}
- 11: Return FAILURE

Another sampling-based planner described by Lynch and Park is the probabilistic roadmap (PRM), which develops a graph for multi-query planning in C-space [28, ch. 10]. This method builds a roadmap in the form of an undirected graph to represent \mathcal{C}_{free} , where the edges are ways the robot can move. Once the roadmap exists the q_{start} and q_{goal} are added to the roadmap by trying to connect them to the nearest nodes. This roadmap can then be searched for a path using a planner such as A*. Algorithm 2 shows how a roadmap is built for N nodes.

Algorithm 2 PRM algorithm adapted from Lynch and Park 2019 [28, ch. 10]

- 1: **for** $i = 1, \dots, N$ **do**
- 2: $q_i \rightarrow$ sample from \mathcal{C}_{free}
- 3: add q_i to R
- 4: **for** $i = 1, \dots, N$ **do**
- 5: $\mathcal{N}(q_i) \rightarrow k$ closest neighbors of q_i
- 6: **for** each $q \in \mathcal{N}(q_i)$ **do**
- 7: **if** there is a collision-free local path from q to q_i and there is not already an edge from q to q_i to the roadmap R **then**
- 8: add an edge from q to q_i to the roadmap R
- 9: Return R

2.2 Humanoid Control and Motion Generation

This section provides the basic background for humanoid robotic control and motion. These robots take the form of tree-like kinematic structures as the legs and arms are all attached to the same base structure (the torso and pelvis). Normally the limbs of these robots are composed of revolute joints making the arms and legs behave very similar to the robotic manipulators previously discussed.

2.2.1 Stability

Several important stability related points exist for humanoid robots. Here, the basics of these stability points are covered for later use in motion development.

Center of Mass and Ground Projected Center of Mass

The center of mass (CoM) is a popular point in humanoid robotics as it can be useful when considering simplified systems like inverted pendulum models. For a humanoid robot, with N masses, m_i , at points, p_i , the center of mass can be calculated as:

$$c = \sum_{i=1}^N \frac{m_i p_i}{M} \quad (2.15)$$

where M is the total mass of the robot [22, ch. 3].

In some cases, the center of mass is projected onto the ground plane. This point is called the Ground Projected Center of Mass (GPCM) and is simply the projection of the CoM, c , using projection matrix, $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, calculated as follows:

$$GPCM = P \sum_{i=1}^N \frac{m_i p_i}{M} \quad (2.16)$$

producing a 2D point.

Zero Moment Point

Another key point for humanoid robots is the zero-moment point (ZMP). As described by Kajita et al., the ZMP is a 2D point on the ground plane that always remains inside the support polygon, the convex hull created by the contacts of the feet with the ground, of the robot when stable [22, ch. 3]. The ZMP, (p_x, p_y) , is calculated as follows:

$$p_x = \frac{Mgx + p_z \dot{P}_x - \dot{L}_y}{Mg + \dot{P}_z} \quad (2.17)$$

$$p_y = \frac{Mgy + p_z \dot{P}_y - \dot{L}_x}{Mg + \dot{P}_z} \quad (2.18)$$

In Equation 2.17 and 2.18 M is the mass, g is gravity, p_z is the height of the ground (zero for flat ground), \dot{P} is the change in linear momentum and \dot{L} is the change in angular momentum. Should the robot be standing still, then both momentum terms are zero and the ZMP simplifies to:

$$p_x = x \quad (2.19)$$

$$p_y = y \quad (2.20)$$

which is equal to the center of pressure of the robot.

Capture Point

Pratt et al. defined the capture point as a 2D point on the ground plane that can be used to determine the step size to be taken for push recovery when the center of mass is moving at a velocity, based on the orbital energy of the linear inverted pendulum model (LIPM) [35]. In the case of a 3 dimensional LIPM with position r_{CoM} and velocity of \dot{r}_{CoM} , the instantaneous capture point, r_{ICP} , is calculated as shown below:

$$r_{ICP} = P(r_{CoM} + \frac{\dot{r}_{CoM}}{\omega_0}) \quad (2.21)$$

where the projection matrix $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ and $\omega_0 = \sqrt{\frac{g}{z_0}}$ for an acceleration of gravity of g and a constant center of mass height of z_0 [36]. An important note is that static stability is preserved if the center of mass is not moving. The important quantities for the capture point can be seen in Figure 2.2, which shows the 3 dimensional linear inverted pendulum model.

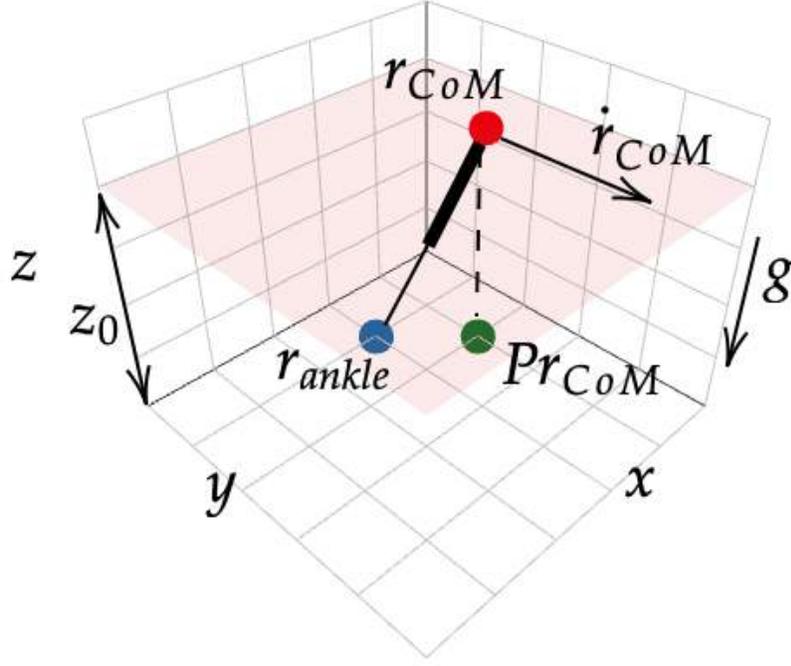


Figure 2.2: The 3 dimensional linear inverted pendulum model used to calculate the capture point. Note that the CoM remains at a constant height indicated by the light red plane.

2.2.2 Walking

Given that humanoid robots have legs, a natural problem is how to make the robot walk. A famous method for walking is the ZMP based walking pattern generator and preview control developed by Kajita et al. [21], typically referred to as Kajita walking, and is described as follows. Considering the 3D LIPM, as seen in Figure 2.3, the governing equations in terms of the CoM position, (x, y) , and the ZMP, (p_x, p_y) , can be written as:

$$\ddot{x} = \frac{g}{z_c}(x - p_x) \quad (2.22)$$

$$\ddot{y} = \frac{g}{z_c}(y - p_y) \quad (2.23)$$

where z_c is the height of the CoM in the constrained plane and g is the acceleration of gravity.

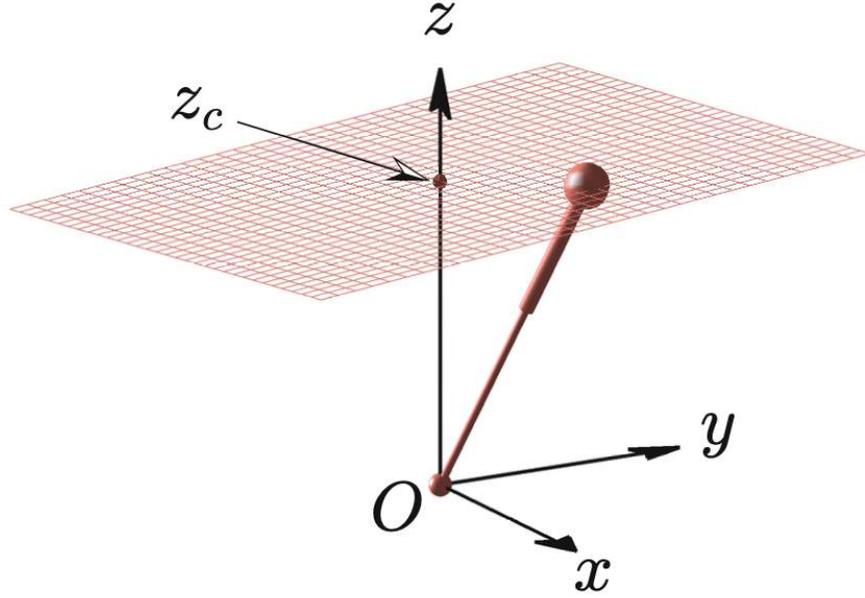


Figure 2.3: A 3D LIPM for a CoM at a constrained height. Adapted from Kajita et al. 2014 [22, ch. 4.3.1].

The walking pattern generator is based on the cart-table model, as seen in Figure 2.4, allowing for the following ZMP calculation in the 3D case for the simplified model as:

$$p_x = x - \frac{z_c}{g} \ddot{x} \quad (2.24)$$

$$p_y = y - \frac{z_c}{g} \ddot{y} \quad (2.25)$$

which can be noticed to be a rewritten version of Equation 2.22 and 2.23 with the cart position as (x, y) . The ZMP can be easily calculated from the motion of the center of mass trajectory; however, for the pattern generator the cart trajectory should be calculated from the ZMP to ensure the ZMP satisfies the stepping pattern prescribed. This leads to the ZMP control problem where the CoM trajectory is generated to calculate the ZMP with a reference ZMP input. To achieve this, a preview controller is used for the servo controller, to ensure the CoM appropriately moves for upcoming steps in the pattern. The control scheme can be seen in Figure 2.5.

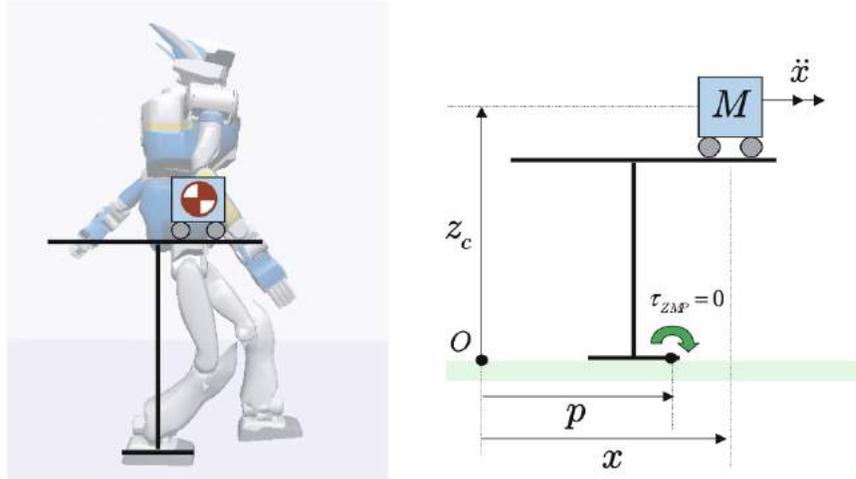


Figure 2.4: A cart of mass M rolling on a mass-less table. Adapted from Kajita et al. 2014 [22, ch. 4.4.1].

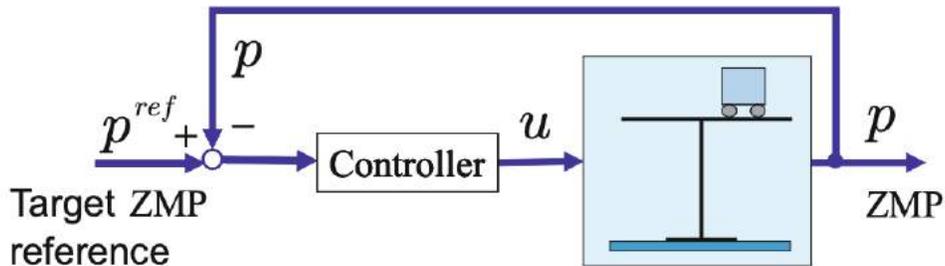


Figure 2.5: ZMP tracking control with a servo controller. Adapted from Kajita et al. 2014 [22, ch. 4.4.3].

2.2.3 Whole-body Control

A popular control method for humanoids due to the highly redundant, floating base, tree-like structure is whole-body control (WBC). This control method allows for all the joints of the robot to be leveraged to coordinate a motion according to a given set of constraints and tasks. Several WBC methods exist and can be classified as either velocity-based or torque-based, where velocity-based controllers give joint velocities and the torque-based give joint torques, and closed-form or optimization-based, where closed-form can be written

and computed algebraically and optimization-based require a solver for the optimization problem described [32]. This discussion is limited to velocity-based WBC and treats both closed-form and optimization-based WBC.

When using closed form techniques for WBC, following the development of Moro and Sentis [32], k control actions denoted as velocities \dot{x}_i can be written as:

$$\dot{x}_i = k_i(x_i^{des} - x_i) \quad (2.26)$$

where a control action aims at achieving a single task. The general definition of the closed-form velocity-based WBC is then:

$$\dot{q} = \tilde{J}_1^\# \dot{x}_1 + \tilde{J}_2^\# \dot{x}_2 + \dots + \tilde{J}_k^\# \dot{x}_k \quad (2.27)$$

where \tilde{J}_i is a matrix that relates joint velocities to task space i velocities, such as the kinematic Jacobian matrix. $J_i^\#$ represents the pseudo-inverse of J_i , where one such calculation of the pseudo-inverse is the Moore-Penrose pseudo-inverse:

$$\tilde{J}_i^\# = \tilde{J}_i^T (\tilde{J}_i \tilde{J}_i^T)^{-1} \quad (2.28)$$

In this formulation, the control action \dot{x}_i can be any velocity measure such as Cartesian velocity, joint velocity, or centroidal momentum.

Some formulations, such as the Stack of Tasks [29, 30], develop the velocity-based WBC in a hierarchical manner for the control actions as:

$$\dot{q} = \tilde{J}_1^\# \dot{x}_1 + \tilde{J}_{2|1}^\# \dot{x}_2 + \dots + \tilde{J}_{k|p(k)}^\# \dot{x}_k \quad (2.29)$$

$$\tilde{J}_i = J_{i|p(i)} = J_i N_p(i) \quad (2.30)$$

where \tilde{J}_i is defined in the null-space, $N_p(i)$ of the tasks with higher priority than task i [32]. $N_p(i)$ can be written as:

$$N_p(i) = \prod_{j=i}^{i-1} N_j = \prod_{j=i}^{i-1} (I - J_j^\# J_j) \quad (2.31)$$

where I is an $n \times n$ matrix for n DoF of the robot [32].

Optimization techniques for velocity-based WBC can also be useful as they allow for the addition of inequality constraints rather than only equality constraints, but at a greater computation cost. In the optimization approach, a hierarchy exists in the form of multiple optimization problems that are solved.

Moro and Sentis define the optimization problems for k velocity tasks or constraints with $x_i = \dot{q}_i$, the linear equality and inequality constraints, where $A_i, C_i \in \mathbb{R}^{m \times n}$ and $b_i, d_i \in \mathbb{R}^m$, are as follows [32]:

$$A_i x = b_i \quad (2.32)$$

$$C_i x \leq d_i \quad (2.33)$$

With the tasks and constraints from Equation 2.32 and 2.33 the optimization problem is defined as:

$$S_i = \{\operatorname{argmin}_{x \in S_{i-1}} \|A_i x - b_i\|^2 + \|w\|^2\} \quad (2.34)$$

$$C_i x - w \leq d_i, w \in \mathbb{R}^{m+} \quad (2.35)$$

where w helps to solve the inequality constraints as a vector of slack variables. To solve the optimization problems for the hierarchical tasks and constraints, the optimization problems are recursively solved beginning with $S_0 = \mathbb{R}^n$ where S_i is a set of solutions x . The optimal solution of S_i is determined using a desired norm that is chosen as:

$$x_i^* = \{\operatorname{argmin}_{x \in S_i} \|x\|^2\} \quad (2.36)$$

It is worth noting that the closed form method could also be expressed such that it could be determined using an optimization solver, though the algebraic method is more efficient.

One Stack of Tasks implementation, named OpenSoT, offers velocity-based WBC using optimization [37]. The OpenSoT software library provides easy to use interfaces for setting up the WBC optimization problem and has been used on robots like COMAN, WALKMAN and CENTAURO [16]. A further development on OpenSoT is the CartesI/O Cartesian control framework that leverages the OpenSoT library with specific task and constraint definitions in the task space, allowing for simple Stack of Tasks implementations in ROS with python scripts and YAML parameter definitions [25].

2.3 Benchmarking

Benchmarking in robotics allows for the comparison of robotic systems under well-defined test scenarios with the goal of quantifying performance using specific metrics defined for the given test scenario. This allows users to compare different control and motion generation strategies on the same robot or compare the performance of different robots for the task being performed. This section provides some related works on benchmarking of robots in locomotion, manipulation and loco-manipulation scenarios.

2.3.1 Locomotion Benchmarking

Torricelli et al. [42] propose a scheme for benchmarking bipedal locomotion. The scheme considers the classification of bipedal balancing and locomotion tasks like stair walking, balancing under surface motions or under constant weight and walking on irregular terrain. They propose a scheme for benchmarks to benchmark bipedal locomotion with two main categories, performance and human likeness, and with metrics including success rate, energetic and mechanical cost of transport and joint states. Furthermore, they offer a method for developing benchmarking protocols for others to create new protocols and iterate on experimental protocols. Other related European projects including H2R, BALANCE, KoroBot, WALKMAN and Biomot motivated the development of this benchmarking scheme. Using principals from the benchmarking scheme of Torricelli et al., locomotion benchmarking was performed for HRP-2 measuring performance indicators like success rate, mechanical joint and actuator energy, cost of transport and duration [40]. Locomotion tasks tested on HRP-2 were inspired by KoroBot with motions like flat ground walking, walking on a beam, step stone walking and stair walking [2]. The KoroBot locomotion task can be seen in Figure 2.6. While the robot used was a humanoid robot, the tasks performed were limited to bipedal locomotion. This leaves benchmarking in the case of manipulation or manipulation during locomotion untreated.

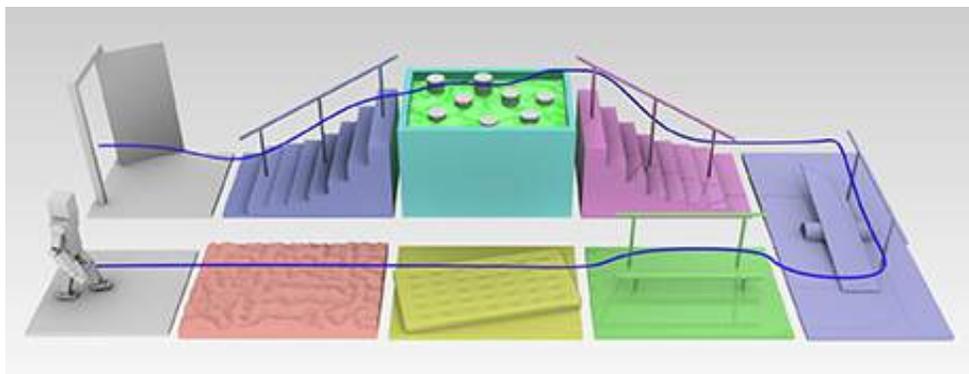


Figure 2.6: The KoroBot locomotion tasks. Adapted from [2].

2.3.2 Manipulation Benchmarking

In the field of manipulation and grasping, various types of benchmarks have been explored. For bimanual manipulation, a benchmark for semi-deformable objects was proposed using

two different tasks, plate assembly in watchmaking and belt assembly of engines. This benchmark offered CAD drawings to reproduce the task and proposed metrics like success rate, completion time and other metrics that considered planning, offline learning, and online learning [8]. Using soft end effectors, an evaluation method for grasping objects like fruits and vegetables was developed to measure grasping success [39]. Other benchmarking in the manipulation space includes planning related work like the evaluation of grasp planning algorithms for a manipulator with a gripper picking up household objects [7]. Other benchmarks focus more on dexterity such as a box and blocks test-based pick and place benchmark for robot manipulators [31] and in-hand manipulation like the benchmark for changing the pose of a household object using fingers or the environment [9]. It is worth noting that many manipulation related benchmarking tasks do not include the manipulation of large objects that require the consideration of dynamics, such as stability.

2.3.3 Loco-manipulation and Whole-body Motion Benchmarking

Recall that loco-manipulation with a humanoid robot involves the combination of walking and manipulating an object. Whole-body manipulation involves the manipulation of an object by a humanoid robot where it performs little to no stepping. For these tasks, benchmarking uses a combination of locomotion and manipulation related considerations. Many competitions have operated in this area like the DARPA robotics challenge where robots competed to complete tasks like driving a vehicle, traversing rough terrain, operating tools, entering a door, climbing a ladder, and closing a valve [10]. While the challenges offered interesting loco-manipulation scenarios, they lacked performance measurements using completion and timing as metrics. Other works in loco-manipulation and whole-body manipulation include those that are not competition based. The ability for a robot to balance during manipulation tasks like pushing or pulling was analyzed using a generalized ZMP approach to evaluate the balance of a humanoid robot during manipulation [15]. One work evaluated a bipedal robot’s walking posture while performing another task using a dynamic reconfiguration manipulability shape index based on manipulability metrics like dynamic and reconfiguration manipulability for greater dynamical shape changeability [38]. For loco-manipulation tasks like pushing a cart and opening a door, a planning method using constrained manifolds to increase planning speed implemented loco-manipulation tasks, but only evaluated the motion plan based on time. Other works offer methods for developing complex loco-manipulation motions for simulated tasks, but again only evaluated the motion based on time [34, 12]. One work benchmarked loco-manipulation tasks for box picking and object grasping on flat ground, uneven terrain and in restricted

space using Valkyrie, but with few metrics as they only measured success rate and the average time, while lacking the strict protocols as it was from a planning perspective [48].

2.3.4 EUROBENCH Benchmarking Framework

To create a unified benchmarking framework for bipedal locomotion in robotic systems, specifically humanoids and wearable robotics in Europe, the EUROBENCH project was created [1]. EUROBENCH addresses the lack of quantifiable performance metrics and standardized, repeatable approaches to measuring the performance of these robots. This project developed a benchmarking framework composed of several experimental test beds hosted in two testing facilities, one for humanoid robotics at Istituto Italiano di Tecnologia (IIT) and one for wearable robotics at Hospital Los Madroños, and a software system for computing robotic performance experimentally and in simulation. The two benchmarking facilities are accessible to the academic and industrial community so that companies and researchers can test the performance of their robots throughout development and access reliable, standardized information when purchasing a new robot for a given task. Each proposed benchmark is composed of three standardized parts: test beds, protocols and KPIs. The benchmarking software developed allows for simple calculation of the KPIs for each benchmark. The benchmarks in the project are inspired by real cases with a focus on locomotion related tasks. Elementary locomotion skills are measured with respect to stability, robustness, and motions on different terrains. In the case of the manipulation related tasks, the manipulation considered is more large object manipulation while balancing and walking. As a result, the object acts as a disturbance on the stability of the robot, rather than small object manipulation, dexterous manipulation or grasping related tasks. Figure 2.7 summarizes the outcomes and Figure 2.8 shows the parties and partners involved in the EUROBENCH project.

2.4 Bimanual Manipulation

This section covers related works in bimanual manipulation and specifically review related works in manipulability and reachability for humanoids and bimanual robots.

Human demonstration of bimanual manipulation tasks has proven useful for learning and developing complex bimanual motions. Performing stir fry cooking with two manipulator arms was learned from offline human demonstration to develop a learned transformer model for one arm and using visual feedback for the other to create motions online [27].

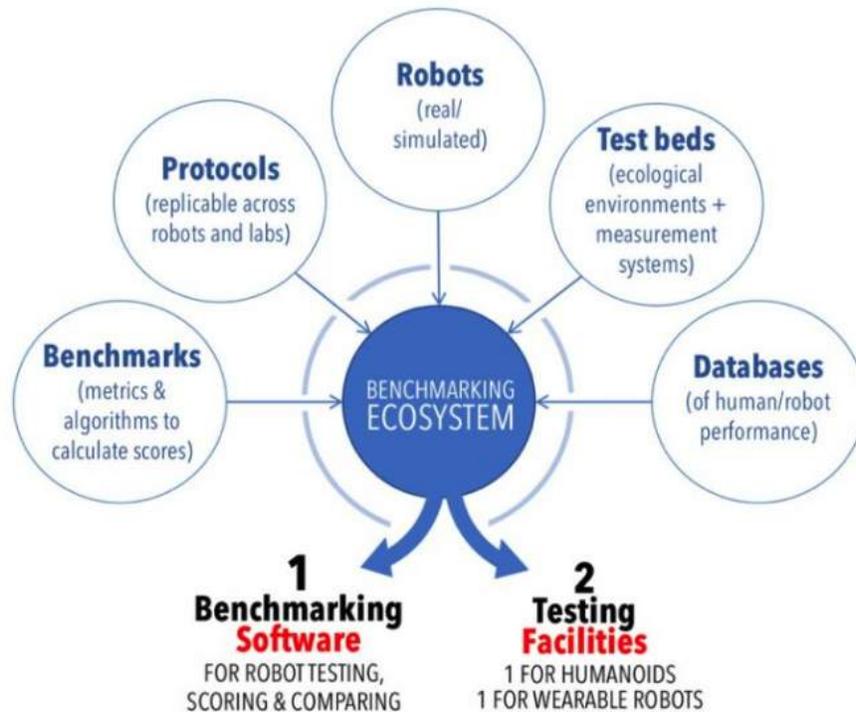


Figure 2.7: The outcomes of the EUROBENCH benchmarking project. Adapted from [1].

Another work involving human demonstrations for kitchen related tasks used learning to identify bimanual tasks according to the hand interactions with various objects [11]. A bimanual data set of 12 motions with 21 household objects for motions like pouring, stirring and cutting were recorded with whole-body motion capture data, finger trajectory data, various camera angle data and human subject IMU data to provide reference motions for study and possibly other learning from demonstration works for robots [23]. Alternate methods for developing bimanual motions rely on control based methods such as the method applied on ARMAR-6 using projected force-admittance control [13]. While these methods can be useful, they do not provide insight in to the kinematic complexity of bimanual tasks.

Reachability is often of interest with bimanual and humanoid robots, normally in terms of reachability maps. One method uses inverse dynamic reachability maps to compute offline quasi-statically stable poses for empty workspaces then online checks the feasibility of the pose and performs the planning to determine valid, reachable robot configurations

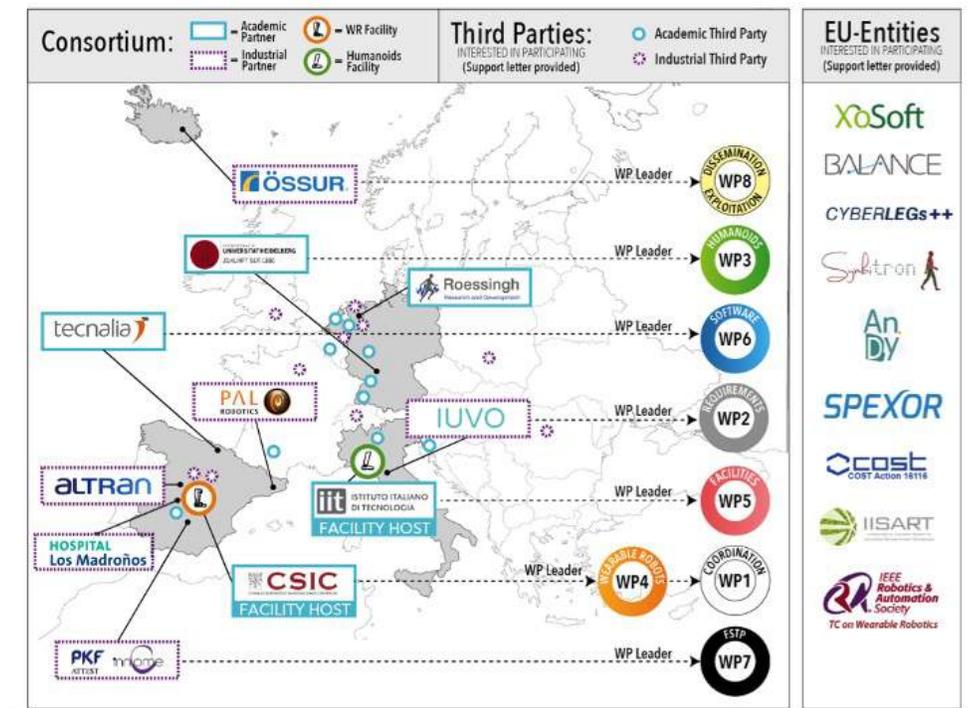


Figure 2.8: The partners and parties involved in the EUROBENCH project. Adapted from [1].

for a desired hand pose [47]. This work was extended to consider uneven terrain using a combination of forward and inverse dynamic reachability maps, where the forward maps were used with the pelvis as the base link and the feet as tip links and the inverse maps were used with a hand as the base link and the torso and other hand as tip links, to develop stances on non-horizontal planes [48]. Another work uses reachability maps in complete contact loco-manipulation motions of humanoid robots moving large objects to evaluate transitions of grasp changes and footsteps developed from a graph search approach [34].

Manipulability for bimanual manipulation is a more complex problem as two manipulators are used. One method proposed to measure bimanual manipulability would be based on the intersection of the manipulability ellipsoids in an attempt to combine kinematic constraints the manipulators would create on each other [26]. Vahrenkamp et al. [46] perform a series of works on bimanual manipulation using the ARMAR-III robot. Initially, a bimanual reachability approach was considered for one hand as the tip link and the other as the base link for a 14 DoF kinematic chain to compute reachable grasps and optimize grasps for increased manipulability. Then an extended manipulability metric

using the Yoshikawa manipulability (see Equation 2.14) and applying costs according to joint limits and nearby obstacles was developed for improved grasps [45]. This metric was then used to compute quality values for an inverse reachability study with ARMAR-III for different base poses given a required hand pose [44]. In terms of the bimanual workspace of ARMAR-III, the extended manipulability metric developed was used for a workspace analysis that considered the position of the two end effectors on either side of a point using spherical coordinates (see Figure 2.9) [43]. It is worth noting that ARMAR-III uses a wheeled base design and does not have any stability concerns so many of the poses searched are achievable on that robot but are possibly unstable on a humanoid robot.

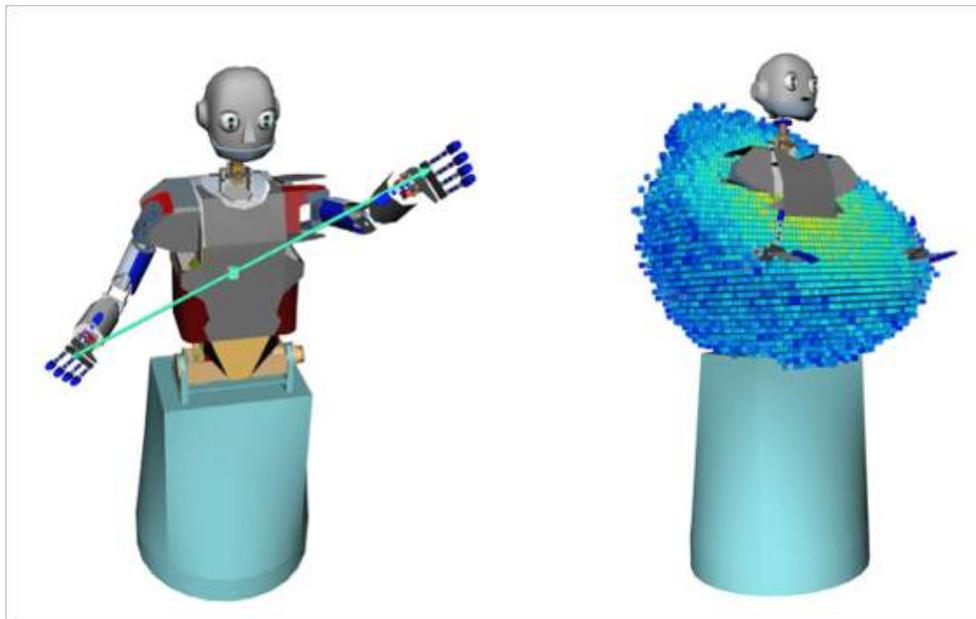


Figure 2.9: The bimanual manipulation workspace of ARMAR-III using the extended manipulability metric. Adapted from Vahrenkamp et al. 2014 [43].

With the ARMAR-4 humanoid robot, Kaiser et al. [19] developed a stability metric for static poses that was based on the distance of the GPCM from the edge of the support polygon divided by the distance of the center of the support polygon to the edge (see Figure 2.10). This metric was used to determine valid whole-body affordances from a visually perceived environment from reachability maps containing quality values from the stability metric. Multiple reachability maps could be combined to improve the selection of whole-body affordances, such as the stability metric with the extended manipulability metric of Vahrenkamp et al. [45] with the stability metric by calculating the product of the

quality values [20]. Though this treatment adds stability considerations necessary when using a humanoid robot, it only considers static stability for the first and last pose then moves between them using a WBC that keeps the robot stable.

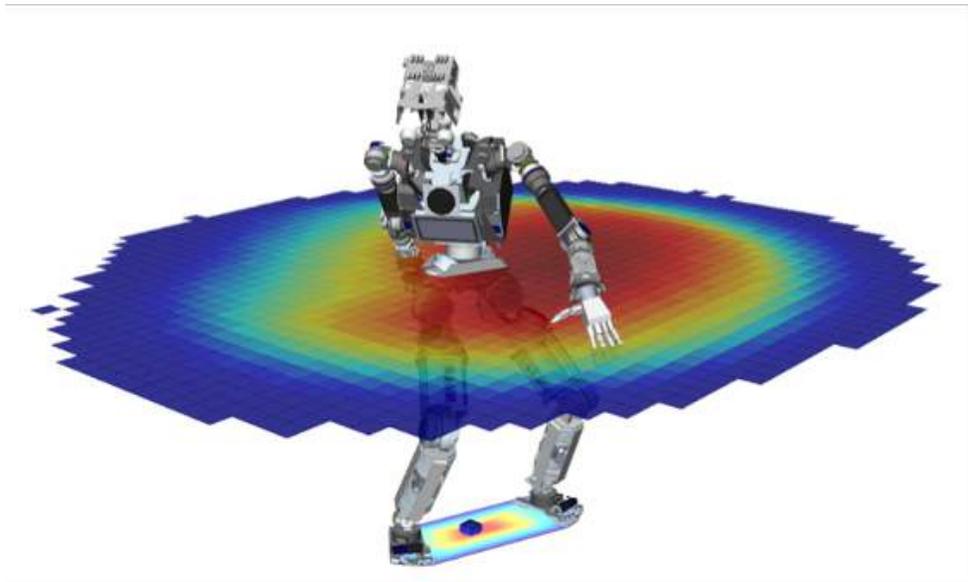


Figure 2.10: The stability metric visualized at feet for ARMAR-4 with extended manipulability metric visualized for left hand. Adapted from Kaiser et al. 2015 [20].

2.5 Summary

This chapter provided some background on robotic manipulation for key concepts like forward and inverse kinematics, manipulability, and planning. It also provided important background on humanoid control and motion generation including points related to stability, walking and whole-body control. Many of the concepts of robotic manipulation are frequently used in humanoid robotics, so together this provides an understanding of topics used in humanoid robotics.

An overview of benchmarking robotics was provided for a variety of cases. Some benchmarks had been developed and studied specifically for locomotion or fine manipulation. Overall, benchmarking was lacking specific benchmarks with detailed performance measurements in cases where both locomotion and manipulation are required, locomanipulation or whole-body manipulation. This thesis aims at providing a detailed bench-

mark for loco-manipulation and whole-body manipulation with specific test bed setups, protocols, and performance metrics to allow for better comparison of humanoid robots and control strategies.

In addition, an overview of bimanual manipulation was provided including recent works and many focused on reachability maps and workspace analysis. Bimanual manipulation involves highly constrained motions, which become even more complex when taking stability into consideration for a humanoid robot. This thesis also aims to develop a workspace analysis for humanoid robots that considers manipulability and non-static stability with specific object considerations and constraints to better inform motion development in bimanual manipulation scenarios.

Chapter 3

Whole-body Manipulation Benchmarking

This chapter develops a benchmark for humanoid whole-body benchmarking for the EU-ROBENCH benchmarking framework, consisting of a test bed, protocols and KPIs for a box manipulation scenario that involves little locomotion. Then, to provide an initial benchmark of a whole-body manipulation motion, experiments are performed for a partial set-up of the benchmark. The partial benchmark uses the REEM-C for two different motion generation and control strategies to provide a comparison for KPI evaluation. An overview of the benchmarking performed in this chapter can be seen in [Figure 3.1](#).

3.1 Whole-body Manipulation Benchmark

This section proposes the whole-body manipulation benchmark. To provide a benchmark that is realistic and could be seen in many applications, the benchmark is developed based on the manipulation of boxes. This sort of larger object manipulation is common in many industrial, logistics or warehouse applications that require tasks like shelving boxes or sorting objects. These tasks involve large whole-body motions where two arms are required to lift the objects and balance must be maintained. These sorts of motions can even translate to tasks for other applications like in the home or in the service industry. As the focus of this benchmark is on manipulation tasks with relatively little locomotion or none at all, the manipulation is restricted to a small workspace with manipulation tasks in the immediate vicinity of the robot.

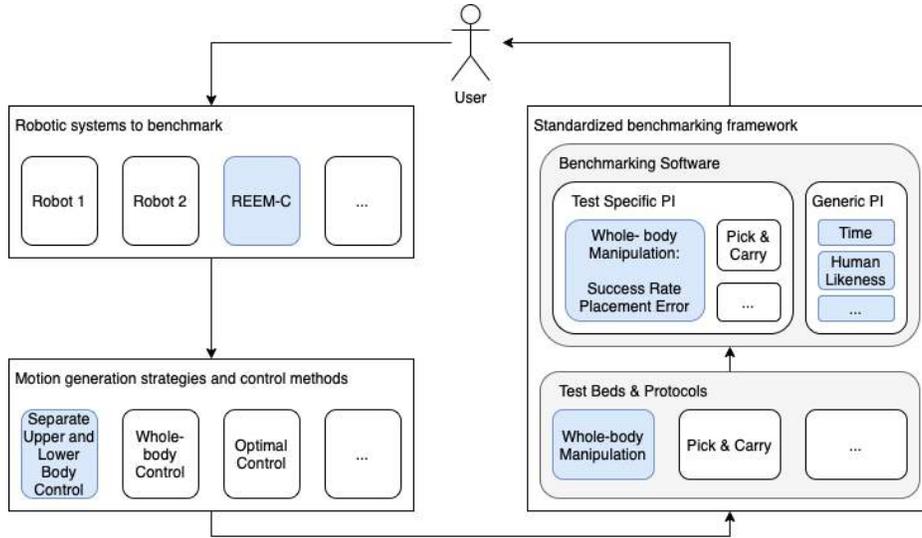


Figure 3.1: EUROBENCH benchmarking work flow with blue highlighting the whole-body manipulation work in this chapter.

3.1.1 Test Bed

This thesis proposes a test bed with a small, constrained workspace similar to a box manipulation scenario in a warehouse setting. This test bed consists of three shelving units surrounding a 1 m^2 floor space on three sides. On each shelving unit there are three shelves that help test pick and place heights like what a human can perform. These heights are selected as below the waist, between the waist and shoulders and above the shoulders at 0.14 m , 0.78 m and 1.42 m , respectively. To inform this decision, the heights are approximated using the REEM-C robot, the standard humanoid robot for the EUROBENCH testing facility (1.64 m tall). With the proposed shelving set-up manipulation motions vertically, laterally and combined vertically and laterally can be performed. To inform the robot’s pick and place locations, multi-colour LED light strips are placed along the edge of each shelf and target placement locations are placed on each shelf. An important part of developing a standardized benchmark is the ability to replicate the set-up worldwide, therefore this test bed uses IKEA IVAR system shelving units with a depth of 0.5 m , as they are available in many locations across the world [17].

The standard manipulation object proposed is a typical milk crate with the dimensions of 0.33 m wide by 0.33 m deep by 0.28 m tall ($13'' \times 13''$ by $11''$). The milk crate is chosen due to the versatility it offers as weight can be easily added to the open top and it allows

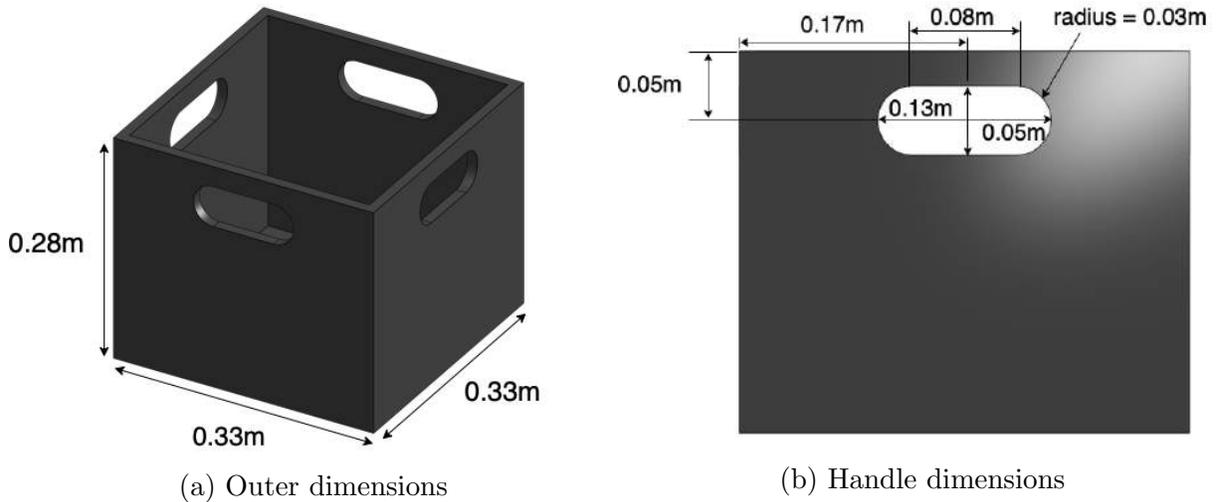


Figure 3.2: Standard whole-body manipulation benchmarking box.

for different grip options for the robot. The robot could grip the crate with the handles, like how a human would, or with a friction grip by pressing on the sides, which is a simpler grip for robots. The test object and key dimensions can be seen in [Figure 3.2](#). Moving forward, the term *box* may be used as well to describe the benchmark.

The test bed with the three shelving units and important dimensions can be seen in [Figure 3.3](#) with the standard manipulation object, the milk crate, and REEM-C, the standard humanoid platform for EURO-BENCH.

3.1.2 Protocols

Four protocols are proposed to evaluate the robot’s performance to offer a complete assessment of the robot in whole-body manipulation scenarios. As this box manipulation scenario involves the pick and place of a large object, these protocols are used to represent a general use case for whole-body manipulation rather than a single specific application. By maintaining generality and performing a comprehensive set of tests, the performance of the robot and its algorithms can translate to a variety of applications rather than a single case.

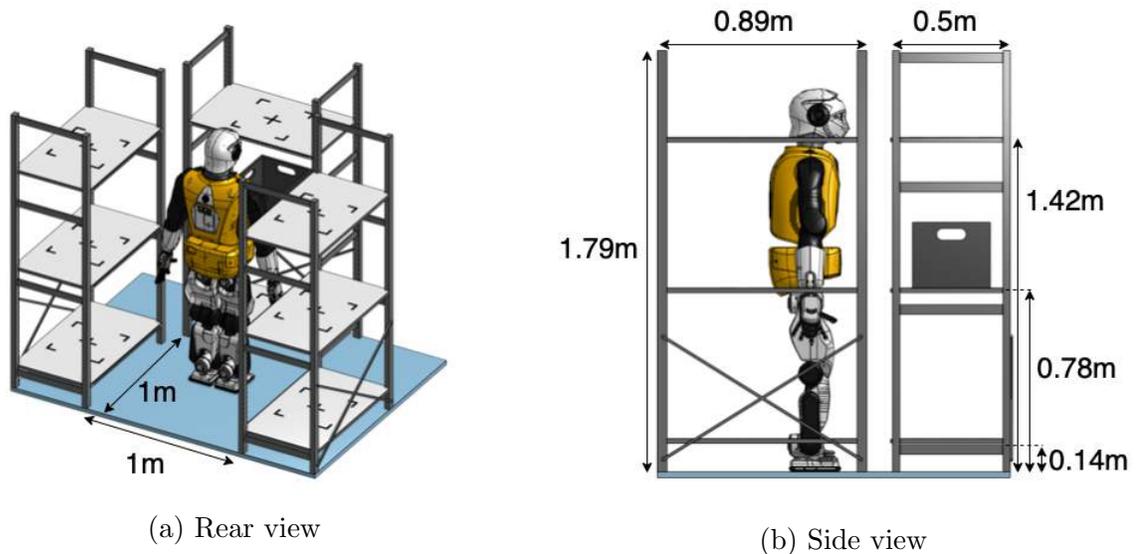


Figure 3.3: Whole-body manipulation test bed model.

Protocol 1: Predefined Frontal Placement

Protocol 1 considers the ability of the robot to perform whole-body box manipulations vertically. In this protocol the robot picks the box from a shelf on the middle shelving unit and places it on a different shelf of the same shelving unit using predefined locations. Both the pick and place location are predefined for each box motion. Over the course of this protocol several are performed, where each run consists of 10 placements. The placements are recorded for evaluation using the KPIs. When a run is successfully completed, the weight of the object is increased according to the user's choice, where weight is added to the box in an even distribution. The maximum time allowed per run is 15 minutes as it would be expected that a human could perform 10 motions within this time and to avoid overly long runs.

Protocol 2: Predefined Lateral Placement

Protocol 2 tests the ability of the robot to perform lateral whole-body box manipulation motions. The motions that the robot must perform consist of manipulations where the robot picks an object from a shelf in front or to the side of itself and must place it on a shelf of the same height on another shelving unit in front or to the side based on the predefined box manipulation. Like Protocol 1, multiple runs are performed with the weight of the box

increasing each time according to the user's choice, with 10 placements and a 15 minute time limit per run.

Protocol 3: Predefined Combined Placement

Protocol 3 tests combined vertical and lateral whole-body box manipulations of the robot. In this protocol the robot picks a box from a shelf in front or to the side of itself and places it on another shelf in front or to the side of itself at a different height for a predefined box manipulation. Similarly, several runs are performed with the weight increasing each run according to the user's choice, with 10 placements and a maximum time of 15 minutes per run.

Protocol 4: Variable Combined Placement

Protocol 4 tests whole-body box manipulations for combined vertical and lateral placement, but in a variable manner. In this protocol the robot must pick and place the box according to the LED lights on the front of the shelves that visually indicate the pick and place location where green lights indicate the shelf to pick the box from and red lights indicate the location to place the box. The LED lights are computer controlled to switch after the end of a manipulation motion and uses a random combination each time. Several shelves contain boxes with varying weights, up to the maximum weight the robot was able to manipulate in previous scenarios. A single run is performed with a fixed time of 20 minutes, where the robot manipulates as many boxes as it can in the given time. The run ends when the 20 minutes elapses or the robot cannot manipulate any more boxes.

Details on Performing a Protocol

All the protocols consist of the run structure, so the rules for performing the protocols are all similar and can be described through the performance of a run. Each run consists of box manipulations where the box of a certain weight is picked from one shelf and placed on another. At the beginning of the run, the robot can be placed anywhere within the 1 m^2 floor space where all segments of the robot must be above this same floor space and not within the shelves. Once the robot starts its motion then the time and joint data of the robot is recorded until the robot completes the box placements or the time limit elapses. During the run, the robot must return to the initial location it started from after each placement. The data recorded for the run is put into the desired CSV file format for the benchmarking software for performance metric processing.

3.1.3 Key Performance Indicators

Key performance indicators are the performance metrics measured for each protocol to evaluate the performance of the robot for a specific task. The KPIs are generated as a set of averages for all the box placements performed in the protocol tested. [Table 3.1](#) summarizes the KPIs used for benchmarking whole-body manipulation with further description for some following the table.

Overall, the goal of these KPIs is to measure how well the box manipulation is performed, so the time, success rate and placement error (position and orientation) are measured. Time helps measure the speed of the manipulation and success rate is important to measure if the task was successfully completed. By measuring placement error, that accuracy of the box manipulation can be measured. However, the strength, versatility and capability of the box manipulation is another important factor to measure, so the maximum weight and shelf heights reached are measured to indicate the limitations on the manipulations that can be performed. KPIs like mechanical work, cost of operation and power consumed offer insight into the energy consumption to perform the box manipulations and help forecast the cost to run a robotic platform. Finally, the human-likeness compares the motion of the humanoid robot to that of a human as it is trying to replicate the performance that a human can achieve.

In the case of the placement position and orientation error it is determined based on the error of the center of the box from the target position as considered in a 2-dimensional plane. The position error is calculated using the Euclidean error:

$$\epsilon_p = \sqrt{(x_{actual} - x_{target})^2 + (y_{actual} - y_{target})^2} \quad (3.1)$$

For the orientation error, there is only one degree of orientation in a 2-dimensional plane, so the orientation error is calculated as:

$$\epsilon_o = |\theta_{actual} - \theta_{target}| \quad (3.2)$$

The image below illustrates the error of a box placement:

The mechanical work is calculated as the work done by the robot on the box to move it from the pick location to the placement location calculated as follows:

$$W = \sum_{j=1}^n \int_0^T |\tau_j \cdot \dot{\phi}_j| dt \quad (3.3)$$

Table 3.1: Key performance indicators for the whole-body manipulation benchmark.

Name	Description/Formulation
Success Rate (%)	The ratio of objects placed on the correct shelf
Placement Position Error (m)	The average position error of the object placed on the shelf from the target location of the shelf as calculated in Equation 3.1
Placement Orientation Error (rad)	The average orientation error of the object placed on the shelf from the target location of the shelf as calculated in Equation 3.2
Time (s)	The average time for a correct box placement
Maximum Weight (kg)	The maximum weight of an object with which the robot can perform a placement
Shelf Heights Reached (m)	A list of the attainable shelf heights that the robot could successfully perform picks from and placements at
Mechanical Work (J)	The average absolute mechanical work calculated as in Equation 3.3
Cost of Operation ($Amps \cdot s/kg$)	The average energy consumed per placement and mass of objects calculated as in Equation 3.4
Power Consumed (W)	The average total amount of power consumed from the battery including actuators, sensors and PCs.
Human-likeness in Time (s)	The average difference in time of the robot motions to human motions performing the same task
Human-likeness in Path (m)	The average difference in position of the humanoid end effector paths to human hand paths performing the same task as in Equation 3.5

where n is the number of joints, τ_j is the torque at joint j , $\dot{\phi}_j$ is the angular velocity at joint j and T is the time to complete the entire box manipulation from initial starting position back to the initial starting position. Note that as this metric requires joint torques, some robots may not be able to record these experimentally, such as the position-controlled robot REEM-C. Alternatively, they can be determined by a dynamic simulation of the robot motion such as with Gazebo [14].

The cost of operation is calculated as the current consumed by the joints of the robot

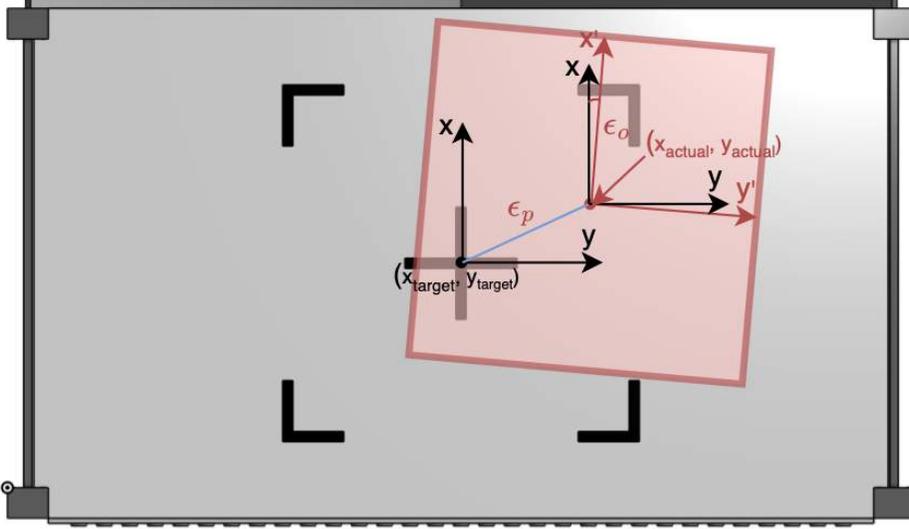


Figure 3.4: A box footprint, in red, misplaced on the shelf with respect to the target.

for a box manipulation divided by the mass of the box manipulated and is calculated as follows:

$$Cost = \sum_{j=1}^n \int_0^T \frac{1}{m_O} |i_j| dt \quad (3.4)$$

where n is the number of current consuming joints, i_j is the current of the actuator for joint j , T is the time to perform the entire box manipulation from initial starting position back to initial starting position, and m_O is the mass of the box being manipulated. Motor currents are typically readily available for most robots, such as in the joint states published with ROS by REEM-C during operation.

The human-likeness in path is calculated as the Euclidean error between the humanoid end effector paths to the human hand paths over the trajectories calculated as:

$$\epsilon_{ee} = \sum_{j=1}^{n=2} \int_0^{\tau} \sqrt{(x_{r,j}(\tau) - x_{h,j}(\tau))^2 + (y_{r,j}(\tau) - y_{h,j}(\tau))^2 + (z_{r,j}(\tau) - z_{h,j}(\tau))^2} d\tau \quad (3.5)$$

where $n = 2$ is the number of robot end effectors and human hands, $(x_{r,j}(\tau), y_{r,j}(\tau), z_{r,j}(\tau))$ are paths of robot end effector j over a normalized time, $(x_{h,j}(\tau), y_{h,j}(\tau), z_{h,j}(\tau))$ are paths of human hand j over a normalized time, τ is the normalized time for the two motions. As the motions of the robot and human are unlikely to take the same time, the normalized time is computed to allow for comparison in this calculation.

Given the number of metrics, the user has access to many measures of performance of the robot performing whole-body manipulation that can be used to determine what robot or algorithm best suits the application. Not all KPIs are as important as others because each application has unique requirements. The user can use the metrics to understand performance like the speed, accuracy, strength, efficiency, and human-likeness and how this may relate to their task. For example, in a situation where a robot is stocking shelves it would likely be important to stock shelves quickly, but also accurately to ensure space is optimized. In this case, metrics like time and placement error would be very important. If the robot is operating independently in a separate environment, then human-likeness metrics are of little concern as how natural the motion matters very little. However, if the environment was collaborative with a human, then human-likeness could be important to help the human in the environment understand the robot's motions better and feel more comfortable working alongside the robot. This demonstrates how a general use case benchmark allows the user to flexibly consider the performance for their specific case and interpret the performance indicators in a way that is useful for their needs.

To calculate the KPIs evaluating the robot's performance a set of CSV files in a specified format containing the robot's joint data is provided to the software framework, which consists of joint angles, joint velocities, joint currents, and the measured or simulated joint torques. For this benchmark a few additional CSV files are provided to add experimental details such as success rate, placement error and object weight. The general structure of the files for a protocol is a set of CSV files for each run that the software splits into box manipulations. The calculations are computed for each box manipulation then are averaged at the end to provide an average value for the KPIs that use average values. [Figure 3.5](#) shows the workflow for the performance metric processing. It is important to note that the KPI calculation is standardized to ensure that KPIs are all calculated identically and measure exactly what the benchmark is defined to measure. This allows the end-user to interpret these standardized metrics to determine what is important for their application. Should the user wish to add to the benchmarking framework or modify part of a benchmark they could consider a collaboration to meet their need.

As this benchmark is part of the EUROBENCH framework, which focuses on benchmarking humanoid robots, this benchmarking set-up could be used for non-humanoid robots, though they would score poorly in human-likeness and the benchmark of other robots is not the main intention. This benchmark within the EUROBENCH framework is specifically used for humanoid robots. There is another whole-body manipulation benchmark that is specifically for wearable robotics.

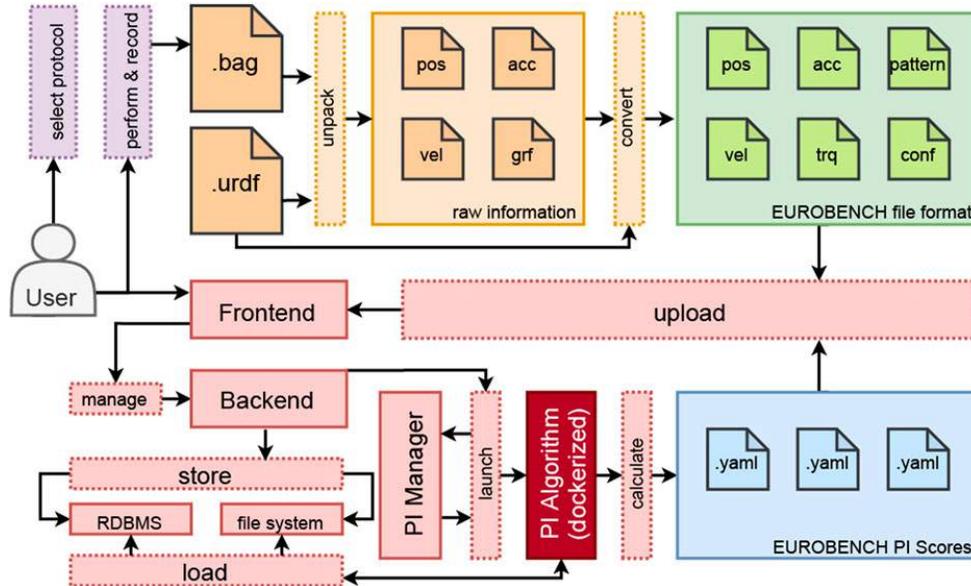


Figure 3.5: EUROBENCH benchmarking performance metric work flow where the protocol is performed and data collected in orange. Green highlights the data converted to the EUROBENCH format and red indicates the performance indicator calculation workflow, with Docker use for the metric calculation is implemented (dark red). The output to the user is shown in blue. Adapted from Aller et al. 2022 [3].

3.2 Partial Benchmark Performed

This section performs a simplified version of the benchmark for Protocol 1 with the REEM-C to begin to evaluate the motion of a humanoid on the whole-body manipulation benchmark proposed. Due to time limitations, certain simplifications were made in the experimental set-up, performance of the protocol and number of KPIs measured; however, it provides a set of initial motions to validate the use of the benchmark and offers insight into possible improvements. In addition, some parts of the benchmark had yet to be finalized, so may differ slightly from the complete benchmark presented earlier.

3.2.1 Experimental Setup

To test Protocol 1, the experiment uses test bed as described in [Subsection 3.1.1](#). For this set-up it uses only the center shelf, but still respects the 1 m^2 floor space. The main reason

for not including the other shelves is that they are not used in this protocol and allow for clearer photos and videos of the motions. Two more alterations to the test bed described exist in this partial benchmark. The LED light strips are not installed on the shelf as Protocol 1 does not require them and the target is altered, which is explained below.

When performing these experiments, the box for manipulation had yet to be finalized, so a cardboard box with dimensions 0.5 m in width, 0.16 m in height and 0.15 m in depth, weighing 0.1 kg was used. This box was selected as appropriate for this partial benchmark as it large enough to require a bimanual pick and place motion but is not exceedingly heavy and allows for a simple friction grip.

See [Figure 3.6](#) for the middle shelving unit with the box on the middle shelf in front of Waterloo’s REEM-C “Seven”.

In the execution of Protocol 1, only vertical placements are performed. For this partial benchmark the box manipulation from the middle shelf to the top shelf is performed. The box starts on the middle shelf, centered horizontally at the front edge, and should be placed on the top shelf in a similar manner at the target location. The box is placed near the edge of the shelf to allow for an easier pick and place for the robot in this partial benchmark.

3.2.2 Performance Metrics

The partial benchmark uses certain KPIs from [Subsection 3.1.3](#) to measure the performance of the motions. To measure the placement error, the experiment uses a simplified scheme, where two taped boundaries are used on the top shelf that indicate the box has met that error level if the robot places the box within the inner edge of the boundary. The green tape indicates a 2.5% error in placement accuracy and the blue tape indicates a 5% error in placement accuracy where the error margin added to the target location is based on a percentage of the shelf’s dimension. Again, this simplified target location and error measurement strategy was used as it had not been finalized when performing the motions. The KPIs selected are seen in [Table 3.2](#), noting that an adjusted success rate is used to accommodate the taped boundary version of placement error.

3.2.3 The Humanoid Robot “Seven”

Seven, who arrived at the University of Waterloo in the summer of 2020, is a REEM-C humanoid robot developed by PAL Robotics and is the 7th REEM-C in the world. It is a

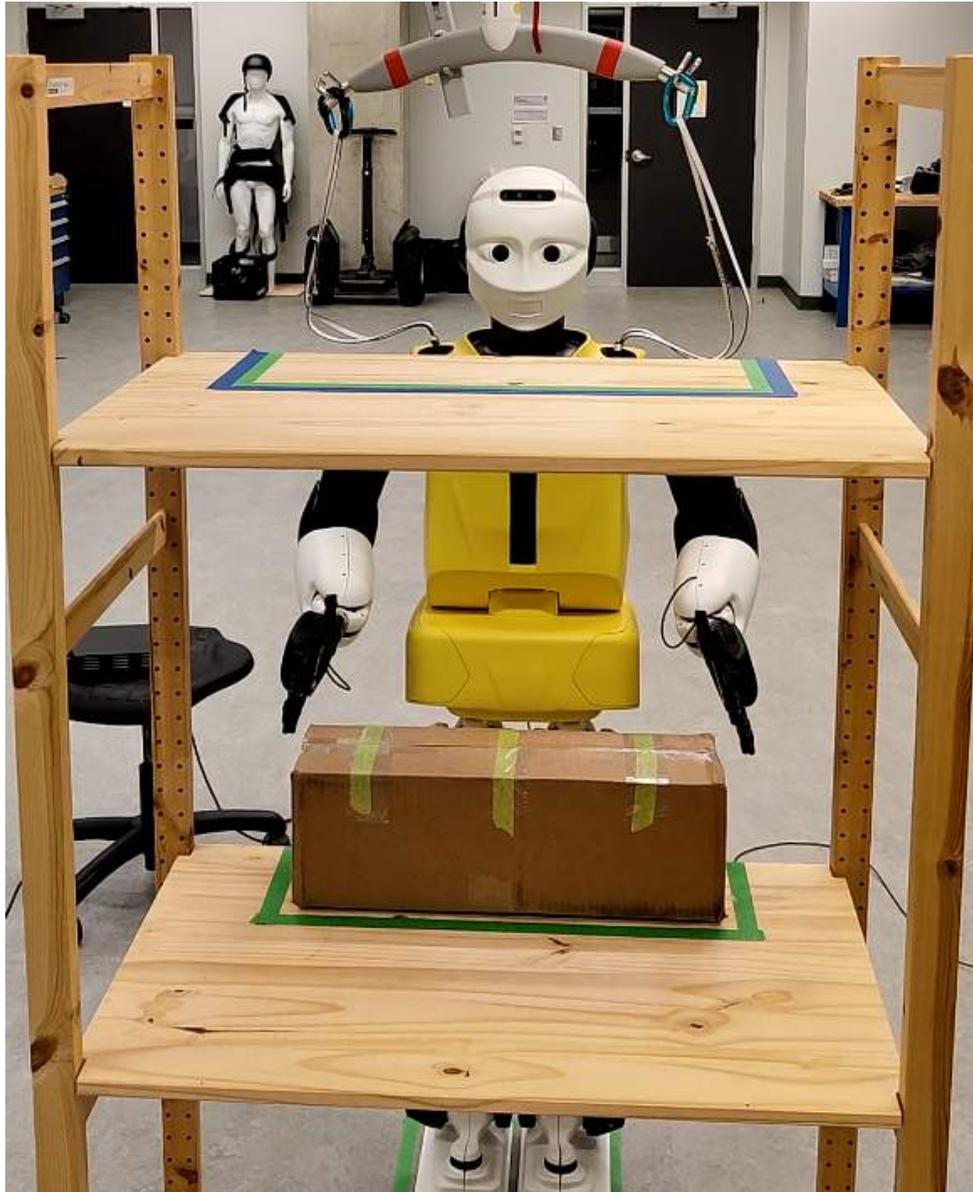


Figure 3.6: Partial whole-body manipulation benchmarking set-up.

Table 3.2: Key performance indicators for the partial whole-body manipulation benchmark.

Name	Description/Formulation
Success Rate of 2.5% Placement Accuracy	The ratio of objects placed correctly within the 2.5% error boundary
Success Rate of 5% Placement Accuracy	The ratio of objects placed correctly within the 5% error boundary
Time (s)	The average time for a correct box placement
Mechanical Work (J)	The average absolute mechanical work calculated as in Equation 3.3
Cost of Operation ($Amps \cdot s/kg$)	The average energy consumed per placement and mass of objects calculated as in Equation 3.4

floating base robot that contains 68 DoF, including the two 19 DoF hands that are under-actuated. Seven is 1.64 m in height and weighs 80 kg . With respect to sensors, there is an IMU in the pelvis, a force torque sensor at both wrists and ankles, lasers in the feet and an Intel RealSense camera in the head with an NVIDIA Jetson TX2 for vision processing. Out of the box, REEM-C comes with certain functionality developed by PAL Robotics including walking, stabilization, grasping, whole-body control and text to speech, with code development available using ROS. Another member of the REEM-C series, the 6th REEM-C, is available at the EURO-BENCH facility available for performing benchmarks by those who wish to use it with code they develop.

3.2.4 Motion Generation and Control Methods

Separate Upper and Lower Body Controllers

The first motion generation method is the separate upper and lower body control method, where the torso and arm joints are controlled using one controller and another is used to control the leg joints. This split control method allows for the highly redundant kinematic structure to be simplified for the controllers and motion planning. The lower body controller used is the default controller developed by PAL Robotics. This controller includes Kajita walking pattern generation [21] with IK for the leg joints and a stabilizer. [Figure 3.7](#) shows an overview of the walking controller control structure. The upper body controller uses joint trajectory controllers for each arm with the torso joints maintained at

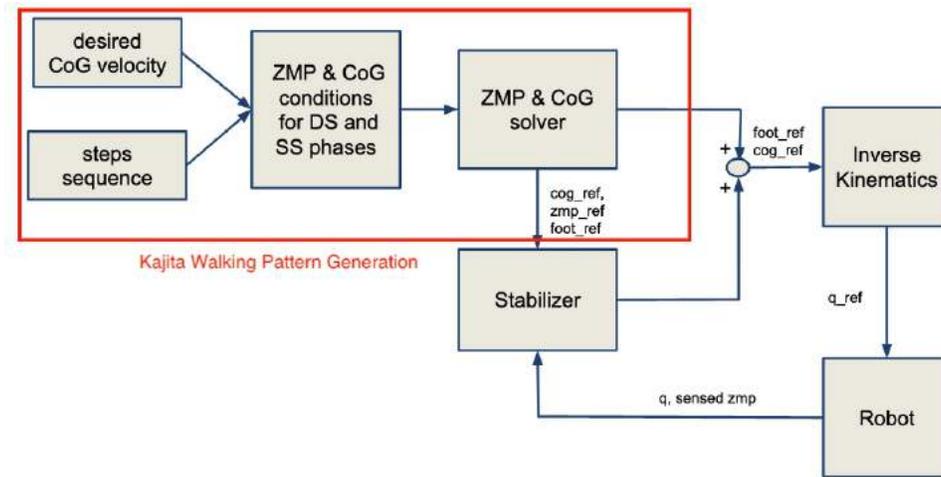


Figure 3.7: PAL Robotics walking controller. Adapted from Baños 2020 [5].

the zero position. The torso joints are kept at zero to stay within a stability region that the stabilizer can compensate. The joint trajectory moves the joints of the arms through joint trajectories consisting of positions at specified times along with velocities and accelerations if provided using spline interpolation. In this case, inverse kinematic planning results from MoveIt! [33], a kinematic motion planning framework, are used to solve joint position trajectories that respect joint limits and are collision free with respect to the rest of the robot. Motion plans with MoveIt! are created with the TRAC-IK kinematics solver using a threaded approach to solve the IK problem with two different solvers where a valid solution is returned from whichever converges first. One solver uses an inverse Jacobian method with Newton’s method convergence that can avoid local minima created from joint limits and the other uses sequential quadratic programming nonlinear optimization with quasi-Newton methods [6]. MoveIt! leverages the Open Motion Planning Library (OMPL) [41] to generate motion plans for moving the arms from one position to another. RRT-Connect was the planner used in this implementation, which is very similar to regular RRT except it grows two trees until they meet with one at the start position and one at the goal position [24]. Due to the fact that MoveIt! does not contain motion planning for tree-like kinematic structures, only chains, the joint positions for a given hand pose were pre-calculated with MoveIt! for a single arm then mirrored for both arms when sending joint trajectories. An overview of the separate upper and lower body control structure can be seen in Figure 3.8 where a predefined arm motion and stepping sequence is provided as input to the motion generation scheme.

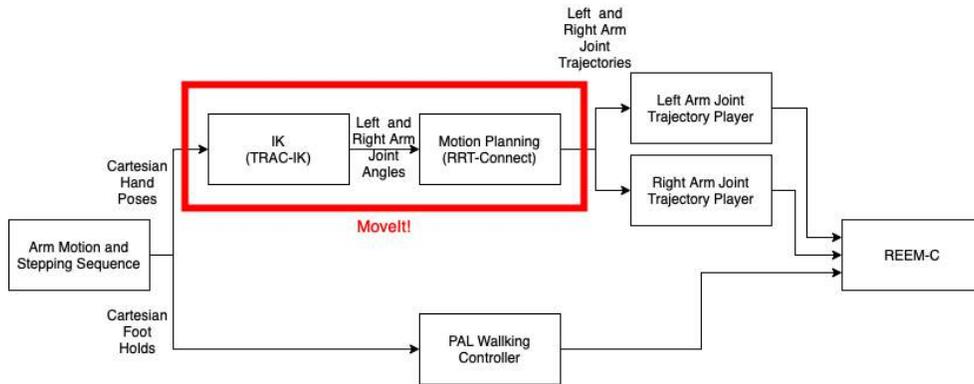


Figure 3.8: Separate upper and lower body control method.

To generate motions with this control method, MoveIt! can be used to determine if the robot can reach the box from the starting position. If the box is out of reach, then the robot can walk forward so that it can reach the box and pick it up. Again, using MoveIt!, it can be determined if the box can be lifted to the top shelf without any collisions. If not, then the robot can step backwards to raise the box then step forward and place it. This methodology allows an arm motion and stepping sequence to be developed as the input to the control method in Figure 3.8. Note that the inverse kinematics for the hand positions to generate the arm joint positions is performed offline. Also, based on the frame definitions of the robot’s arms, the same joint angles can be provided to both arms to achieve a mirrored arm pose. Based on this motion generation approach, the resultant motion requires stepping and is shown in Figure 3.9.

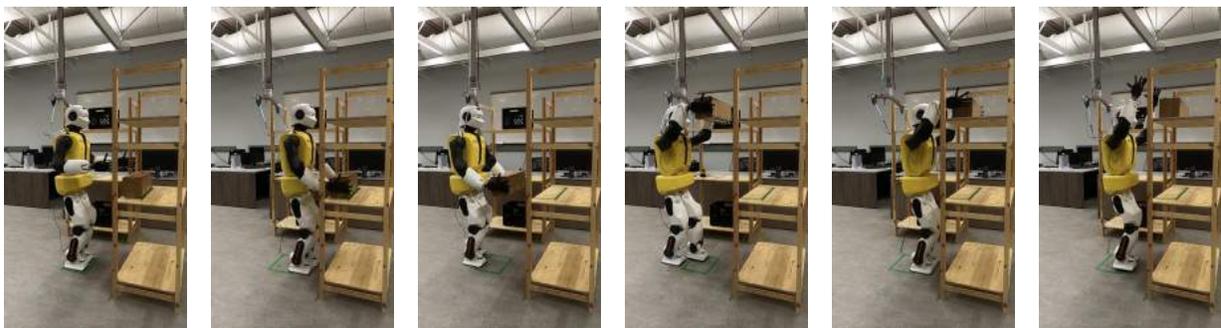


Figure 3.9: Separate upper and lower body controllers method motion sequence.

In this motion, the robot takes steps towards the shelf and away from the shelf depend-

ing on the stage of the motion sequence. A benefit of this motion generation method is that it uses the ability of humanoid robots to move around in the environment to use a wider workspace for obstacle avoidance and reachability issues experienced in this scenario.

Whole-body Controller

The second control method and motion generation approach is a whole-body control method using Cartesi/O [25], the Cartesian control framework that uses the OpenSoT Stack of Tasks library, which uses a velocity-based optimization approach. This approach allows for the creation of a hierarchical Stack of Tasks approach with tasks defined in the Cartesian world frame as well as the ability to specify constraints. In this implementation, joint and velocity limits were set as constraints and the task list was set in the following order, descending in priority:

1. Left foot Cartesian pose, right foot Cartesian pose
2. Center of mass XY position, waist yaw
3. Left hand Cartesian pose, right hand Cartesian pose, torso posture, left arm joint 2 posture, right arm joint 2 posture
4. Head posture, left leg posture, right leg posture

To generate motions with this method, Cartesian poses were sent to the left and right hand along with center of mass XY positions. This generates whole-body motions for the robot online with joint states composed of position and velocity being published in ROS. These joint states can then be sent to the joint trajectory controllers for all the joints to play motions on the robot. An overview of the whole-body control method can be seen in Figure 3.10.

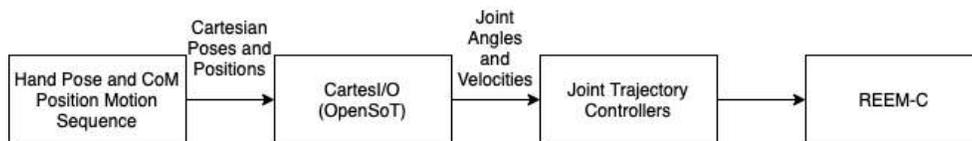


Figure 3.10: Whole-body control method.

Due to the ability to move all joints simultaneously, the robot can bend and lean forward to grab the box then pull it back and raise the box to place it on the top shelf without

taking steps or encountering collisions with the shelf. The motions generated using this approach can be seen in [Figure 3.11](#).



Figure 3.11: Whole-body controller method motion sequence.

It can be seen that the robot makes use of the legs to crouch then reach further forward with the arms by using the torso joint. An increased number of possible motions with the robot is offered with this WBC method allowing for a valid box manipulation to be found in the constrained workspace. In this solution, the general trajectory of the box was selected to avoid collisions as no collision avoidance was implemented or available with CartesI/O. To bias the elbows out to avoid self-collision while lifting the box, the left and right joint 2 postures were used. Other postures like waist yaw, torso posture, head posture and left and right leg postures were used to produce favourable positions of certain joints during the motion, such as keeping the torso upright while raising the box, or limiting the motion altogether, like keeping the waist from turning and the head moving for no reason. Another task limitation was that there was no stabilization task available or implemented, so the center of mass position was shifted to maintain balance. It is worth noting that even with the ease of assembling the task list this motion still requires the fine tuning of parameters in the WBC definition on the robot to achieve a feasible, stable motion, which is a common problem encountered with Stack of Tasks methods.

3.2.5 Experimental Results & Discussion

To collect data that could be processed for the KPIs of the partial benchmarking scenario, each of the two motion generation and control approaches were performed 10 times for the box manipulation from the middle shelf to the top shelf on the center shelving unit for the 0.1 kg cardboard box. [Table 3.3](#) shows the results for the KPIs with mechanical work and

cost of operation computed as described in [Subsection 3.1.3](#). Time, mechanical work and cost of operation are all computed as average values with standard deviations for the 10 box manipulations performed with each motion.

Table 3.3: Results from 10 trials of two motion generation approaches for the partial whole-body manipulation benchmark

Performance Metric	Separate Upper and Lower Body Control Method	Whole-body Control Method
Success Rate of 2.5% Placement Accuracy	60 %	100%
Success Rate of 5% Placement Accuracy	100%	100%
Average Time \pm Standard Deviation (s)	52.96 \pm 0.14	62.41 \pm 0.12
Mechanical Work \pm Standard Deviation (J)	2585.25 \pm 775.33	561.69 \pm 7.75
Cost of Operation \pm Standard Deviation ($Amps \cdot s/kg$)	10802.08 \pm 15.14	20245.68 \pm 9.67

In the results, it can be seen that the separate upper and lower body control method was approximately 10 seconds faster than the whole-body control method. This increased speed came at the expense of a 40% lower success rate for the 2.5% placement accuracy. The whole-body method achieved excellent accuracy due to the non-stepping approach it used as taking steps introduces error in the position of the robot with respect to the shelf. Both methods did achieve perfect 5% placement accuracy though. In terms of mechanical work, the separate upper and lower body control method was almost five times that of the whole-body method as it does more work on the box by taking steps during the motion than the non-stepping whole-body method. However, the cost of operation for the separate upper and lower body control method is approximately half that of the whole-body control method. The reason for this is that much more current is consumed in the whole-body method due to the longer time to perform the box manipulations and the increased use of joints in the robot. Performing crouching motions with the robot consumes high amounts of current to support the robot’s weight and remain stable compared to the current required for stepping. The differences in performance metrics measured offer a trade-off for the user

depending on the needs of their application as certain KPIs may matter more to them than others. Therefore, if a user required a very accurate motion, the whole-body control method would be the better algorithm to use, but if a faster or more energy efficient algorithm is important then the separate upper and lower body control method would be more desirable. Overall, the results between trials were very similar as indicated by the small standard deviations in time and cost of operation, with no noticeable decrease in performance. However, larger standard deviations can be observed for the mechanical work, especially for the separate upper and lower body control method. The larger variability is owing to the variability of the results obtained through Gazebo simulations and the variation in the way the PAL walking controller performs the stepping action to remain stable.

When performing the experiments, a few other observations were made that should be considered regarding the whole-body manipulation benchmark. In a few cases the robot would bump the shelf on the approach for placing the box using the separate upper and lower body control method. This only affected the success rate of the box placement. While this did not damage the shelf or the robot, this is an undesirable result as it could damage the robot in other real applications. Also, depending on the way the box is manipulated the object or shelves could get damaged due to motions like dragging a heavy object across a shelf, banging the object into the shelf, or dropping the object. Furthermore, if the object was something delicate or an open top box like the milk crate with contents that could be spilled then the contents of the box being manipulated is of greater concern. In the partial benchmark here, the box was shaken around during the walking motions of the separate upper and lower body control method, which could be damaging for the object. To treat these motion safety concerns, a metric for how carefully the object is manipulated and if a collision occurred could be valuable. A possible solution to monitor how carefully the object is manipulated could be to add an IMU to the manipulated object and measure the change in acceleration. With respect to collisions, a possible solution could be a metric that counts the number of collisions of the robot or box with the environment (aside from placement).

Furthermore, while the cost of operation does give a sense of the current consumed at the joints this can be challenging to interpret in terms of battery usage. If the robot is always plugged in or performing the motions with a full battery, then the battery usage is of little concern. However, the battery life of most humanoid robots is very short and operating with a cable attached is generally not desirable due to tripping and tangling hazards. Using the power consumption metric proposed as part of the complete benchmark would help give a better understanding of the power required for the task, but it may also be valuable to provide a metric in terms of the number of batteries used or charge cycles required as

this is a very real problem when using humanoids in real world applications. One other note on battery use is that some systems may experience a degradation in performance as the battery depletes. To identify this, the performance of the robot would need to be observed over time. Possible solutions would be to perform a time-based comparison of the motions or perform the same motion at two different charge states, like full battery and low battery, to compare the performance.

A final note on the motions in this partial benchmark is that the grasping approach used was a simple friction grasp. REEM-C does have hands with five fingers that allow for the robot to grasp objects and even handles on a box like the milk crate; however, the kinematic structure adds complications. On the arms of REEM-C there are 7 DoF, but joint 5 (the first joint after the elbow) and joint 7 can each rotate the hand and joint 6 can tilt the hand. This means that the robot cannot simultaneously tilt the hands up or down and bend them in or out at the wrist while keeping the palms of the hands parallel to each other the way humans can. A single hand grasping the milk crate handle can be seen in [Figure 3.12](#).



Figure 3.12: REEM-C grasping milk crate handle.

As a result, picking a handled box and lifting it using the arms is more constrained than expected resulting in a very small bimanual workspace for this box, not to mention the added constraints involving possible collisions and stability of the robot. This challenge motivates the work presented in [Chapter 5](#), which explores the workspace of humanoid robots in bimanual manipulation scenarios with object imposed constraints.

Chapter 4

Loco-manipulation Benchmarking

This chapter develops a benchmark for humanoid loco-manipulation benchmarking for the EURO-BENCH benchmarking framework, which like the whole-body manipulation benchmark, consists of a test bed, protocols and KPIs for a box manipulation scenario though with a longer locomotion phase. Then it constructs and tests a partial set-up of the benchmark to provide an initial benchmark of the scenario. Also similar to the whole-body manipulation, the chapter performs KPI evaluation using two motion generation and control strategies. Unfortunately, due to the REEM-C being damaged in another experiment, results remain limited. An overview of the benchmarking performed in this chapter can be seen in [Figure 4.1](#).

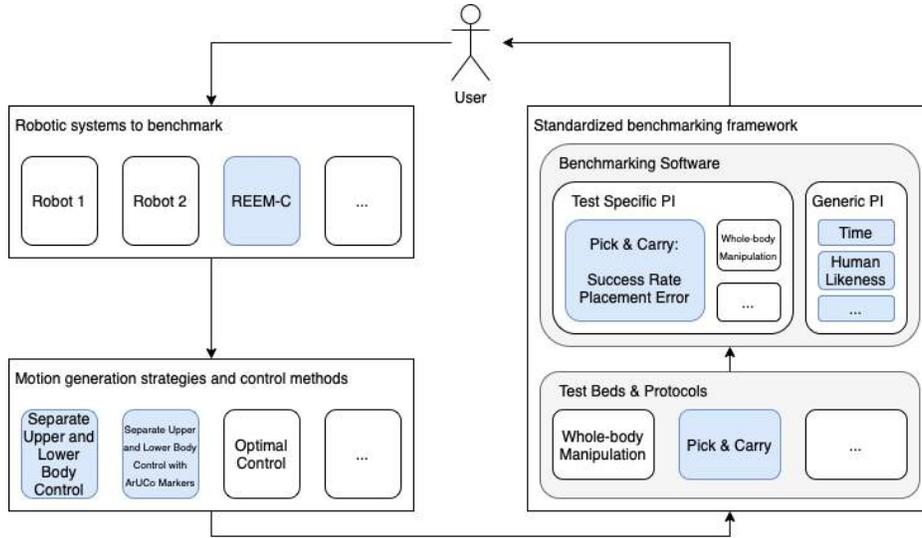


Figure 4.1: EUROBENCH benchmarking work flow with blue highlighting the loco-manipulation benchmarking work in this chapter.

4.1 Pick and Carry Benchmark

This section proposes the pick and carry benchmark for loco-manipulation. Similar to the whole-body manipulation benchmark, a realistic benchmark is inspired by industrial, logistics and warehouse applications. Often materials or large objects need to be lifted and carried to different locations like transporting packages or moving objects on a factory floor. This involves lifting and carrying these objects to place them at more distant locations. Again, these sorts of motions are applicable in the home or service industry like carrying a basket of laundry or a tray of food. This benchmark focuses on manipulation tasks that require significant locomotion, so the workspace is larger to force the robots to walk during the manipulation of the objects.

4.1.1 Test Bed

The proposed test bed replicates a common pick and place scenario like in a warehouse setting. The loco-manipulation benchmark, which will also be referred to as the pick and carry benchmark, is essentially an extension of the whole-body manipulation benchmark. It uses two of the same IKEA IVAR shelving units, but instead they are separated by a larger

distance, 3 *m*, which requires locomotion from the robot to perform the box manipulation task of picking and carrying a box to the other shelf. The shelving units are set up with the same shelving heights as the whole-body manipulation test bed based on human sized heights below the waist, between the waist and shoulders and above the shoulders at 0.14 *m*, 0.78 *m* and 1.42 *m*, respectively. The shelves contain the same targets with LED light strips on the edges too. It also uses the milk crate as the standard manipulation object (see [Figure 3.2](#)). [Figure 4.2](#) shows the two shelving unit test bed with the milk crate and REEM-C.

4.1.2 Protocols

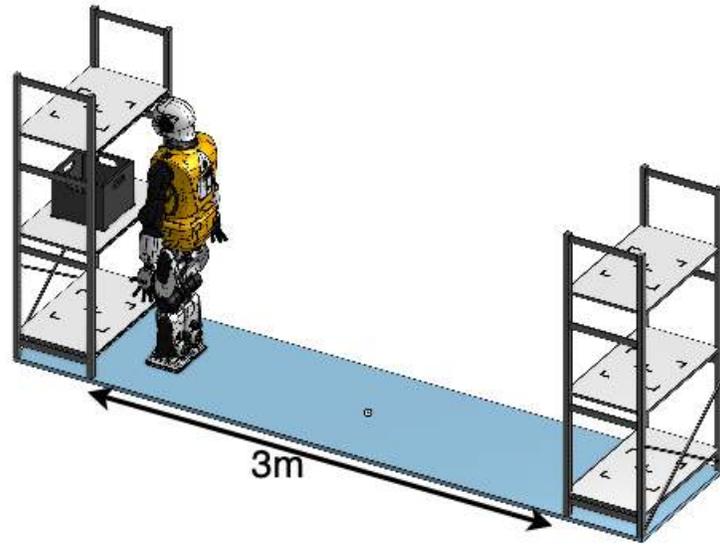
To evaluate the robot’s performance in the pick and carry scenario four protocols are proposed. Again, generality is maintained with the definition of these protocols to ensure the performance is measured in a way that will allow for the extension to a variety of applications.

Protocol 1: Predefined Basic Carrying

Protocol 1 tests the ability of the robot to pick and carry objects from one shelving unit to another. For predefined pick and place locations, the robot picks up the box from one shelf and transports it to a specified shelf on the other shelving unit. This protocol requires multiple runs to be performed, with 10 placements per run. After the successful completion of a run, the weight of the box is increased based on the user’s choice. The maximum time permitted per run is 20 minutes as it is expected that a human could perform a run in this time, and it helps avoid extremely long runs.

Protocol 2: Predefined Basic Carrying at Specified Velocity

Protocol 2 tests the ability of the robot to perform pick and carry motions at specified velocities. In this protocol the robot performs pick and carry motions from one shelving unit to the other with predefined pick and place locations using the heaviest weight the robot can carry in the box. Multiple runs will be performed, which again contain 10 placements with a maximum time of 20 minutes per run. For each run the locomotion speed for carrying the box is increased based on the user’s choice.



(a) Rear view



(b) Side view

Figure 4.2: Pick and carry test bed model.

Protocol 3: Endurance Carrying

Protocol 3 tests how long the robot can carry the heaviest object it can. The pick and carry test bed is not used for this protocol. Instead, the robot must walk and carry the object as long as it can around the track available at the EUROBENCH facility for humanoids. This protocol only has one run, that consists of 10 trials. The 10 trials are used to generate an average result. Each run ends when the robot can no longer carry the object.

Protocol 4: Variable Endurance Pick and Carry

Protocol 4 tests how long the robot can perform pick and carry motions. In this protocol the LED lights indicate pick and place locations, with green indicating the pick location and red indicating the placement location, that the robot must perform pick and carry motions for. Multiple shelves will contain boxes with different weights, up to the maximum weight the robot can handle. This protocol only uses a single run of 20 minutes, where the test ends when the time runs out or the robot cannot manipulate anymore objects.

Details on Performing a Protocol

The same run structure is used in all these protocols and therefore all require very similar rules. The robot must be placed in the center of the two shelving units when starting a run and after each placement must return to the initial position after each placement. The only exception to this is Protocol 3 that continuously carries the object on the track and begins holding the object. Similar to the whole-body manipulation benchmark, the time and joint data are recorded once the motion starts until the completion of the required motion or time elapses. Again, the recorded data is placed in the required CSV format for KPI processing with the benchmarking software.

4.1.3 Key Performance Indicators

The key performance indicators used for this benchmark are very similar to the ones for the whole-body manipulation benchmark in [Table 3.1](#) and are calculated similarly, though now walking speed is more of a concern than manipulation locations. Due to the longer locomotion phase the walking speed will directly relate to the performance of the robot in this scenario. [Table 4.1](#) shows the KPIs that will be calculated for the protocols. See [Subsection 3.1.3](#) for full calculation details and equations.

Table 4.1: Key performance indicators for the pick and carry benchmark.

Name	Description/Formulation
Success Rate (%)	The ratio of objects placed on the correct shelf
Placement Position Error (m)	The average position error of the object placed on the shelf from the target location of the shelf as calculated in Equation 3.1
Placement Orientation Error (rad)	The average orientation error of the object placed on the shelf from the target location of the shelf as calculated in Equation 3.2
Time (s)	The average time for a correct box placement
Maximum Weight (kg)	The maximum weight of an object with which the robot can perform a placement
Maximum Walking Speed (m/s)	The maximum walking speed of the robot while carrying the heaviest object that it can handle
Mechanical Work (J)	The average absolute mechanical work calculated as in Equation 3.3
Cost of Operation ($Amps \cdot s/kg$)	The average energy consumed per placement and mass of objects calculated as in Equation 3.4
Power Consumed (W)	The average total amount of power consumed from the battery including actuators, sensors and PCs.
Human-likeness in Time (s)	The average difference in time of the robot motions to human motions performing the same task
Human-likeness in Path (m)	The average difference in position of the humanoid end effector paths to human hand paths performing the same task as in Equation 3.5

4.2 Partial Benchmark Performed

This section presents a partial benchmark with REEM-C for a simplified version of Protocol 1 to begin evaluating motions on the pick and carry test bed. Due to issues with the robot being damaged the results are limited but are currently in progress. This partial benchmark shares many similarities with the whole-body manipulation partial benchmark in [Section 3.2](#) and is referred to throughout the section.

4.2.1 Experimental Setup

To test Protocol 1, the test bed proposed in [Subsection 4.1.1](#) is used except the shelves do not have LED light strips. The motion tested involves picking up a box from the middle shelf of the shelving unit that the robot is facing and placing it on the top shelf of the other shelving unit behind the robot. In a similar manner to the whole-body partial benchmark, a cardboard box with dimensions 0.5 m in width, 0.16 m in height and 0.15 m in depth, weighing 0.1 kg was used with a simple friction grip. The placement location at the front edge of the shelf uses the taped boundaries as seen in [Figure 3.6](#). The set up for the pick and carry partial benchmark can be seen in [Figure 4.3](#).



Figure 4.3: Partial pick and carry benchmarking set-up.

4.2.2 Performance Metrics

For the partial benchmark the same KPIs are used as in [Subsection 3.2.2](#), with the exception that a success rate for placing the box on the correct shelf not considering placement accuracy is added. This is due to the fact that the placement accuracy may be tougher to achieve when performing locomotion over a larger distance.

Table 4.2: Key performance indicators for the partial pick and carry benchmark.

Name	Description/Formulation
Placement Success Rate	The ratio of objects placed on the correct shelf (no error considerations)
Success Rate of 2.5% Placement Accuracy	The ratio of objects placed correctly within the 2.5% error boundary
Success Rate of 5% Placement Accuracy	The ratio of objects placed correctly within the 5% error boundary
Time (s)	The average time for a correct box placement
Mechanical Work (J)	The average absolute mechanical work calculated as in Equation 3.3
Cost of Operation ($Amps \cdot s/kg$)	The average energy consumed per placement and mass of objects calculated as in Equation 3.4

4.2.3 Motion Generation and Control Methods

Separate Upper and Lower Body Controllers

The first motion generation method is the separate upper and lower body control method as described in [Subsection 3.2.4](#), see this section for full details. [Figure 4.4](#) shows an example of this motion under development.

In this motion, the robot walks towards the shelf and picks up the box then retreats and turns around. After a few side steps to re-center, the robot then approaches the other shelf to place the box. It is worth noting that due to the PAL walking controller using the underlying Kajita pattern generation [\[21\]](#), the robot cannot turn on the spot. Turning involves walking in a small arc.

Separate Upper and Lower Body Controllers with ArUco Markers

The second motion control method uses the same separate upper and lower body controllers as the first method, except prior to the stage of approaching the second shelf to place the box, the robot uses the camera in the head to measure how centered the robot is with respect to the shelving unit based on ArUco markers on the shelving unit. It reads the ArUco markers seen by the camera and turns or takes a sidestep accordingly to center itself

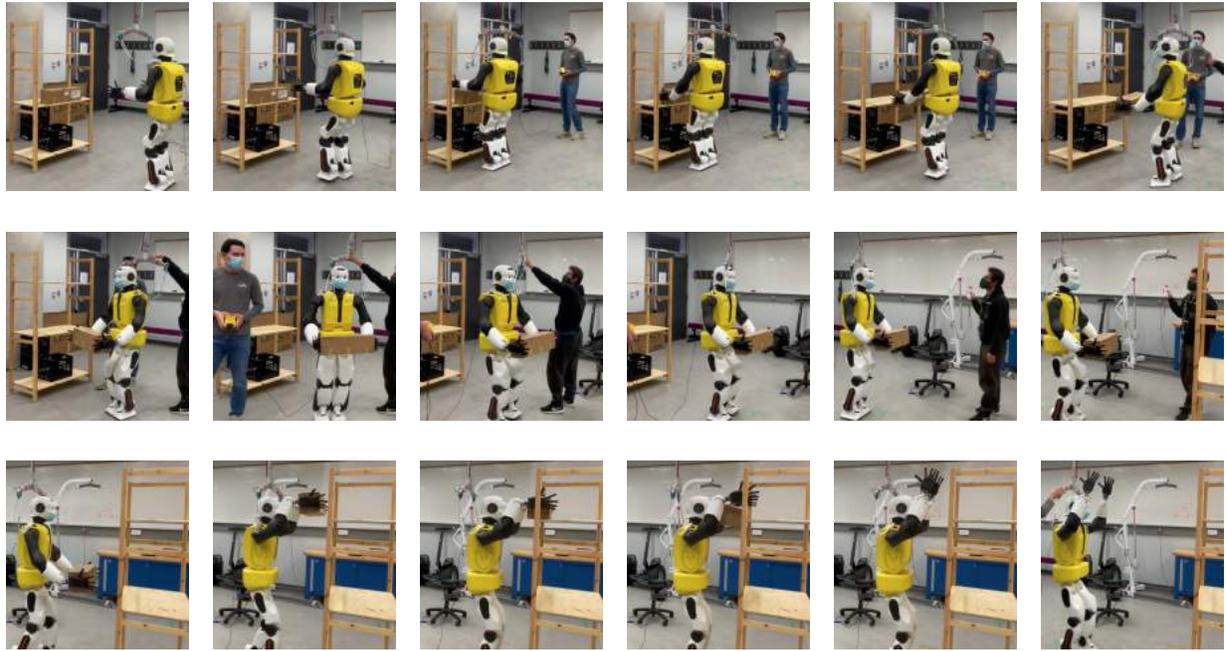


Figure 4.4: Separate upper and lower body controllers method pick and carry motion sequence ordered across then down.

with respect to the shelf. Once reoriented and re-centered, the robot can walk towards the shelf for the box placement. This method was motivated to correct the error in the robot's stepping motions that are slightly inaccurate due to foot slippage and stepping inaccuracies. Current observations of the robot walking with the PAL walking controller show that when the robot is walking over a larger distance, it tends to drift. This error is due to slipping of the feet on the floor and the inaccuracy in the actual distance travelled compared to the footsteps commanded. In part, this can be tuned through experimentation, but it is not always repeatable. By adding in ArUco marker error correction, the robot can determine if it is centered with respect to the shelf and if it is oriented to walk directly at the shelf. This error correction should produce a higher success rate for placement accuracy and less potential misplacement, such as the box missing the shelf. [Figure 4.5](#) shows the ArUco markers on the shelf used for error correction as seen by the robot.

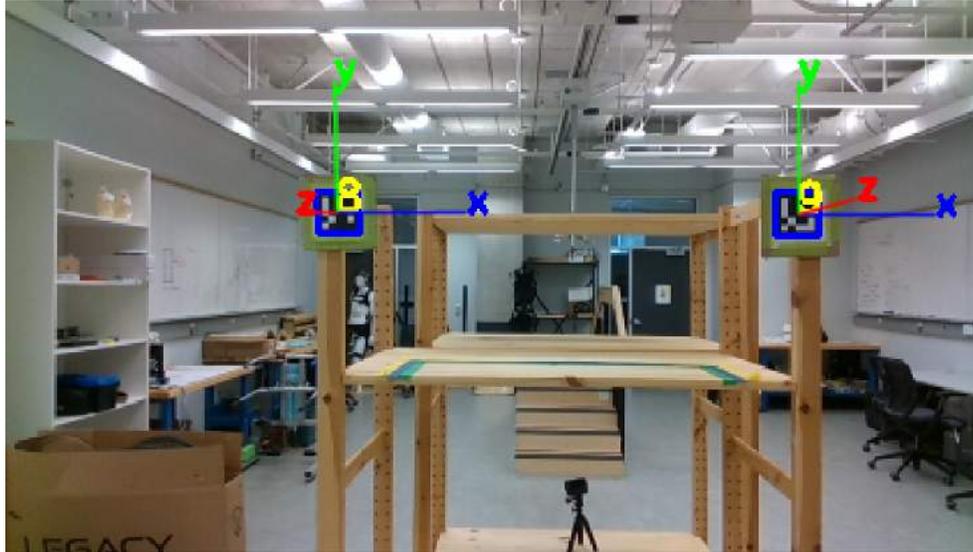


Figure 4.5: ArUco markers on shelving unit as seen by the robot.

4.2.4 Experimental Results & Discussion

Unfortunately, the REEM-C used in these experiments was damaged in another experiment that caused a broken force torque sensor in the right foot and in turn the robot could not walk. Experimental results are limited to the motion generated and shown above, but further experimentation is in progress.

Chapter 5

Bimanual Manipulation Workspace Analysis

This chapter performs an in-depth analysis of the workspace of humanoid robots for bimanual motions. As identified at the end of [Chapter 3](#), bimanual box manipulation can result in highly constrained motions with some that are possibly infeasible, motivating this work. Building off previous work in bimanual workspace analysis, as discussed in [Section 2.4](#), this chapter develops a new quality metric to represent the manipulability and stability of a bimanual pose. Next it uses this metric to perform a general workspace analysis similar to the work of Vahrenkamp et al. [\[43\]](#). Then workspace analyses are performed for specific bimanual object grasps studying the impact of CoM velocity of the robot, object mass, object grasp and robot kinematic and dynamic structure on the workspace. The work presented in this chapter is the culmination of an equal collaboration with Vidyasagar Rajendran.

5.1 Workspace Metrics

This section develops a quality metric for workspace analysis that accounts for bimanual manipulability and stability of the humanoid. Individual manipulability and stability metrics are first developed followed by a method for combining them into a single quality metric.

5.1.1 Manipulability

Manipulability can often be measured with manipulability metrics like those presented in [Equation 2.12](#), [2.13](#) and [2.14](#) presented in [Subsection 2.1.2](#), but these metrics are intended for the manipulability of single chains. In the case of bimanual manipulability, the manipulability of two arms must be considered, which makes it more challenging to define a manipulability metric. One way to define the bimanual manipulability, denoted as c_m is by selecting the lower manipulability as shown below:

$$c_m = \min(\mu(A_{LeftArm}), \mu(A_{RightArm})) \quad (5.1)$$

where $\mu(A_{LeftArm})$ and $\mu(A_{RightArm})$ are the manipulabilities of the left and right arm, respectively, as calculated using [Equation 2.14](#) for the volume of the manipulability ellipsoid. The bimanual manipulability metric is defined as such because in a bimanual task the less manipulable arm will be the one that limits the motion making it more representative of the true manipulability of the object being manipulated.

5.1.2 Stability

To define a stability metric for a humanoid robot in a given pose, the capture point as described in [Equation 2.21](#) in [Subsection 2.2.1](#) can be considered as it allows for the inclusion of a velocity at the center of mass. To determine if a pose is stable with the capture point, the capture point should be found inside the support polygon. Therefore, the closer to the edge of the support polygon, the closer the humanoid is to becoming unstable. This leads to a definition of the stability metrics as the minimum distance to the edge of the support polygon:

$$c_s = \min(d_f, d_b, d_l, d_r) \quad (5.2)$$

where d_f , d_b , d_l and d_r are distances from the capture point to the front edge, back edge, left edge and right edge of the support polygon, respectively, when the feet are in a static side by side pose. Only the non-stepping case will be considered in this analysis for simplicity. These distances and the support polygon for a humanoid robot can be seen in [Figure 5.1](#), where the stability metric would select the distance d_f marked in red for the support polygon marked in blue. Note that when the velocity is zero it reduces to the static stability criterion using the GPCM.

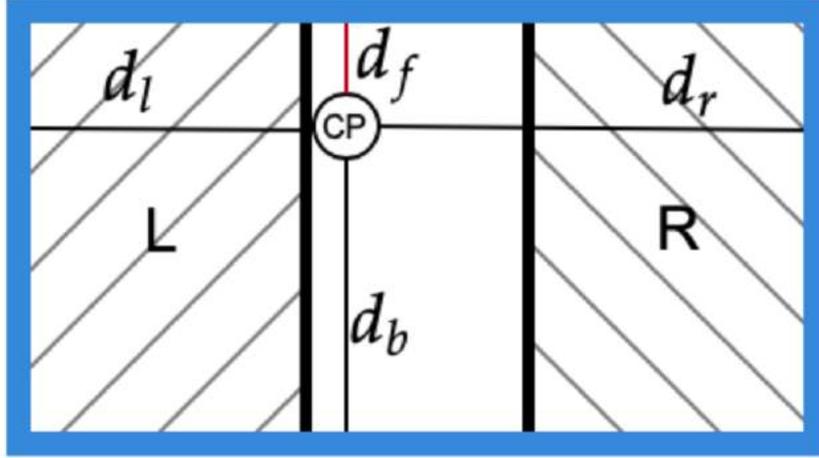


Figure 5.1: A simplified support polygon showing the capture point and the selected minimum distance for the stability metric, d_f .

5.1.3 Combined Manipulability-Stability Metric

A combined manipulability metric can be defined to allow for the evaluation of reachable points with both considerations. As values of different scales are produced in [Equation 5.1](#) and [5.2](#) they can be normalized using the largest value found for each metric during the workspace analysis as follows:

$$c'_{mi} = \frac{c_{mi}}{\max(c_{m0}, c_{m1}, \dots, c_{mn})} \quad (5.3)$$

$$c'_{si} = \frac{c_{si}}{\max(c_{s0}, c_{s1}, \dots, c_{sn})} \quad (5.4)$$

where n is the number of reachable points in the workspace analysis and i is the i^{th} reachable point.

Using the normalized metrics in [Equation 5.3](#) and [5.4](#), a weighted sum can be used to create the combined manipulability-stability metric for a reachable point as:

$$c_{vi} = \alpha c'_{mi} + \beta c'_{si} \quad (5.5)$$

where α and β are weights between 0 and 1 that when added together equal 1. Weight values of 0.5 each are proposed for the initial evaluation of the metric.

5.2 Workspace Generation and Visualization

5.2.1 Whole-body Inverse Kinematics

In order to solve the inverse kinematics problem of reaching a point in space with both arms of a humanoid, a whole-body inverse kinematics approach must be used due to the tree-like kinematic structure of the robot, unlike the approaches typically used with chain-like kinematic structures for typical manipulators. This problem can be solved using an optimization approach with the **Extensible Optimization Toolset** (EXOTica) framework, that defines an optimization problem with a cost function constructed from task maps and allows for constraints [18]. Task maps can be used to define desired functionality like joint limits, end effector position and orientation and center of mass position. For the whole-body inverse kinematics optimization, the task maps are defined as seen in Table 5.1.

Table 5.1: EXOTica task maps in whole body inverse kinematics formulation.

Task Map	Formulation
Joint position	$\Phi_{Ref}(\mathbf{x}) = \mathbf{x} - \mathbf{x}_{ref}$
Joint limits	$\Phi_{Bound}(x) = \begin{cases} x - x_{min} - \epsilon, & \text{if } x < x_{min} + \epsilon \\ x - x_{max} + \epsilon, & \text{if } x > x_{max} - \epsilon \\ 0, & \text{otherwise} \end{cases}$
End effector frame	$\Phi_{EffFrame}(x) = \mathbf{M}_A^B$ where $M_A^B \in SE(3)$
End effector position	$\Phi_{EffPos}(x) = \mathbf{P}_A^B$ where $P_A^B \in SE(3)$
End effector orientation	$\Phi_{EffRot}(x) = \mathbf{R}_A^B$ where $R_A^B \in SE(3)$
Center of mass	$\Phi_{CoM}(\mathbf{x}) = \sum_i (\mathbf{P}_{CoM_i}^{world} m_i)$

With these task maps, the positions of the tool center point (TCP), defined as the left and right hand of the humanoid, can be queried. In this problem the model of the REEM-C, Seven, uses simplified hands since use of the fingers and grasping is not considered. As the task maps form an optimization problem, they require weights. The weights in descending priority are respecting joint limits, restricting position of the feet, limiting unnecessary CoM motion, and reaching desired end effector poses.

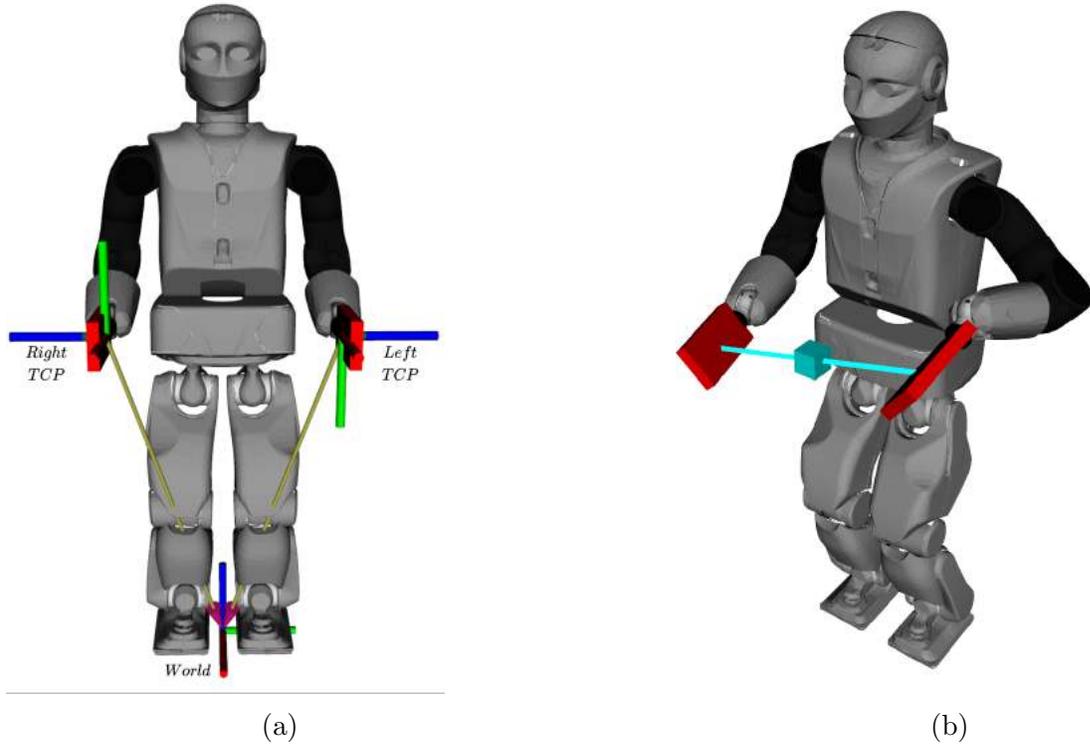


Figure 5.2: a) The REEM-C humanoid with *World*, left and right TCP frames with x in red, y in green and z in blue. b) A single point on the workspace grid indicates the center between the two end effectors of the robot for general workspace generation.

5.2.2 General Workspace Generation

The general workspace analysis is similar to that of Vahrenkamp et al. [43], which specifies a center point and then sets the hands to positions around the point for different distances. In this situation includes no considerations for objects, such as the specific hand poses to manipulate the object. In the generation of this thesis, a 6D voxelized workspace is created by discretizing the area in front of the robot where most manipulation tasks would occur for humanoids. The center points for the workspace analysis are the points in the 3D grid with a range of spherical coordinates at each point. This point is described by a 3D position vector in the *World* frame, which is depicted in Figure 5.2a, and the end effector positions are encoded with the spherical coordinate (r, θ, ϕ) as seen in Figure 5.2b. Note that no end effector orientation is stored in this method. The lack of orientation greatly

increases the speed of the generation because of the simpler position-only IK performed.

For the workspace generation, let $j = [v_0, v_1, v_2]$ be position of a 6D voxel, v , in the *World* frame. If the spherical coordinates are $s = [r, \theta, \phi]$ then another Cartesian position, $j' = [x', y', z']$, can be calculated as:

$$x' = r \sin(\theta) \cos(\phi) \quad (5.6)$$

$$y' = r \sin(\theta) \sin(\phi) \quad (5.7)$$

$$z' = r \cos(\theta) \quad (5.8)$$

which allows for the Cartesian positions of the left and right TCPs to be calculated in the *World* frame as follows:

$$TCP_{left} = j - j' \quad (5.9)$$

$$TCP_{right} = j + j' \quad (5.10)$$

As a result, the TCPs can be positioned around the point at different distances, where they are always across from each other. For an IK solution to be considered valid during the generation of the workspace it must be reachable using this bimanual configuration, the capture point must lie within the support polygon, the humanoid’s configuration must be collision free and obey the joint limits. These requirements are performed with checks after the whole-body IK is performed as the optimization approach does not have any constraints applied, so invalid solutions can be generated.

[Algorithm 3](#) details the generation method for the general workspace analysis. This algorithm takes the range of each voxel dimension, the discretization steps, and the velocity of the center of mass and then computes the metric c_v as defined in [Equation 5.5](#) for a valid IK state of the robot. Note that q represents the whole-body position as calculated by the inverse kinematics in `calcBothArmsIK` that takes in the left and right TCP positions. Also, the algorithm is designed to compute both sides of the workspace without exploiting symmetry. While this does result in double the generation time, it allows for the consideration of asymmetric models, which is more realistic for a model that represents the real robot very well.

5.2.3 Workspace Generation with Objects

The general workspace analysis generated in the previous section provides a generous view of manipulating an object in a bimanual scenario, but this is not very realistic as objects introduce various constraints. Constraints related to objects include the fact that the

Algorithm 3 General Bimanual Workspace Generation

Input: x, y, z, r, θ, ϕ ranges and discretization steps, V_{COM}

Parameters:

```
1: for  $x \in \{x_{min}, \dots, x_{max}\}$  do
2:   for  $y \in \{y_{min}, \dots, y_{max}\}$  do
3:     for  $z \in \{z_{min}, \dots, z_{max}\}$  do
4:       for  $r \in \{r_{min}, \dots, r_{max}\}$  do
5:         for  $\theta \in \{\theta_{min}, \dots, \theta_{max}\}$  do
6:           for  $\phi \in \{\phi_{min}, \dots, \phi_{max}\}$  do
7:              $j \leftarrow \text{Vector}(x, y, z)$ 
8:              $j' \leftarrow \text{calcSphericalCoordOffsets}(r, \theta, \phi)$ 
9:              $q \leftarrow \text{calcBothArmsIK}(j + j', j - j')$ 
10:             $\text{stateValid} \leftarrow \text{isStateValid}(q)$ 
11:            if  $\text{stateValid}$  then
12:               $c'_m \leftarrow \text{calcNormManipulability}(q)$ 
13:               $c'_s \leftarrow \text{calcNormStability}(q, V_{COM})$ 
14:               $c_v \leftarrow \alpha c'_m + \beta c'_s$ 
15:               $\text{writeVoxelToDatabase}(x, y, z, r, \theta, \phi, c_v)$ 
16: Return  $\text{workspaceDataFile}$ 
```

distance of the hand separation is fixed according to the bimanual grasp of the object, the orientation of the hands with respect to each other matters based on the grasp of the object and certain objects have additional orientation constraints like how open top boxes or crates must be carried close to horizontal to avoid spilling the contents. Therefore, this section considers a bimanual workspace analysis that is object specific. The workspace generation procedure is similar to that of [Algorithm 3](#), where the space in front of the robot is discretized into a 6D voxelized workspace, except the voxel parameters are position and orientation. To compute the object specific workspace an object specific grasp must be defined to provide offsets for each TCP. This replaces the spherical coordinates used with the position-only IK of the general workspace analysis. Consequently, to meet the grasp requirements the IK must be full position and orientation. The use of orientation at the voxels then allows for different orientations of the object to be considered at that point in the workspace as it could be useful when manipulating the object for a certain task.

The details of the object specific workspace map generation is shown in [Algorithm 4](#). The input to the algorithm is the range of the voxel positions and orientations, the discretization steps, the velocity of the center of mass (\dot{r}_{COM}) and the TCP offsets based on the grasp and object geometry. The algorithm computes the metric c_v from [Equation 5.5](#).

In this workspace generation the kinematic grasp is being considered for the object, but also the dimensions of the object to consider collisions and the mass of the object at the TCPs to consider for stability. This provides a much more comprehensive treatment of manipulating an object than the general workspace analysis did. Note that the hands are still simplified for computational speed and only consider simplified fixed contact grasps. The objects and the grasps used by the robot can be seen in [Figure 1.2](#). This algorithm also does not exploit symmetry. While this could be performed for symmetric grasps if the robot was perfectly symmetric, the longer generation time is accepted to allow for the treatment of asymmetric grasps and more realistic robot models.

Algorithm 4 Object Bimanual Workspace Generation

Input: $x, y, z, roll, pitch, yaw$ ranges and discretization steps, \dot{r}_{CoM} , `object_geometry`

Parameters:

```

1: for  $x \in \{x_{min}, \dots, x_{max}\}$  do
2:   for  $y \in \{y_{min}, \dots, y_{max}\}$  do
3:     for  $z \in \{z_{min}, \dots, z_{max}\}$  do
4:       for  $roll \in \{roll_{min}, \dots, roll_{max}\}$  do
5:         for  $pitch \in \{pitch_{min}, \dots, pitch_{max}\}$  do
6:           for  $yaw \in \{yaw_{min}, \dots, yaw_{max}\}$  do
7:              $j \leftarrow \text{Vector}(x, y, z, roll, pitch, yaw)$ 
8:              $j_l, j_r \leftarrow \text{calcGraspOffsets}(\text{object\_geometry})$ 
9:              $q \leftarrow \text{calcBothArmsIK}(j + j'_l, j + j'_r)$ 
10:            stateValid  $\leftarrow \text{isStateValid}(q)$ 
11:            if stateValid then
12:               $c'_m \leftarrow \text{calcNormManipulability}(q)$ 
13:               $c'_s \leftarrow \text{calcNormStability}(q, \dot{r}_{CoM})$ 
14:               $c_v \leftarrow \alpha c'_m + \beta c'_s$ 
15:              writeVoxelToDatabase( $x, y, z, roll, pitch, yaw, c_v$ )

```

5.2.4 Workspace Visualization

To visualize the results of c_v for the various workspace analyses, the voxels in front of the robot use a colour spectrum. Low quality is shown in blue then green as medium quality to yellow-orange for high quality. Since not all voxels explored will be feasible as they may be unreachable or unstable, they are shown as light gray. Also, since only 3D voxels can be shown, the average of the multiple results is displayed at the voxel by overlapping the voxels and using a transparency.

5.3 Results

This section presents the results of the various workspace analysis scenarios with the proposed manipulability-stability metric for use with humanoid bimanual manipulation.

5.3.1 General Workspace Maps

First, this section studies the general workspace map for a 0.8 m^3 workspace with 5 cm voxel discretization. This workspace is the main area where the manipulation task of the robot would be performed, and no CoM velocity is applied in this case. In [Figure 5.3](#), the full workspace and a horizontal cut of the workspace can be seen for the metric c_v of varying weights to show the manipulability, stability and combined manipulability-stability quality metrics. The manipulability metric, where $\alpha = 1$ and $\beta = 0$ in [Equation 5.5](#), is shown in [Figure 5.3a](#) and [5.3d](#). It can be seen that the values closer to the robot are higher quality, or more manipulable, and the further from the robot the lower the quality, or less manipulable, they become. This is expected since the arms would be more extended and closer to singularity. The stability metric, where $\alpha = 0$ and $\beta = 1$ in [Equation 5.5](#), is seen in [Figure 5.3b](#) and [5.3e](#). In this case, the voxels near the robot are higher quality so more stable as expected since the CoM remains close to the center of the support polygon. The quality is lower near the edges as the CoM would have to shift closer to the edge of the support polygon to reach those poses. The combined manipulability-stability metric, where $\alpha = 0.5$ and $\beta = 0.5$ for equal weighting in [Equation 5.5](#), is displayed in [Figure 5.3c](#) and [5.3f](#). Again, the higher quality poses are for the voxels closer to the robot. The visualization for the combined metric provides a good overview of the locations where the most stable and manipulable configurations are in the workspace. Note that in the workspace generation REEM-C could reach 22% of the queried voxels in the workspace.

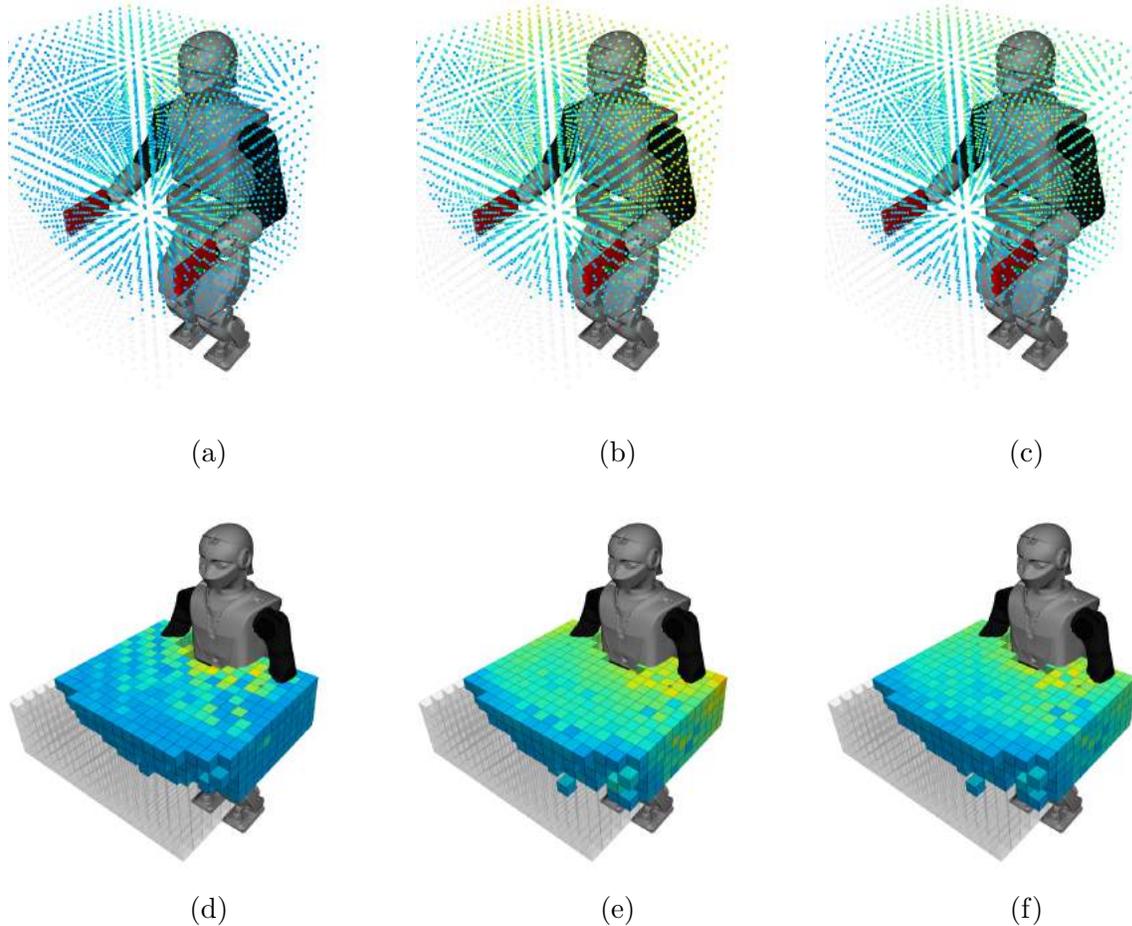
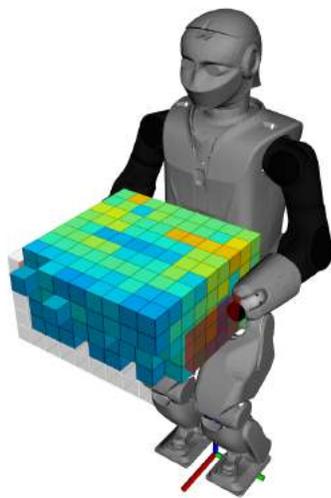


Figure 5.3: Workspace visualizations for the general generation method. Low quality voxels are blue and high quality voxels are yellow/orange. Voxels that are unreachable or unstable are translucent grey. The first row shows the entire workspace that was considered and the second row shows a horizontal cut through the workspace. a) and d) show the manipulability metric, b) and e) show the stability metric and c) and f) show the combined manipulability-stability metric as detailed in [Equation 5.5](#).

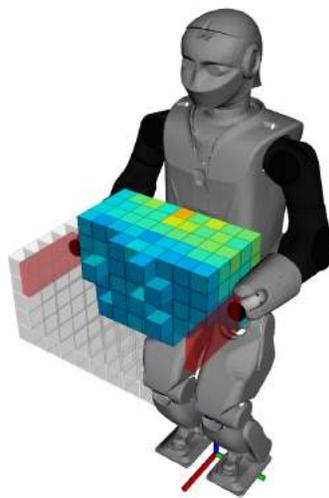
5.3.2 Workspace Maps with Varying Center of Mass Velocities

When developing the new metric, the stability metric used was based on the distance of the capture point to the nearest edge of the support polygon, which allows for CoM velocities

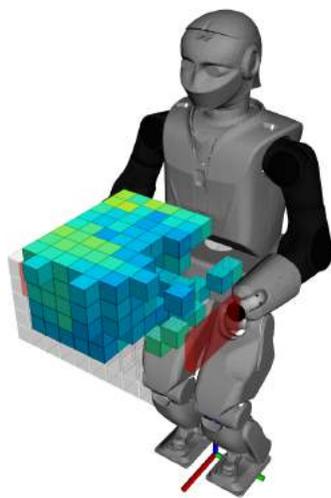
to be considered. Including CoM velocity in the workspace visualization helps provide a deeper understanding of the robot’s bimanual manipulation abilities while moving. A set of workspace analyses with different CoM velocities was generated to see how this affects stability and voxel reachability. To illustrate the effect of CoM velocities, this analysis considers a smaller subset of the general workspace presented earlier with four different CoM velocities to compare using the stability metric, where $\alpha = 0.0$ and $\beta = 1.0$ in [Equation 5.5](#). First, the analysis considers a workspace with no CoM velocity in [Figure 5.4a](#). In this case most of the voxels can be reached and the voxels closer to the robot are higher quality, so more stable. Applying a CoM velocity in the x -direction, as seen in [Figure 5.4b](#), causes a portion of the workspace to become unreachable because the robot is no longer stable in those positions as the capture point would leave the support polygon and require a step. In like manner, a portion of the workspace becomes unreachable when as CoM velocity in the y -direction is applied, as seen in [Figure 5.4c](#). Only a small portion of the workspace is reachable when a CoM velocity is applied in the x and y -directions, which is shown in [Figure 5.4d](#). This sort of analysis can be useful for bimanual manipulation scenarios beyond the simple static case as CoM velocity plays a key role in planning locomanipulation motions.



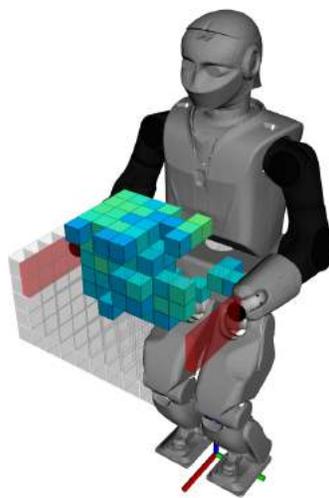
(a)



(b)



(c)



(d)

Figure 5.4: Various CoM velocity cases visualized for a subset of the general workspace with the stability metric. a) $\dot{r}_{CoM} = [0.0, 0.0, 0.0]$ b) $\dot{r}_{CoM} = [0.0, 0.2, 0.0]$ box c) $\dot{r}_{CoM} = [0.0, 0.6, 0.0]$ d) $\dot{r}_{CoM} = [0.2, 0.6, 0.0]$

5.3.3 Workspace Maps with Different Object Masses

Here, two workspace scenarios for box manipulation with REEM-C are analyzed. The two situations involve boxes of different dimensions, which will be considered for collision checking, and mass, which will be considered for stability.

The first box under consideration is 1 kg with dimensions $0.1\text{ m} \times 0.4\text{ m} \times 0.1\text{ m}$. A workspace of 0.8 m^3 with a voxel discretization of 5 cm was generated with slices shown in [Figure 5.5](#). It can be seen that a symmetric workspace is generated for a box manipulated with a symmetric grasp. Due to the weight of the box, the feasible workspace is reduced in the x -direction because the capture point leaves the support polygon at the outer voxels. In this case, 18% of the checked workspace was found to be feasible and is less than the result from the general workspace.

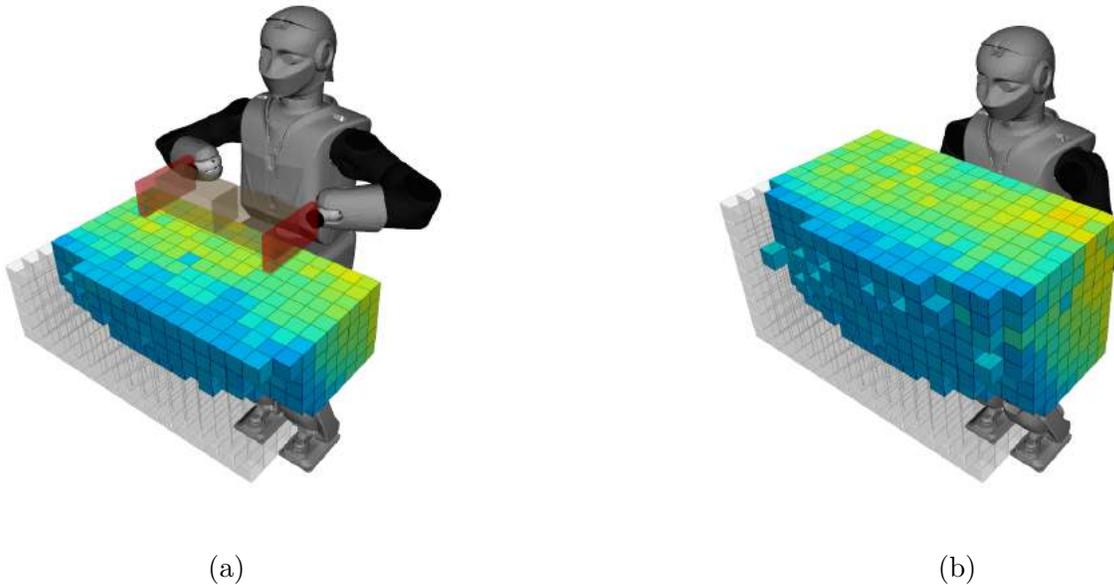


Figure 5.5: Visualization of the bimanual manipulability-stability metric for the REEM-C manipulating a 1 kg box. a) A slice at $z = 1.0\text{ m}$ b) A slice at $z = 1.2\text{ m}$.

The second box under consideration is 7 kg with dimensions $0.2\text{ m} \times 0.6\text{ m} \times 0.4\text{ m}$. In [Figure 5.6](#) it can be seen that the large mass significantly impacts the feasible workspace causing a large portion of the outer voxels to be infeasible. Compared to the first box, the quality of the feasible voxels is also lower due to the challenge of manipulating the larger

object as more collisions and unreachable poses occur. The checked workspace for this scenario yielded only 10% as feasible, a marked decrease from the first box. In [Figure 5.6a](#), the box is held at 0.2-0.3 m from the torso and the capture point, which is displayed between the feet in green, is very close to the edge of the support polygon, shown in blue.

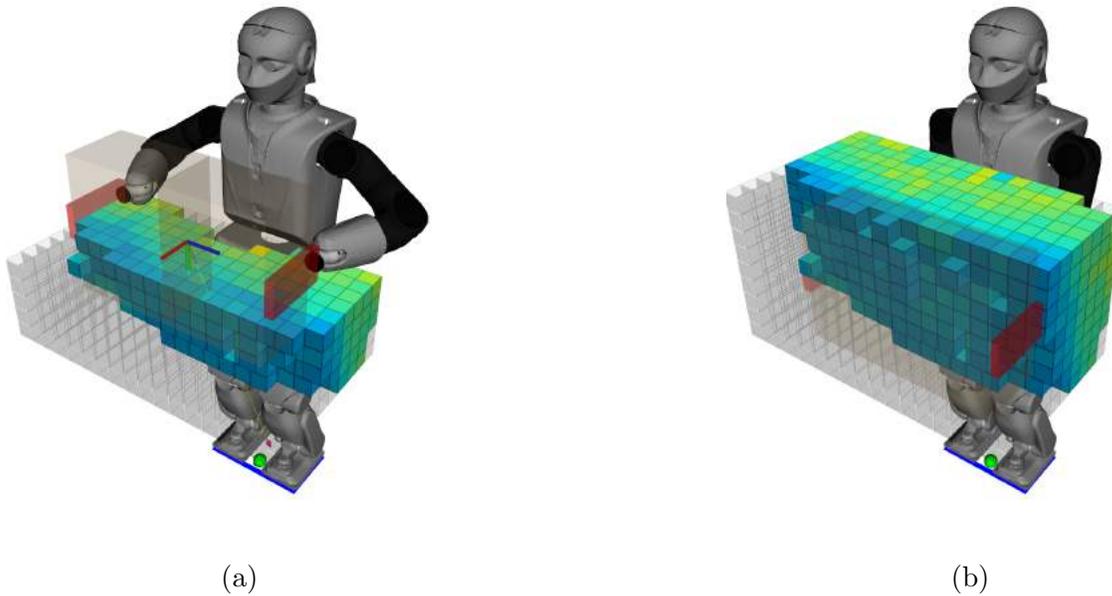


Figure 5.6: Visualization of the bimanual manipulability-stability metric for the REEM-C manipulating a 7 kg box. a) A slice at $z = 1.0 m$ b) A slice at $z = 1.2 m$.

5.3.4 Workspace Maps with Different Objects

Next, two other objects of different grasping requirements, a rolling pin and a broom, are considered.

The grasping configuration for a rolling pin requires gripped handles or a palms down hand contact with the rolling pin, which is how the grasp is represented for REEM-C's simplified hand contacts. The analysis uses a small workspace directly in front of the robot with all orientations limited to 7.5° . In [Figure 5.7a](#) the rolling pin workspace is shown at the slice height of $z = 0.9 m$ from the ground. Only the middle portion of the workspace is feasible. The area close to the robot is infeasible as self-collisions of the arms and torso occur and the area further from the robot near the edges requires the significant

arm extension and torso bending that would result in stability issues. Considering the height of $z = 1.1 \text{ m}$ in [Figure 5.7b](#), the workspace is much larger and the quality values of the middle region are higher. This results from the configuration of the robot with its elbows out helping to avoid self-collisions, so evidently the workspace for this grasp is significantly impacted by height. This would be important when selecting a table for REEM-C to perform a rolling pin task, as the height that allows for more highly stable and manipulable configurations would allow for much easier motion planning. Note that contact and force are not considered in this workspace analysis, only the kinematics for the constrained bimanual task.

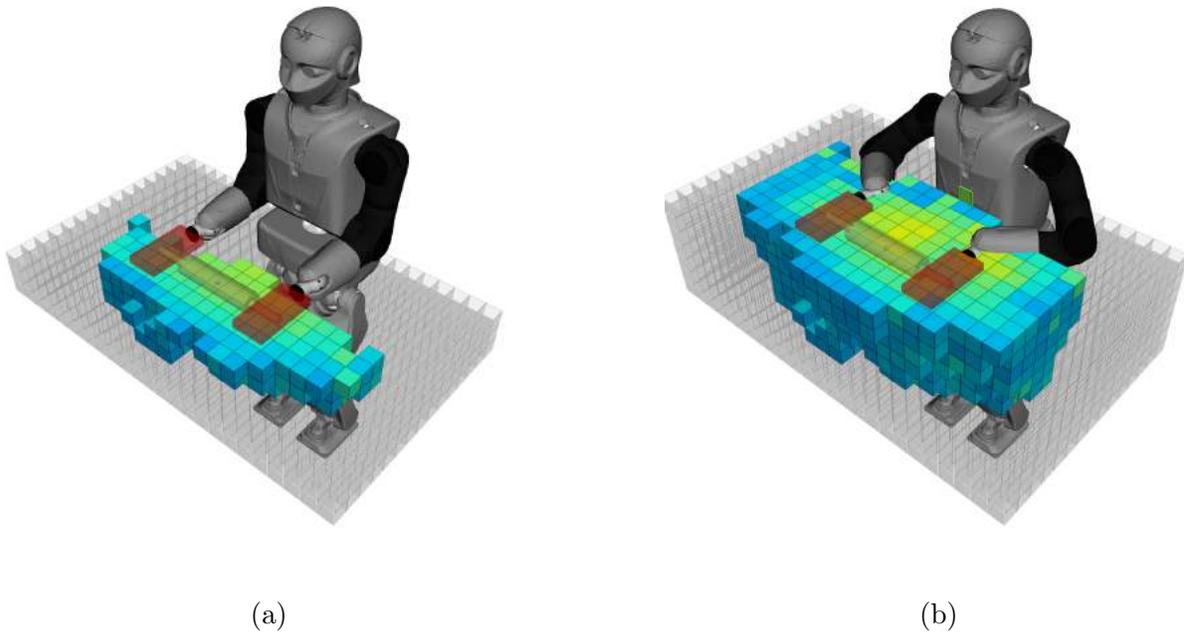


Figure 5.7: Visualization of the bimanual manipulability-stability distribution for REEM-C manipulating a rolling pin. a) A cut through at $z = 0.9 \text{ m}$ b) A cut through at $z = 1.1 \text{ m}$

In the case of a broom, the grasping configuration requires a certain handed pose as one hand is placed below the other. This analysis considers the left-handed grasp with the left hand below the right hand when grasping the broom. When grasping a broom the palms face in opposite directions, but in different locations along the broom handle. The workspace for the broom can be seen in [Figure 5.8](#). Due to the left-handed grasp, higher quality voxels and more feasible voxels exist of the left side and most of the right side is infeasible. This is sensible as reaching across the body to reach the robot's right side with

the left-handed grasp would result in collisions between the left arm and the torso. This helps inform the area the robot can easily manipulate the broom in a sweeping task and if the robot should turn or switch grips to sweep a different area.

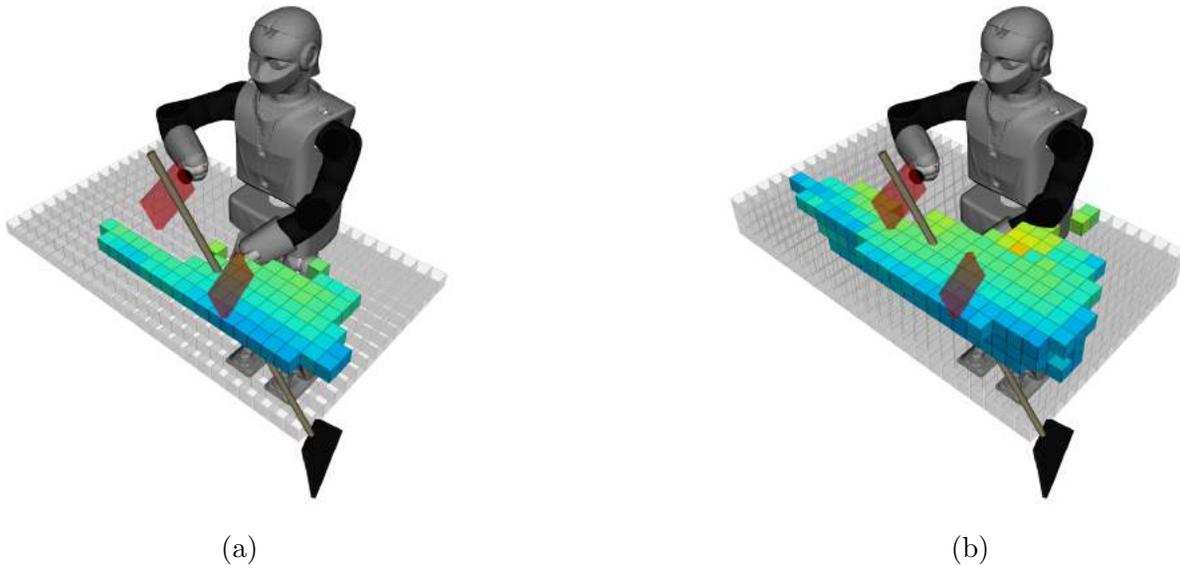
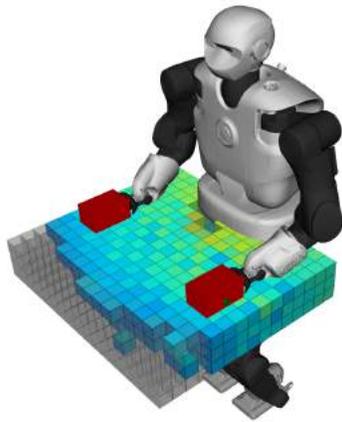


Figure 5.8: Visualization of the bimanual manipulability-stability distribution for REEM-C manipulating a broom. a) cut through at $z = 0.9 \text{ m}$ b) cut through at $z = 1.0 \text{ m}$

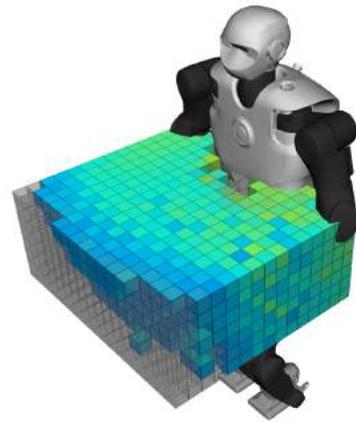
5.3.5 Workspace Maps of TALOS

The final workspaces considered use the humanoid robot TALOS. Similar to the general workspace performed with REEM-C, [Figure 5.9](#) shows the manipulability-stability metric for a general workspace using the spherical coordinates and the position-only IK for a workspace of 0.8 m^3 and 5 cm discretization. As expected, the higher quality voxels are closer to TALOS and the lower quality voxels are further from the robot. In this workspace generation 34% of the voxels were feasible compared to the 22% for REEM-C. This significant increase is owing to the longer arms and alternate shoulder configuration of TALOS, allowing for more range of motion than REEM-C.

In addition, the 1 kg box with dimensions $0.1 \text{ m} \times 0.4 \text{ m} \times 0.1 \text{ m}$ is visualized with TALOS in [Figure 5.10](#). Of the voxels checked 26% were feasible compared to the 18% with REEM-C again because TALOS is a larger robot with an alternate kinematic structure.

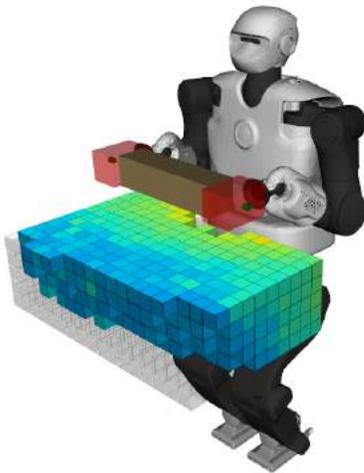


(a)

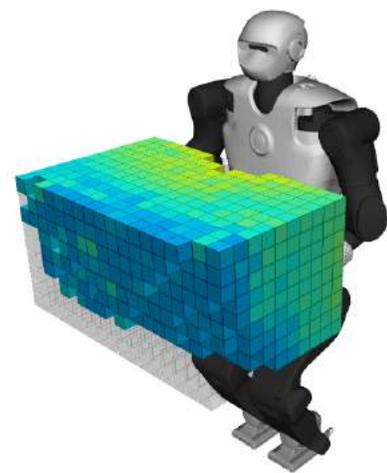


(b)

Figure 5.9: Visualization of the bimanual manipulability-stability metric with TALOS, for the general workspace generation with spherical coordinates. a) A slice at $z = 0.9 \text{ m}$ c) A slice at $z = 1.1 \text{ m}$.



(a)



(b)

Figure 5.10: Visualization of the bimanual manipulability-stability metric for the TALOS manipulating a 1 kg box. a) A slice at $z = 1.0 \text{ m}$ b) A slice at $z = 1.2 \text{ m}$.

5.3.6 Applications of the Combined Manipulability Metric and Workspace Visualization

Through these visualizations, the highly constrained nature of bimanual manipulation tasks became more apparent and what regions of the workspace were completely infeasible. A direct application of this visualization would be for the box manipulation benchmarking with REEM-C as described at the end of [Chapter 3](#). Not only could the visualizations help determine the feasibility of a desired motion, but the metric could even be added as a KPI to rate the quality of the manipulation being performed. Another application of these workspace visualizations could be to design or redesign workspaces to be more optimal for the robot. As some motions may be infeasible for the robot, the workspace could be designed to allow for the motions required to remain in the more stable and manipulable regions of the robot's workspace. Considering the example of picking and placing a box on a shelf, like in the whole-body manipulation benchmarking scenario, altering the pick and place location of the boxes to be at the edge of the shelf and using a shelf that was less deep would keep the robot from having to reach too far where the workspace becomes unreachable or unstable.

Chapter 6

Thesis Conclusions

6.1 Summary of Findings

This thesis developed two manipulation benchmarks for humanoid robots, one for whole-body manipulation and another for loco-manipulation. These two benchmarks were proposed as part of the EURO-BENCH unified robotic benchmarking framework by developing a standardized test bed, comprehensive protocols, and insightful key performance indicators to evaluate performance. A partial initial benchmark was performed for the whole-body benchmark using the REEM-C robot Seven to evaluate two different motion generation and control strategies where one used a separate controller approach of the upper and lower body and the other used a whole-body controller. The separate upper and lower body control approach was approximately 10 seconds faster but had a 2.5% placement accuracy rate that was 40% lower. Also, this control approach had a mechanical work five times that of the whole-body control approach, but half the cost of operation. This evaluation provided insights into trade-offs in accuracy and cost of operation for speed. The challenge of performing bimanual manipulation tasks was identified leading to the bimanual manipulation workspace analysis that followed. Similarly, a partial benchmark was performed for the loco-manipulation benchmark; however, testing was interrupted due to a damaged force torque sensor in the foot of Seven. This partial benchmark also compared two motion generations and control strategies where one uses ArUco marker detection for error correction and the other does not. Together these two manipulation benchmarks offer a well-defined, comprehensive approach to measuring the performance of humanoid robots and the motion generation algorithms in loco-manipulation scenarios.

Another development of this thesis was the bimanual manipulation workspace analysis

for humanoid robots. This workspace analysis developed a new quality metric to evaluate the bimanual manipulation of humanoids based on bimanual manipulability and capture point stability. This metric was then visualized for a general workspace generation method, based on the work in [43], and an object specific workspace generation method. In the general workspace of 0.8 m^3 with 5 cm voxel discretization 22% of the voxels were reachable with the REEM-C. The object specific workspace generation method allowed for the evaluation of different objects that impose different constraints on the bimanual manipulation task like collisions, mass, and grasp requirements. The objects considered were boxes of two different sizes, a rolling pin, and a broom. The workspace analysis also allowed for the impact of CoM velocity of the robot on the feasible workspace to be evaluated due to the use of the capture point based stability metric. An evaluation of the workspace of TALOS and REEM-C was performed in the same scenario showing that the size of TALOS and shoulder structure result in a workspace with more feasible configurations. For the general workspace case, TALOS had 34% voxels reachable compared to the 22% of the REEM-C and for the 1 kg box 26% of voxels were reachable compared to 18% with REEM-C. Using these visualizations and the new metric greater understanding for bimanual tasks and the feasible regions can be gained. The new metric also offers another possibility for evaluating the quality of a bimanual manipulation task being performed, such as for benchmarking.

6.2 Limitations

While comprehensive benchmarks were provided for whole-body manipulation and locomanipulation, only partial benchmarks could be performed to provide an initial benchmark of different motions. This helped provide insight into comparisons that could be made but does not provide the complete picture that completely performing the benchmark would. Unfortunately, the REEM-C robot Seven was broken for approximately 6 months limiting the testing that could be performed. Further work investigating motions and comparing different robots should be performed to validate the benchmarks.

For the bimanual manipulation workspace analysis, the workspace generation methods are based on an optimization method using EXOTica to perform whole-body inverse kinematics with many pose requirement checks to ensure the motion is feasible. One drawback of these methods was the long workspace generation time, which often took several hours to run with many voxels being infeasible. Another was the variation in some of the workspace results regarding voxel symmetry and smoothness of the coloured regions. As the generation uses an optimization approach necessary to handle the complex task requirements and is useful for tree-like kinematic structures such as those of humanoid robots, some config-

urations found could have been higher quality or even feasible though the solver may not have converged on that result. To improve the symmetry and smoothness of the figures, the optimization approach would have to be run for many more iterations to ensure the solver converges to the correct points. Both generation time and voxel symmetry could have been improved by exploiting symmetry in the algorithm; however, this trade-off was made to allow for use with asymmetric robot models and asymmetric grasps.

6.3 Future Work

Future work related to the benchmarking would be to perform the complete benchmarking protocol for the two manipulation benchmarks and calculate the key performance metrics to have a complete initial benchmark. This would involve adding the LED lights to the shelves for the variable protocols and adding a placement error tracking system like a vision based one with ArUCo markers. Also, to compute human likeness performance, human data of the motions being benchmarked is required, so human motion capture should be performed to have a complete human motion set for the benchmarks. One last area to investigate for the benchmarks in future work is to compare the performance of two different humanoid robots, like Waterloo’s REEM-C and TALOS.

With respect to the bimanual manipulation workspace analysis work, future work will involve the development of a cost-based planner that uses the combined manipulability-stability metric to plan motions for bimanual tasks with objects. This would allow for highly stable and manipulable motion plans to be created for bimanual manipulation tasks. Another topic for future work is exploring stability metrics using angular momentum with the zero moment point for the linear inverted pendulum model.

References

- [1] Eurobench project. <https://eurobench2020.eu/>. Accessed: 2022-07-15.
- [2] Koroibot project. <https://orb.iwr.uni-heidelberg.de/koroibot/>. Accessed: 2022-07-15.
- [3] Felix Aller, Monika Harant, and Katja Mombaur. Optimization of dynamic sit-to-stand trajectories to assess whole-body motion performance of the humanoid robot REEM-C. *Frontiers in Robotics and AI*, 9, 2022.
- [4] Felix Aller, David Pinto-Fernandez, Diego Torricelli, Jose Luis Pons, and Katja Mombaur. From the state of the art of assessment metrics toward novel concepts for humanoid robot locomotion benchmarking. *IEEE Robotics and Automation Letters*, 5(2):914–920, 2020.
- [5] Narcis Miguel Baños. *REEM-C 7 Training: "Walking"*. PAL Robotics, Barcelona, Spain, Sept. 16, 2020.
- [6] Patrick Beeson and Barrett Ames. Trac-ik: An open-source library for improved solving of generic inverse kinematics. In *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*, pages 928–935, 2015.
- [7] Yasemin Bekiroglu, Naresh Marturi, Máximo A. Roa, Komlan Jean Maxime Adjigble, Tommaso Pardi, Cindy Grimm, Ravi Balasubramanian, Kaiyu Hang, and Rustam Stolkin. Benchmarking protocol for grasp planning algorithms. *IEEE Robotics and Automation Letters*, 5(2):315–322, 2020.
- [8] Konstantinos Chatzilygeroudis, Bernardo Fichera, Ilaria Lauzana, Fanjun Bu, Kunpeng Yao, Farshad Khadivar, and Aude Billard. Benchmark for bimanual robotic manipulation of semi-deformable objects. *IEEE Robotics and Automation Letters*, 5(2):2443–2450, 2020.

- [9] Silvia Cruciani, Balakumar Sundaralingam, Kaiyu Hang, Vikash Kumar, Tucker Hermans, and Danica Kragic. Benchmarking in-hand manipulation. *IEEE Robotics and Automation Letters*, 5(2):588–595, 2020.
- [10] Defence Advanced Research Projects Agency. Darpa robotics challenge. <https://www.darpa.mil/program/darpa-robotics-challenge>. Accessed: 2022-07-15.
- [11] C. R. G. Dreher, Mirko Wächter, and Tamim Asfour. Learning object-action relations from bimanual human demonstration using graph networks. *IEEE Robotics and Automation Letters*, 5(1):187–194, 2020.
- [12] Paolo Ferrari, Marco Cognetti, and Giuseppe Oriolo. Humanoid whole-body planning for loco-manipulation tasks. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4741–4746, 2017.
- [13] Jianfeng Gao, You Zhou, and Tamim Asfour. Projected force-admittance control for compliant bimanual tasks. In *2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids)*, pages 1–9, 2018.
- [14] Gazebo. <https://gazebo.org/home>. Accessed: 2022-07-28.
- [15] Kensuke Harada, Shuuji Kajita, Kenji Kaneko, and Hirohisa Hirukawa. Dynamics and balance of a humanoid robot during manipulation tasks. *IEEE Transactions on Robotics*, 22(3):568–575, 2006.
- [16] Enrico Mingo Hoffman, Alessio Rocchi, Arturo Laurenzi, and Nikos G. Tsagarakis. Robot control for dummies: Insights and examples using opensot. In *2017 IEEE-RAS 17th International Conference on Humanoid Robotics (Humanoids)*, pages 736–741, 2017.
- [17] Ivar shelf unit, pine, 35x19 5/8x70 1/2 in (89x50x179 cm). <https://www.ikea.com/ca/en/p/ivar-shelf-unit-pine-s69251345/>. Accessed: 2022-07-28.
- [18] Vladimir Ivan, Yiming Yang, Wolfgang Merkt, Michael P. Camilleri, and Sethu Vijayakumar. EXOTica: An Extensible Optimization Toolset for Prototyping and Benchmarking Motion Planning and Control. In Anis Koubaa, editor, *Robot Operating System (ROS)*, volume 778, pages 211–240. Springer International Publishing, Cham, 2019. Series Title: Studies in Computational Intelligence.
- [19] Peter Kaiser, David Gonzalez-Aguirre, Fabian Schültje, Júlia Borràs, Nikolaus Vahrenkamp, and Tamim Asfour. Extracting whole-body affordances from multimodal

- exploration. In *2014 IEEE-RAS International Conference on Humanoid Robots*, pages 1036–1043, 2014.
- [20] Peter Kaiser, Nikolaus Vahrenkamp, Fabian Schültje, Júlia Borràs, and Tamim Asfour. Extraction of whole-body affordances for loco-manipulation tasks. *International Journal of Humanoid Robotics*, 12:1550031, 06 2015.
- [21] S. Kajita, F. Kanehiro, K. Kaneko, K. Fujiwara, K. Harada, K. Yokoi, and H. Hirukawa. Biped walking pattern generation by using preview control of zero-moment point. In *2003 IEEE International Conference on Robotics and Automation (Cat. No.03CH37422)*, volume 2, pages 1620–1626 vol.2, 2003.
- [22] Shuuji Kajita, Hirohisa Hirukawa, Kensuke Harada, and Kazuhito Yokoi. *Introduction to Humanoid Robotics*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014.
- [23] Franziska Krebs, Andre Meixner, Isabel Patzer, and Tamim Asfour. The kit bimanual manipulation dataset. In *2020 IEEE-RAS 20th International Conference on Humanoid Robots (Humanoids)*, pages 499–506, 2021.
- [24] J.J. Kuffner and S.M. LaValle. Rrt-connect: An efficient approach to single-query path planning. In *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No.00CH37065)*, volume 2, pages 995–1001 vol.2, 2000.
- [25] Arturo Laurenzi, Enrico Mingo Hoffman, Luca Muratore, and Nikos G. Tsagarakis. Cartesi/o: A ros based real-time capable cartesian control framework. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 591–596, 2019.
- [26] S. Lee. Dual redundant arm configuration optimization with task-oriented dual arm manipulability. *IEEE Transactions on Robotics and Automation*, 5(1):78–97, 1989.
- [27] Junjia Liu, Yiting Chen, Zhipeng Dong, Shixiong Wang, Sylvain Calinon, Miao Li, and Fei Chen. Robot cooking with stir-fry: Bimanual non-prehensile manipulation of semi-fluid objects. *IEEE Robotics and Automation Letters*, 7(2):5159–5166, 2022.
- [28] Kevin M. Lynch and Frank C. Park. *Modern Robotics: Mechanics, planning, and Control*. Cambridge University Press, Cambridge, United Kingdom, 2019.
- [29] Nicolas Mansard, Oussama Khatib, and Abderrahmane Kheddar. A unified approach to integrate unilateral constraints in the stack of tasks. *IEEE Transactions on Robotics*, 25(3):670–685, 2009.

- [30] Nicolas Mansard, Olivier Stasse, Paul Evrard, and Abderrahmane Kheddar. A versatile generalized inverted kinematics implementation for collaborative working humanoid robots: The stack of tasks. pages 1 – 6, 07 2009.
- [31] Andrew S. Morgan, Kaiyu Hang, Walter G. Bircher, Fadi M. Alladkani, Abhinav Gandhi, Berk Calli, and Aaron M. Dollar. Benchmarking cluttered robot pick-and-place manipulation with the box and blocks test. *IEEE Robotics and Automation Letters*, 5(2):454–461, 2020.
- [32] Federico L. Moro and Luis Sentis. *Whole-Body Control of Humanoid Robots*, pages 1161–1183. Springer Netherlands, Dordrecht, 2019.
- [33] Moveit! motion planning framework. <https://moveit.ros.org/>. Accessed: 2022-07-15.
- [34] Masaki Murooka, Iori Kumagai, Mitsuharu Morisawa, Fumio Kanehiro, and Abderrahmane Kheddar. Humanoid loco-manipulation planning based on graph search and reachability maps. *IEEE Robotics and Automation Letters*, 6(2):1840–1847, 2021.
- [35] Jerry Pratt, John Carff, Sergey Drakunov, and Ambarish Goswami. Capture point: A step toward humanoid push recovery. In *2006 6th IEEE-RAS International Conference on Humanoid Robots*, pages 200–207, 2006.
- [36] Jerry E. Pratt, Sylvain Bertrand, and Twan Koolen. *Stepping for Balance Maintenance Including Push-Recovery*, pages 1419–1466. Springer Netherlands, Dordrecht, 2019.
- [37] Alessio Rocchi, Enrico Mingio Hoffman, Darwin G. Caldwell, and Nikos G. Tsagarakis. Opensot: A whole-body control library for the compliant humanoid robot coman. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6248–6253, 2015.
- [38] Keli Shen, Xiang Li, Hongzhi Tian, Takayuki Matsuno, and Mamoru Minami. Analyses of biped walking posture by dynamical-evaluating index. *Artificial Life and Robotics*, 23(2):261–270, 2018.
- [39] Panagiotis Sotiropoulos, Maximo A. Roa, Murilo F. Martins, Werner Fried, Hussein Mnyusiwalla, Pavlos Triantafyllou, and Graham Deacon. A benchmarking framework for systematic evaluation of compliant under-actuated soft end effectors in an industrial context. In *2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids)*, pages 280–283, 2018.

- [40] Olivier Stasse, Kevin Giraud-Esclasse, Edouard Brousse, Maximilien Naveau, Rémi Régnier, Guillaume Avrin, and Philippe Souères. Benchmarking the hrp-2 humanoid robot during locomotion. *Frontiers in Robotics and AI*, 5, 2018.
- [41] Ioan A. Şucan, Mark Moll, and Lydia E. Kavraki. The Open Motion Planning Library. *IEEE Robotics & Automation Magazine*, 19(4):72–82, December 2012. <https://ompl.kavrakilab.org>.
- [42] Diego Torricelli, Jose Gonzalez-Vargas, Jan F. Veneman, Katja Mombaur, Nikos Tsagarakis, Antonio J. del Ama, Angel Gil-Agudo, Juan C. Moreno, and Jose L. Pons. Benchmarking bipedal locomotion: A unified scheme for humanoids, wearable robots, and humans. *IEEE Robotics Automation Magazine*, 22(3):103–115, 2015.
- [43] Nikolaus Vahrenkamp and Tamim Asfour. Representing the robot’s workspace through constrained manipulability analysis. *Autonomous Robots*, 38:1–14, 01 2014.
- [44] Nikolaus Vahrenkamp, Tamim Asfour, and Rüdiger Dillmann. Robot placement based on reachability inversion. In *2013 IEEE International Conference on Robotics and Automation*, pages 1970–1975, 2013.
- [45] Nikolaus Vahrenkamp, Tamim Asfour, Giorgio Metta, Giulio Sandini, and Rüdiger Dillmann. Manipulability analysis. In *2012 12th IEEE-RAS International Conference on Humanoid Robots (Humanoids 2012)*, pages 568–573, 2012.
- [46] Nikolaus Vahrenkamp, Markus Przybylski, Tamim Asfour, and Rüdiger Dillmann. Bimanual grasp planning. In *2011 11th IEEE-RAS International Conference on Humanoid Robots*, pages 493–499, 2011.
- [47] Yiming Yang, Vladimir Ivan, Zhibin Li, Maurice Fallon, and Sethu Vijayakumar. idrm: Humanoid motion planning with realtime end-pose selection in complex environments. In *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*, pages 271–278, 2016.
- [48] Yiming Yang, Wolfgang Merkt, Henrique Ferrolho, Vladimir Ivan, and Sethu Vijayakumar. Efficient humanoid motion planning on uneven terrain using paired forward-inverse dynamic reachability maps. *IEEE Robotics and Automation Letters*, 2(4):2279–2286, 2017.