

# Humanoid Locomotion and Manipulation: Current Progress and Challenges in Control, Planning, and Learning

Zhaoyuan Gu<sup>1</sup>, Junheng Li<sup>2</sup>, Wenlan Shen<sup>3</sup>, Wenhao Yu<sup>4</sup>, Zhaoming Xie<sup>5</sup>, Stephen McCrary<sup>6</sup>, Xianyi Cheng<sup>7</sup>, Abdulaziz Shamsah<sup>1,8</sup>, Robert Griffin<sup>6</sup>, C. Karen Liu<sup>9</sup>, Abderrahmane Kheddar<sup>10,11</sup>, Xue Bin Peng<sup>12</sup>, Yuke Zhu<sup>13,14</sup>, Guanya Shi<sup>15</sup>, Quan Nguyen<sup>2</sup>, Gordon Cheng<sup>3</sup>, Huijun Gao<sup>\*16</sup>, and Ye Zhao<sup>\*1</sup>

<sup>1</sup>Georgia Institute of Technology <sup>2</sup>The University of Southern California <sup>3</sup>Technische Universität München  
<sup>4</sup>Google DeepMind <sup>5</sup>The AI Institute <sup>6</sup>The Institute for Human and Machine Cognition <sup>7</sup>Duke University  
<sup>8</sup>Kuwait University <sup>9</sup>Stanford University <sup>10</sup>CNRS-University of Montpellier LIRMM <sup>11</sup>CNRS-AIST Joint  
 Robotics Laboratory <sup>12</sup>Simon Fraser University <sup>13</sup>The University of Texas at Austin <sup>14</sup>NVIDIA <sup>15</sup>Carnegie  
 Mellon University <sup>16</sup>Harbin Institute of Technology

**Abstract**—Humanoid robots hold great potential to perform various human-level skills, involving unified locomotion and manipulation in real-world settings. Driven by advances in machine learning and the strength of existing model-based approaches, these capabilities have progressed rapidly, but often separately. This survey offers a comprehensive overview of the state-of-the-art in humanoid locomotion and manipulation (HLM), with a focus on control, planning, and learning methods. We first review the model-based methods that have been the backbone of humanoid robotics for the past three decades. We discuss contact planning, motion planning, and whole-body control, highlighting the trade-offs between model fidelity and computational efficiency. Then the focus is shifted to examine emerging learning-based methods, with an emphasis on reinforcement and imitation learning that enhance the robustness and versatility of locomotion and manipulation skills. Furthermore, we assess the potential of integrating foundation models with humanoid embodiments to enable the development of generalist humanoid agents. This survey also highlights the emerging role of tactile sensing, particularly whole-body tactile feedback, as a crucial modality for handling contact-rich interactions. Finally, we compare the strengths and limitations of model-based and learning-based paradigms from multiple perspectives, such as robustness, computational efficiency, versatility, and generalizability, and suggest potential solutions to existing challenges.

**Index Terms**—Humanoid robotics, Loco-manipulation, Model predictive control, Whole-body control, Imitation learning, Foundation models, and Whole-body tactile sensing.

## I. INTRODUCTION

A humanoid robot refers to any anthropomorphic robot that resembles the form of a human [7]. Typically, a humanoid robot possesses a torso, two arms, and two legs, though the degree of anthropomorphism may vary. For instance, some humanoid robots feature simple hands or wheeled legs [8]. The primary focus of this review is on humanoid robots that replicate human morphologies and functionalities, rather than those that closely mimic human-like visual appearance and external aesthetics. Specifically, this survey examines bipedal humanoid robots, with an emphasis on the whole-body motion of humanoid robots instead of on the dexterous manipulation of multi-finger hands.

\*co-corresponding authors

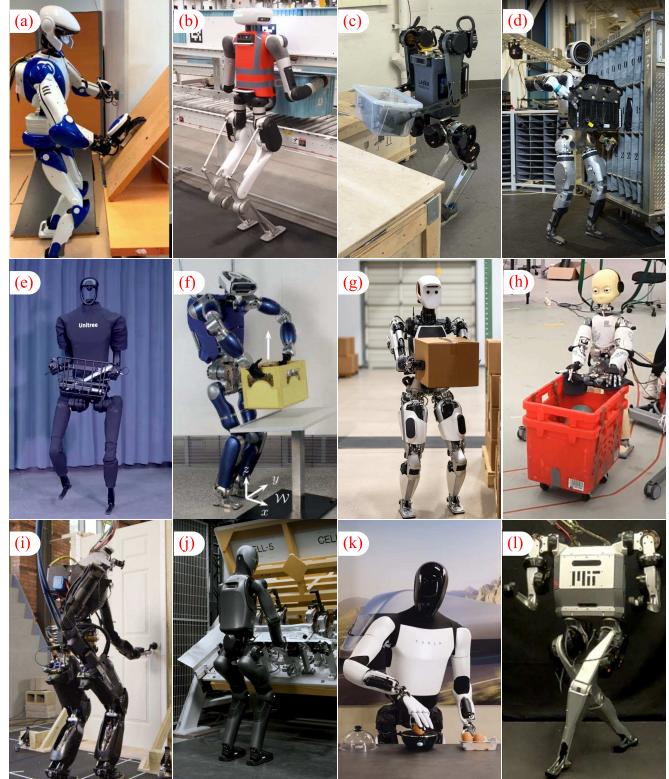


Fig. 1. Humanoids executing locomotion and manipulation tasks: (a) HRP-4 wipes a wood board while adapting to terrain [1]; (b-g) Object pick and place by Digit, Hector [2], Atlas, H1, TORO [3], and Apollo; (h) iCub pushes a cart [4]; (i) Nadia opens a door [5]; (j-k) Object manipulation by Figure 02 and Optimus; (l) MIT humanoid whole-body push recovery [6].

Humanoid robots are well suited for executing human-level tasks, as they are built to (ideally) replicate human motions and achieve various whole-body loco-manipulation tasks, *e.g.*, applications ranging from manufacturing to services, as shown in Fig. 1. Their anthropomorphism makes them stand out from other robot forms in performing human-like tasks, such as payload carrying over stairs, especially in environments designed for humans. Furthermore, humanoid robots can interact with humans for physical collaboration tasks, such as collaboratively moving a large table upstairs with human

assistance. However, simultaneously achieving these intricate tasks while addressing highly complex robot dynamics is still challenging, let alone safe physical collaborations with humans and/or operations in unstructured environments. As a promising direction to solve this problem, humanoids can exploit the abundance of data collected from humans to quickly acquire motor and cognitive skills due to their human-like morphology. Therefore, scaling human data and computational resources for humanoid embodiment is potentially a fast route to embodied intelligence.

Embodied intelligence at large is thriving at an unprecedented pace. Perception algorithms can detect, classify, and segment a wide variety of objects in real time. Model-based methods that leverage both predictive and reactive control have enabled agile and reliable locomotion and manipulation. Meanwhile, deep learning policies have demonstrated convincing control results on robot hardware through exploration and imitation. Large foundation models trained on massive, internet-scale datasets are beginning to show cognitive capabilities of open-world reasoning. Consequently, the building of autonomous humanoid robots for real-world applications has become possible, leading to the emergence of many humanoid robot companies and concrete deployment applications. Especially, companies with powerful GPU-based parallelization capabilities, such as NVIDIA, and companies building physical humanoid robots, such as Boston Dynamics, Tesla, and Figure, have begun collaboration on the embodied intelligence of humanoid robots.

Acknowledging the rapid advancements in humanoid robotics, this article reviews recent developments in Humanoid Locomotion and Manipulation (HLM). As laid out in Fig. 2, humanoid robotics is a multidisciplinary field that spans domains in design, actuation, sensing, control, planning, and decision-making. In this survey, we mainly examine task planning, motion planning, policy learning, and control from the perspective of model-based and learning-based methods. We first synthesize traditional model-based methods for planning and control. Then, we shift our focus to more recent learning-based approaches, especially those leveraging reinforcement learning, imitation learning, and foundation models.

Model-based methods serve as the cornerstone for enabling HLM capabilities. These methods depend critically on physical models, which can significantly influence the quality, speed, and guarantees of motion generation and control. Over the past decade, planning and control techniques have shown a trend of converging to the predictive-reactive control hierarchy, employing a whole-body model predictive controller (MPC) or simplified model (centroidal dynamics) MPC coupled with local task-space Whole-Body Controllers (WBC) [9]. These planning and control techniques are usually formulated as Optimal Control Problems (OCPs) that are solved by off-the-shelf or customized numerical solvers. Although these numerical optimization methods are well-established, research continues to focus on enhancing their computational efficiency, numerical stability, robustness, and scalability for high-dimensional systems.

Learning-based approaches have witnessed a rapid surge in humanoid robotics and achieved impressive results that attract

an increasing number of researchers to the field. Among the diverse learning approaches, Reinforcement Learning (RL) has proven its ability to achieve robust motor skills. However, despite its ability to discover novel behaviors via trial and error, pure RL without demonstration data is often prohibitively inefficient for HLM tasks, which are characterized by high degrees-of-freedom robots and sparse reward settings. Therefore, sim-to-real RL, where training is conducted in simulation before transferring to the physical robot, has become a prevalent method, though it faces the challenge of bridging sim-to-real gaps. From the existing literature, we observe that sim-to-real RL has gained significant traction for achieving humanoid whole-body motion. There is an ongoing debate between model-based and learning-based control, and therefore we highlight two key points: (1) sim-to-real RL that relies on the accurate dynamics models in a simulator can also be viewed as a “model-based,” method, even though the RL algorithm itself does not explicitly model the dynamics; and (2) model-based control and sim-to-real RL do not conflict; on the contrary, they often complement each other and can be combined to achieve a better performance than either method alone.

Imitation learning (IL) from expert demonstrations has proven to be an efficient method of acquiring motor skills. IL techniques, such as behavior cloning [10], have shown impressive abilities to mimic a wide range of skills. In pursuit of versatile and generalizable policies through IL, many researchers and companies have focused on scaling data. While robot experience data can be diverse and high-quality, its acquisition is both expensive and time-consuming. Thus, learning from human data, which is abundant and readily available from Internet videos and public datasets, emerges as a pivotal strategy for humanoid robotics. Learning from humans is a unique advantage exclusive to humanoid robots. However, even though humanoid robots may attain human-level motor skills, a deeper question of embodied intelligence persists: how to learn the intentions (source) behind human actions rather than merely replicating the observed motions (outcome). It is hypothesized that understanding human intention is achieved via Foundation Models (FMs) that are capable of semantic interpretation of the environment and the task. This hypothesis motivates us to include FMs as a part of our survey.

The remarkable success of FMs has sparked a surge in *general* robotics research, as highlighted by several comprehensive surveys [11, 12]. This paper surveys the application of FMs specifically for humanoid robots. FMs offer a promising solution to the persistent challenge of generalizability in robotics by efficiently harnessing internet-scale datasets to acquire extensive knowledge. A pre-trained FM exhibits capabilities for open-world reasoning and multimodal semantic understanding. These capabilities are invaluable for robots engaged in complex environments requiring long-term, logically coherent task planning. Within the realm of humanoid robots, FMs have been successfully implemented as task planning modules in hierarchical planning and control frameworks. Training end-to-end FMs for humanoid applications is an emerging area, poised for significant future advancements.



Fig. 2. This survey begins by defining relevant concepts of humanoid robots and their locomotion and manipulation capabilities. Centered around achieving humanoid loco-manipulation tasks, the core of this survey delves into two main categories of methods: the traditional planning and control approaches, such as contact planning, motion planning, and control, as well as the emerging learning-based approaches, including skill learning and foundation models. In addition, this survey highlights whole-body tactile sensing as a crucial modality to achieve contact-rich loco-manipulation.

### A. Survey Roadmap

For the selection criteria of papers cited in this survey, we prioritize topics most relevant to humanoid loco-manipulation capabilities, including tactile sensing [13], model-based planning [14] and control [9], as well as learning-based methods such as RL [15, 16], IL [17, 18], and FMs [11, 12]. Distinct from survey papers on these topics, our paper aims to present a broad overview with a focus on humanoid loco-manipulation.

Each of these topics has extensive studies. For each topic, we aim to spotlight three categories of papers: (i) foundational studies that established key principles in these areas, (ii) recent works representing the state of the art, and (iii) survey papers for further reading. Our survey spans 30 years of research. Specifically, model-based planning and control papers have been collected since 1990, and learning-based methods predominantly span the past 20 years.

Compared with existing survey reviews on humanoid robotics [7, 19], this survey collects up-to-date papers and introduces in detail the new paradigm of learning-based methods.

As shown in Fig. 2, this survey is organized in the following order. We first establish the background, defining humanoid robots and the key capability of locomotion and manipulation

in Sec. II. We detail whole-body tactile sensing in Sec. III. We then present traditional approaches that achieve loco-manipulation, including contact planning (Sec. IV), motion planning (Sec. V), and control (Sec. VI).

We then examine the state-of-the-art learning-based algorithms. In Sec. VII, we explore approaches using reinforcement learning and imitation learning to acquire loco-manipulation skills. In Sec. VIII, we discuss how foundation models have become the backbone of semantic understanding and decision-making for effective humanoid task planning. Finally, we highlight significant challenges in this field and present our perspectives on potential future research directions and emerging opportunities in Sec. IX.

### B. Survey Contributions

We summarize our core contributions as follows:

- The survey serves as an effective resource for graduate students and researchers new to the field, offering a comprehensive review of humanoid technical methods, while also providing perspectives for humanoid experts in academia and industry with the latest advancements.
- We advance the understanding of loco-manipulation in bipedal humanoid robots by clearly defining the prob-

lem, surveying state-of-the-art methods and capabilities, and outlining remaining challenges along with promising solutions.

- We provide a clear comparison between model-based and learning-based methods, and their corresponding loco-manipulation skills achieved on humanoid robots. This comparison offers valuable insights for researchers deciding their methods.

## II. BACKGROUND

In this section, we focus on the humanoid robot's main capabilities: bipedal locomotion and whole-body manipulation. Finally, we detail the combined loco-manipulation skills with the state-of-the-art methods and current challenges.

### A. Bipedal Locomotion

Bipedal locomotion is a significant characteristic of humanoid robots. Therefore, in the past three decades, bipedal locomotion has been a prolific field of humanoid research. Interested readers can refer to the excellent reviews (most of which are recent) [20, 21, 22], and the monographs [23, 24]. In summary, model-based bipedal locomotion has evolved significantly, progressing from passive walking [25, 26] to quasi-static walking [27], and then to dynamic walking [23]. Bipedal walking on flat surfaces has been well explored and mastered through periodic motions with model-based methods [23, 28]. These approaches have also expanded to more agile motions such as jumping [29, 30] and back-flipping [31].

Bipedal locomotion under external perturbations and force loads has been extensively studied. Such capabilities lay the foundation for simultaneous locomotion and manipulation, the focus of this survey. Model-based methods are among the early works that demonstrate these capabilities. For example, [32] introduced a passivity-based controller with task space dynamics that treats external forces as part of the robot's generalized forces. [2] incorporated a payload into a simplified rigid body model to enable dynamic walking while carrying a load, and [33] integrated external forces into a simplified linear inverted pendulum model.

In addition to model-based methods, bipedal locomotion has also been successfully addressed by learning-based methods [34, 35, 36], particularly in the context of periodic motions on flat surfaces. Furthermore, learning-based approaches have also demonstrated capabilities in more complex settings, such as running [37], jumping [38], and handling non-periodic motions such as stair climbing [39] and parkour [40]. Similar to the trend in model-based methods, learning-based methods have further extended their capabilities to handle external forces and payloads [41, 42].

Researchers have explored energy efficiency in bipedal locomotion by examining the Cost of Transport (COT), which designates energy expenditure per unit distance traveled, normalized by body weight. Early work in passive dynamic walking [25, 26] achieved higher efficiency than humans but offered limited versatility. Meanwhile, today's humanoid robots ( $COT > 0.7$ ) are significantly less efficient than humans ( $COT = 0.2$ ) [43]. Therefore, efficient walking continues to

motivate hardware designs that incorporate passive-compliant components for energy storage [44, 45] and control algorithms that exploit the system's natural dynamics [46, 47].

### B. Bipedal Navigation

Proficiency in bipedal locomotion has naturally progressed to humanoid robots' ability to effectively navigate complex environments, including indoor and outdoor areas with uneven terrain and dynamic obstacles. A navigation stack often incorporates a hierarchical structure: a global path planner and a local step planner. The global path planner [48, 49, 50, 51, 52] is typically responsible for understanding the overall navigation task and generating a path that avoids obstacles and reaches the target location. On the other hand, local step planners, *e.g.*, [53, 54, 55] focus on determining the precise foot placements that adhere to the bipedal dynamics within the immediate surroundings of the robot while also tracking the global path.

From the aforementioned navigation stack, bipedal navigation capabilities have progressed from static obstacle avoidance on flat terrain [56] to more challenging scenarios, including locomotion through height-constrained space [51, 57], avoiding dynamic obstacles in a constrained environment [58], navigating dynamic social environments using reachable sets [59], or deep RL that incorporates the perceived pedestrian emotions into the navigation plan [60], and traversing rough terrains [48, 49, 50, 61, 62, 63]. A persistent challenge for these methods is that they are tailored to specific use case scenarios and lack the versatility to handle a wide range of different situations.

While bipedal locomotion and navigation have been widely studied, real-world deployment remains a significant challenge due to inherent uncertainties. Uncertainty can arise from the environment and the robot model. Real-world environments have uneven, varying terrain, dynamic obstacles, and occlusion, making it difficult to ensure the safety and robustness of bipedal navigation. On the other hand, model uncertainty arises from discrepancies in the mathematical representation of the robot model and the physical system. Model uncertainty also exists in most current navigation frameworks that employ reduced-ordered models at the high level for collision avoidance and goal-reaching tasks and a full-order model at the low level for tracking high-level commands. A coupled framework that considers both the navigation task and whole-body control stability and accuracy is under-explored. Although previous works have addressed various aspects of environment uncertainties [64] and model uncertainties [65], a comprehensive navigation stack capable of handling the full spectrum of real-world uncertainties is still essential.

### C. Whole-body Manipulation

Anthropomorphic manipulation has been the inspiration for bimanual manipulation [66], loco-manipulation, and dexterous manipulation [67]. The ultimate form of anthropomorphic manipulation is whole-body manipulation, referring to the ability to manipulate objects using any part of one's body. For example, humans use their elbows or hips to hold a

door open for convenience; humans use their palms or fists instead of fingertips to provide large forces; humans curl their little fingers to hold a small object while still using other fingers for manipulation. In comparison, most robots often have predefined end-effectors, such as foot soles or fingertips, as the only parts allowed to physically interact with the world. Whole-body manipulation is a grand problem that shares challenges in bimanual manipulation, loco-manipulation, and dexterous manipulation. This general ability has yet to be developed, but its emergence will indicate a breakthrough for robotic manipulation.

The idea of whole-body manipulation was originally studied in the whole-arm manipulation community [68]. Whole arm manipulators were designed and built to explore the benefit of manipulating objects with all surfaces of a robot manipulator [69]. This brings a unique challenge that manifests itself in all the system levels in perception, estimation, planning, and control. Since there are an infinite number of such contacts, the planning complexity suffers from the combinatorial explosion of contact modes [70] and exponential computational costs from the high degree of freedom of the system [66].

Numerous breakthroughs in mechanical design, control, and planning have been achieved in the endeavor to address the challenges of whole-body manipulation. On the mechanical design side, robots made with soft materials and full-body sensing, such as Punyo [71], provide whole-body manipulation capability in a built-in manner.

For control, the coordinative and contact-rich nature requires forceful and compliant control. Traditionally, robot arms were hard-coded to switch across different control strategies according to task requirements [66]. Different task requirements, such as reaching a point or wiping a table, require different control strategies, such as pure position control or hybrid force position control. However, it is still unclear how to define and enumerate the control strategies for whole-body manipulation. In addition, a general control framework that can take in the sensor data, perform state estimation, and reactively control each body contact has yet to exist [72]. Such general frameworks require innovations in advanced hardware and algorithm architecture, including whole-body sensing [73] and robot designs with compliance and force control capabilities for reactive manipulation [74].

From the planning perspective, the challenge of whole-body manipulation can be potentially alleviated via human behavior imitation algorithms [75, 76, 77]. Most of these works focus on simple manipulation strategies such as whole-body grasping and pushing. To enable the robot to mimic more complex human whole-body manipulation behaviors, it is important to address the cross-morphology gap between humans and humanoids.

To achieve humanoid whole-body manipulation, full-stack system integration at all system levels is crucial. In the future, we expect to see hardware advances in whole-body sensing, compliant materials, and force-transparent mechanism design. Significant improvements on the algorithm side will also be needed. While classical planning and control approaches suffer from huge complexity issues, pure learning methods lack the flexibility to react to contacts and adapt to different

TABLE I  
TAXONOMY OF WHOLE-BODY LOCOMOTION AND MANIPULATION

	(a) Whole-body Manipulation	(b) Whole-body Loco-manipulation	(c) Loco-manipulation
Object movement (Manipulation)	✓	✓	✓
Robot self mobility (Locomotion)	✗	✓	✓
All surface interaction (Whole-body)	✓	✓	✗

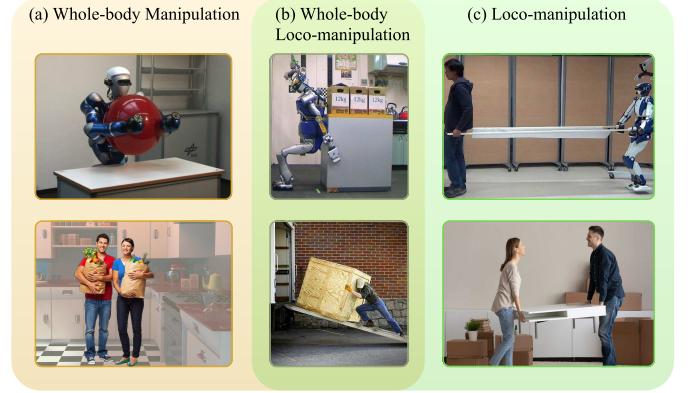


Fig. 3. (a) Whole-body manipulation exemplified by human and humanoid Justin [78] interacting with objects using all surfaces. (c) Loco-manipulation involves simultaneous locomotion and manipulation, as shown in the collaborative tasks performed by humans and a humanoid [33]. (b) Whole-body loco-manipulation is an intersection of (a) and (c), as exemplified by a human and a humanoid HRP-4 [79] pushing heavy objects using their legs and arms.

tasks. We foresee that the solution will be an integrated approach, which combines the strength of both. Ultimately, this could lead to more complex, human-like capabilities in humanoid robots, merging improved control, adaptive learning, and comprehensive sensing. Furthermore, addressing the core issues in loco-manipulation will also shed light on whole-body manipulation, as both areas involve handling complex, contact-rich interactions on different body parts.

#### D. Loco-manipulation

One of the key features of humanoid robots is their ability to simultaneously perform locomotion and manipulation (abbreviated as loco-manipulation hereafter) tasks. As suggested by its name, loco-manipulation involves both the movement of objects through manipulation and the mobility of the robot itself through locomotion. In a more general case of *whole-body* loco-manipulation, the *whole-body* refers to the use of all body surfaces to interact with the environment. We summarize the relationship between loco-manipulation, and whole-body manipulation in Table I. Both whole-body manipulation and loco-manipulation highlight the importance of utilizing physical contact. As shown in Fig. 3, loco-manipulation considers the movement of the robot itself while it manipulates an object, whereas whole-body manipulation emphasizes leveraging all accessible robot contact surfaces, such as using the chest as extra support to move large objects.

Loco-manipulation capability has been widely demonstrated on quadruped robots, specifically those achieving loco-manipulation capability by using their limbs as manipula-

tors [80, 81, 82]. For quadrupeds with upper-body manipulators, whole-body control is widely adopted for pick and place tasks from the model-based [83, 84] and learning-based community [85, 86].

Loco-manipulation for humanoid robots is particularly challenging, compared with quadrupeds. Humanoid robots have a smaller support region on the ground and a higher center of mass, which is challenging for dynamic balance. Therefore, early humanoid frameworks focus on separate control for locomotion and manipulation. For example, in locomotion tasks, most studies constrain the upper body to remain upright, which simplifies the whole-body problem to a bipedal locomotion problem that considers only the lower limbs. Conversely, in most table-top manipulation tasks, the lower body of the humanoid remains stationary [87, 88]. In such cases, any external force exerted on the upper body is treated as a disturbance to the legs, whose goal is to solely maintain balance. On the contrary, in [1], there is no such categorization of contacts: all contacts contribute simultaneously to achieve the task and balance.

Humanoid loco-manipulation requires a holistic and strategic use of the entire body to explore the humanoid's full behavioral capability space. Additionally, whole-body loco-manipulation needs to schedule contact for all limbs to simultaneously achieve robust movement and safe object interaction. Acquiring this technique unlocks a broad range of useful tasks such as opening doors [89, 90], pushing trolleys [91, 92], rolling large bobbins [93], or climbing ladders [94, 95].

### III. TACTILE SENSING

Humanoid locomotion and manipulation involve extensive physical interactions with the environment and objects, requiring multimodal sensing for understanding the environment, tracking manipulated objects, and evaluating how contact impacts the balance of both the robot and the objects. Visual sensors have shown effectiveness in object tracking and simultaneous localization and mapping (SLAM) [93], while proprioceptive sensors are usually combined to estimate contact information in contact-rich tasks [89]. These sensory modalities have been widely adopted in existing systems and have been thoroughly reviewed in the literature [19].

In contrast, tactile sensing, despite being uniquely suited for capturing detailed contact information critical for both locomotion and manipulation, has received relatively limited attention in existing humanoid literature. This survey specifically highlights tactile sensing to address this gap and to underscore its complementary role alongside other sensing modalities in humanoid systems.

Mimicking the human sense of touch, tactile sensing provides more accurate and comprehensive contact information over large areas of robot skin compared to proprioceptive sensors [96], and allows the robot to perceive complex environments and assess object properties through physical interactions, especially in scenarios where vision is occluded [97]. Additionally, tactile sensing can be used to estimate contact-based object properties such as roughness, texture, and weight, complementing traditional visual information such as location,

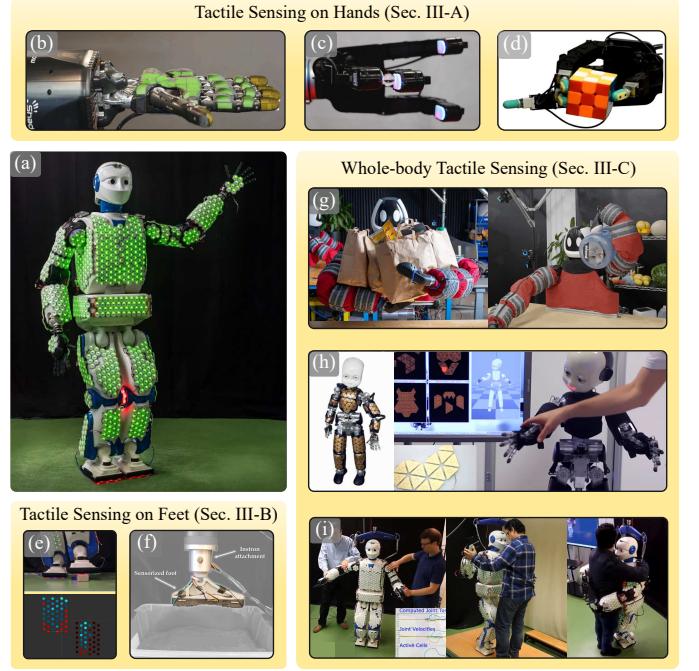


Fig. 4. Tactile sensing on humanoid robots, exemplified by (a) REEM-C fully covered with artificial skin [73] (image copyright: A. Eckert), which cover three body regions: hand, feet, and the whole body. (i) Hand tactile sensors demonstrated by (b) Shadow-Dexterous-Hand equipped with tactile sensors on palm and fingertips [111], (c) Allegro Hand equipped with DIGIT sensors [109], and (d) BioTac [123] tactile sensors for dexterous manipulation; (ii) Tactile sensors on foot soles for (e) obstacle recognition [120] and (f) terrain classification [121]; (iii) Whole-body tactile sensors for (g) whole-body manipulation by Punyo-1 [71] and whole-body human-robot interaction by (h) iCub [124] and (i) REEM-C [125, 126]

shape, and color [98]. A combination of tactile with other sensory modalities can significantly enhance humanoid perception capabilities in solving complex loco-manipulation tasks.

Numerous studies have developed tactile sensors based on various transduction principles that can sense normal and tangential forces, vibration, temperature, and pre-contact proximity information. Comparative studies of various sensor designs can be found in [13, 99, 100]. This survey instead focuses on their application in humanoid loco-manipulation, categorized into three areas: (i) tactile sensing on hands, (ii) tactile sensing on foot soles, and (iii) whole-body tactile sensing. The following sections review recent advancements in each domain, emphasizing their roles in balancing control, scheduling contacts, and enhancing interaction capabilities, as illustrated in Fig. 4 and Table II.

#### A. Tactile Sensing on Hands

Tactile sensors on dexterous hands provide contact information, addressing the challenges in object manipulation such as grasped object controllability and object property estimation. In this subsection, we survey studies that integrate tactile perception into the control, planning, and learning of complex manipulation tasks. Due to the similar nature of contact-rich interactions, the tactile sensing techniques on hands also offer valuable insights for whole-body tactile sensing and manipulation, which is discussed in Sec. III-C.

TABLE II  
ROLES AND CHALLENGES OF TACTILE SENSING IN LOCOMOTION AND MANIPULATION APPLICATIONS

Application	Roles of Tactile Sensing	Body Region Involved	Challenges	Solutions
Manipulation	• Grasping force and stability regulation [101, 102, 103, 104]	• Hand (in IHM*) • Whole-body (in WHM*)	• High-dimensional contact space • Tactile physics simulation • Sim-to-real transfer	• Dimensionality reduction [111, 115, 116] • Tactile simulators [117, 118, 119]
	• Object property and state estimation [105, 106, 107]			
	• Contact dynamics estimation [108, 109]			
	• Input modality for RL/IL/FM training [110, 111, 112, 113, 114]			
Locomotion	• GRF and support region estimation [120]	• Feet	• Sensor durability against heavy weight and various terrains	• Durable materials and mechanical design [120, 121, 122]
	• Terrain property recognition [121, 122]			
Loco-manipulation	• All of above • Unified locomotion and manipulation planning	• Whole-body	• All of Above • Real-time contact reasoning and adaptation	• To be explored

\* IHM - *In-hand Manipulation*; WBM - *Whole-body Manipulation*.

To achieve the grasping objectives, sensed contact forces serve as real-time feedback in force or impedance control loops to regulate the desired object behavior [101]. Moreover, slip detection and prediction based on tactile sensor data are used to adapt grasping forces, thereby enhancing grasp stability [102, 103, 127].

More complex in-hand manipulation tasks demand interactive perception beyond static object models. Dynamic contact information, including real-time tracking of object states, monitoring contact stability [106], and predicting interaction outcomes [108, 109], *i.e.*, how contact forces affect the balance of both objects and the robot, are crucial to achieving complex interactive behaviors. However, due to the inherent complexity of multi-contact dynamics and increased dimensionality of the contact state space, model-based methods still struggle to match human-level dexterity and versatility in multi-finger manipulation.

Alternatively, model-free Reinforcement Learning (RL) has shown the ability to address complex contact interactions. These approaches integrate tactile measurements directly into the state space to train end-to-end policies [110, 111]. Other learning methods besides task-specific RL have been sought for more generalizable policies. [112] employs diffusion policy to achieve complex and long-horizon bi-manual manipulation tasks, while recent work has integrated tactile sensing into foundation models alongside vision and language [113, 114]. Though limited to simple control tasks, these models may eventually enable more natural and versatile physical interactions in humanoid robots.

Advancing robot hands with tactile sensing for humanoid tasks requires addressing the dual demands of high dexterity for delicate manipulation and high payload capacity for heavy object lifting. While human hands naturally achieve this balance, most robotic hands prioritize dexterity but support limited payloads. In the short term, swappable modular hands tailored to specific tasks are practical, but the long-term goal should be a unified hand combining both capabilities. A promising approach involves multimodal sensing modules, integrating sensors optimized for different force ranges and resolutions. Progress in sensor design, material science, sensor fusion, and high-fidelity simulation is critical to this effort.

### B. Tactile Sensing on Feet

Other than manipulation, tactile sensing has started to gain traction for locomotion problems. For legged locomotion, estimation of Ground Reaction Forces (GRFs) and terrain properties is critical for maintaining whole-body stability on diverse, uneven surfaces. While vision and proprioception sensors can provide an indirect estimation of the terrain, these sensing modules lack the capability of accurately estimating GRFs and various terrain properties. Tactile sensing on foot soles has the potential to provide direct, unobstructed, and accurate contact measurements, but remains largely underexplored.

To measure GRFs, existing works use Force/Torque sensors mounted on ankles [128, 129] or load-cell sensors for point-wise measurement [130]. However, such methods inform only the zero moment point and lack accurate information on contact patch location, force distribution, and detailed terrain properties. To obtain such information, contact sensing arrays [122] and multimodal sensing suites [121, 131, 132] have been integrated into legged robotic systems for diverse contact information.

To date, tactile sensors for legged systems have been primarily developed for monopods, quadrupeds, and hexapods. Only a few studies have built tactile feet sensors for humanoid robots, with applications such as terrain classification [121] and ground slope recognition [122]. The sensed tactile information should aid the humanoid control and enhance the locomotion performance. A notable work in this direction [120] reconstructs the pressure shape of the foothold, enabling the recognition of uneven terrain and footstep replanning in real time.

Building tactile sensors for humanoid feet is more challenging due to larger impulse and shear force during intermittent ground contact caused by fewer legs and heavier robot weight. Another challenge lies in developing robust and reliable sensors capable of withstanding various terrains, prompting researchers to seek durable materials and dependable mechanical designs. In addition, humanoids have stricter requirements for system integration. For example, the computing and power units of an adult-size humanoid robot are potentially more

distal from the foot.

To enable robust humanoid locomotion in the wild, future directions for tactile sensing of feet need to address the following challenges: (i) how to accurately estimate more terrain properties such as stiffness, damping, plasticity, heterogeneity, and porosity; (ii) what are the appropriate metrics to measure the level of terrain complexity such as density, height, slickness, and roughness (*e.g.*, size and wavelength of rocks in terrain), and the effect induced by weather and lighting conditions (*e.g.*, rainy, snowy, sunny, night); and (iii) how to fuse terrain tactile sensing with other conventional sensing modules such as proprioception and visual perception to jointly inform postures, speeds, and gaits for intelligent, terrain-aware locomotion.

### C. Whole-body Tactile Sensing

Whole-body tactile sensing extends the aforementioned single-body sensing to all parts of the body, enabling humanoid robots to interact with unknown environments not only by the fingertips or foot soles but also by the arms, legs, and torso.

With explicit tactile feedback, humanoid robots such as iCub and REEM-C have achieved whole-body compliance [125, 133], controlling the contact force from whole-body regions. This level of contact awareness facilitates safe and intuitive physical human-robot interactions including dancing with human [126]. Contact awareness is also useful for improving balance and collision avoidance in unstructured environments.

Large-area tactile sensing significantly enhances a robot’s ability to handle *large* objects, including object identification through tactile exploration and whole-body manipulation. For example, [107] enables a humanoid robot NAO, covered with artificial skin over its entire upper body, to classify large, heavy objects with different weights and textures. [97] demonstrates whole-arm tactile sensing by reaching objects in cluttered spaces while regulating contact forces across its arms. Close-proximity whole-body capacitive sensing is implemented in [134], enabling a cobotic humanoid with workers close-presence awareness. The same technology is used to draw semantics in human-humanoid physical interaction in [135]. Moving away from traditional methods [136] that prioritize collision-avoiding trajectories, [137] uses tactile feedback to detect and clear movable obstacles, thereby solving the problem of navigation among movable objects. Additionally, with tactile sensors covering their arms and chest, humanoid robots HRP-2 and Punyo-1 can use their entire upper body to grasp and lift large, heavy boxes [104, 138] or various household items [71].

### D. Challenges in Tactile Sensing

Current works of tactile-based manipulation, whether in-hand or whole-body, are still largely limited to grasping or simple motions like rolling, pushing, or pick-and-place. Although tactile sensing holds great promise, its integration into more dexterous manipulation or loco-manipulation tasks that involve more dynamic interactions and contact shifting over the full body continues to pose substantial challenges. These include understanding the complex multi-contact dynamics, handling

the high dimensionality of the sensor data, and addressing the sim-to-real gap.

Several approaches have been explored to address these difficulties. For instance, dimensionality reduction techniques such as spectral clustering, principal component analysis [115], and autoencoders [116] have been employed to manage the high-dimensional input space. Others have proposed simplifying sensor outputs by using boolean contact states (*i.e.*, contact or no-contact) instead of continuous force values in RL training [111]. Furthermore, advances in tactile simulation [117, 118] are making simulated tactile data more accessible, thereby supporting zero-shot sim-to-real transfer [119].

One of the remaining fundamental challenges is to dynamically reason about contacts. This involves not only estimating contact states and static object models but, more crucially, understanding how these contacts and changes of contacts impact the system in real time, including the balance of both the robot and the object. Such information is vital for a planner to make informed decisions and, in a learning framework, can enhance sample efficiency. Addressing this challenge calls for deeper collaboration between the sensing and control communities to determine what tactile information should be abstracted and how to tightly integrate real-time tactile perception into control and planning for humanoid tasks.

*Key Takeaways:* Tactile sensing is yet an underexplored modality for advancing humanoid loco-manipulation, providing direct contact information necessary for tasks involving complex interactions with environments and objects. While it has improved task performance, achieving human-level dexterity and versatility requires advances in dynamic perception, and multi-modal sensing integration to enable systematic, real-time decision-making during interactions. Future directions include optimizing whole-body contact scheduling based on object properties and understanding how contact dynamics affect robot and object balance during loco-manipulation. Moreover, the design of whole-body tactile systems should account for varying sensor resolutions and load requirements, *i.e.* hands need higher resolution for delicate tasks, while body skin can operate at lower resolution but withstand higher payloads. For further reading, we recommend a survey paper on humanoid tactile sensing [13], and a book chapter on tactile sensing technologies with an emphasis on deployment on humanoid robot [139].

## IV. MULTI-CONTACT PLANNING FOR HUMANOIDS

Multi-contact planning remains one of the most challenging tasks in robotics. Specifically, in the context of humanoid whole-body loco-manipulation, a planner ought to solve trajectories that handle rich and intricate interactions with environments or objects. Particularly, beyond robot state trajectories, the planner is also expected to decide *contact position* (or *contact location*), *contact mode*, and *contact force* in a loco-manipulation task. Given the underactuated nature of humanoid robots and the addition of manipulation interaction dynamics, maintaining balance and manipulating objects rely solely on these contact interactions. This already makes multi-contact planning a challenging problem. Moreover, the diverse

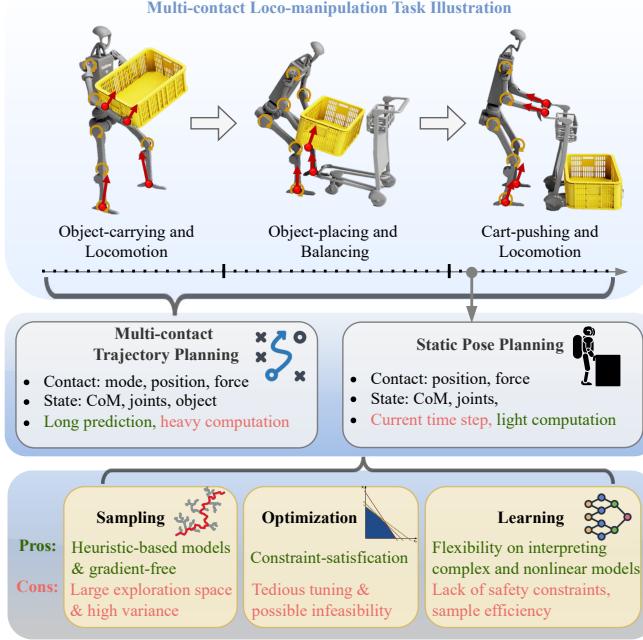


Fig. 5. An illustration of a task sequence for loco-manipulation planning in humanoid robots, involving carrying and placing a box and pushing a cart. The planning techniques explored include (a) multi-contact trajectory planning and (b) whole-body pose planning, highlighting their contact and state planning strategies. Additionally, the pros and cons of categorized approaches in (i) sampling-based, (ii) optimization-based, and (iii) learning-based methods are summarized.

physical properties of environments and objects (*e.g.*, rigid or soft, fixed or movable) complicate the problem even further.

Over the past decade, the field has produced fruitful results in multi-contact humanoid planning, demonstrating promising potential across various locomotion and manipulation tasks [3, 140, 141, 142]. However, these works require pre-planned contact mode sequences before planning robot whole-body motion trajectories [143]. This raises an open problem: how to solve the locomotion and manipulation contact planning problem simultaneously with the whole-body trajectory planning in a unified fashion, *a.k.a.*, Contact-Implicit Planning (CIP) [144, 145]. The primary challenge of this CIP lies in its high computational burden and combinatorial complexity of identifying potential contact mode sequences. Therefore, selecting suitable approaches depends on the specific problem requirements, including factors such as solving time, numerical robustness of the solution, resolution of the solution, and dependency on numerical models.

State-of-the-art procedures select planning algorithms for the underactuated system with three main categories: (i) searching, (ii) optimization, and (iii) learning, as illustrated in Fig. 5.

#### A. Search-based Contact Planning

Search-based approaches employ state expansion that allows exploring configurations to make and break contacts; collisions and kinematic feasibility are often checked during each search step. Heuristics can be applied in search-based methods for efficient exploration. The search result is an optimal sequence of

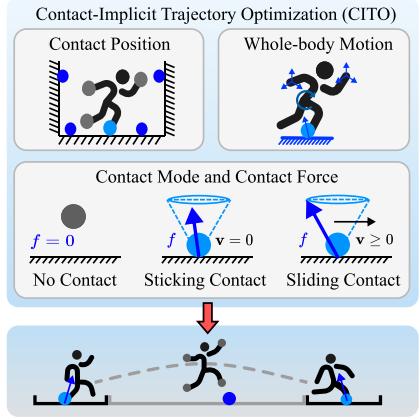


Fig. 6. An illustration of Contact-Implicit Trajectory Optimization (CITO) that simultaneously plans contact mode, contact position, contact force, and whole-body motion. However, solving CITO problems online for humanoid loco-manipulation tasks still poses a challenge.

contact modes that ensure stability and task efficiency. Whole-body motions can be optimized during the search to verify the dynamic feasibility of candidate contact sequences [146] or after the search in a *contact-before-motion* style [147]. Search-based methods are commonly used for gait planning in legged robot locomotion [148, 149, 150, 151]. Expanding their capabilities in more intricate multi-contact loco-manipulation planning, [152] seeks to implement a graph search method for humanoid grasp contact planning and replanning, and [153] introduces a contact-before-motion planner for multi-contact behaviors.

However, search-based methods usually struggle to cover the entire exploration space in a limited time budget for online planning and may result in solutions with high variance. To tackle this, [154] incorporates control variate and importance sampling as statistical variance-reduction techniques for faster solution convergence. [155] avoids the time-consuming re-planning by incorporating only forward path expansion with informed possible paths to achieve reliable online kinodynamic motion planning.

Furthermore, the feasibility guarantee of the results via search-based methods can be made through Pose Optimization (PO), a subset of multi-contact planning in humanoid loco-manipulation. This holds true when the contact locations, timings, and manner of interaction are predetermined—such as in scenarios where a humanoid safely assists a person [142] with feasible contact locations through point cloud processing. PO focuses on leveraging optimization-based techniques to plan whole-body poses and kinematic configurations at specific time steps, given a predefined contact mode. While PO is limited to handling discrete keyframes and does not account for continuous dynamics, this makes it highly suitable as a subsequent pose generator for gradient-free multi-contact planners, reducing the kinematic computation load during the search process. Furthermore, task-oriented objectives can be incorporated during PO, such as to maximize the interaction force [156, 157] and to efficiently retarget the operator's motion into safe and feasible robot poses [158]. Given a nominal pushing pose, Farnioli et al. [159] optimizes the

distribution of reaction forces among all contacts to guarantee the friction constraints in heavy object pushing. Kinematics-and-mass-model-based posture generator is employed on HRP-4 humanoid to leverage leaning pose and body contacts to improve force in a heavy object pushing task [79]. A kinodynamics-based PO approach is used in generating optimal humanoid pushing poses for dynamic non-prehensile locomotion [160]. Search-based multi-contact planning and PO are often paired with online whole-body control that effectively tracks the optimal pose while adaptively interacting with the environment and objects. We detail the whole-body control strategies in Sec. VI.

### B. Optimization-based Contact Planning

Unlike search-based contact planning, which primarily checks kinematic feasibility for expansion and often requires additional lower-level planning to generate dynamically feasible motion, optimization-based contact planning [144] offers the possibility of *simultaneous* planning of whole-body motion and contact interactions, as illustrated in Fig. 6. This approach integrates dynamics directly into the contact planning process, eliminating the need for a hierarchical structure. A Contact-Implicit Trajectory Optimization (CITO) is formed by incorporating contact dynamics into the trajectory optimization formulation, allowing the solver to determine the contact modes, contact forces, contact positions, and whole-body motions all at once [161, 162, 163].

Due to the inherently large problem size, CITO often rely on speed-up strategies, such as warm-starting with reasonable initial guesses for fast convergence [164] and separating into contact planning and whole-body motion planning subproblems in a hierarchical fashion [165]. With the increasing demand for computation efficiency, CITO have witnessed a rise in computation speed via sequential quadratic programming (SQP) (e.g., [144]), differential dynamic programming (DDP) (e.g., [166]), and iterative linear quadratic regulator (e.g., [167]). These improvements have even enabled the use of CITO in a Model Predictive Control (*a.k.a.* CI-MPC) framework for real-time planning on quadruped robots [168, 169] and robotic arms [170, 171]. However, for the humanoid robots, applying CITO to loco-manipulation has yet to be achieved.

Migrating such CITO as real-time CI-MPC to humanoid loco-manipulation presents its own set of challenges, including high-dimensional space of optimization variables, complex/undifferentiable contact dynamics models, proper modeling of interaction dynamics, resolution of initial guess, and tedious tuning. Esteban et al. [172] has made an initial effort to provide reasonable reference trajectories with the Hybrid Linear Inverted Pendulum (HLIP) model to allow natural contact behaviors at both hand and foot locations.

### C. Learning-based Contact Planning

Learning-based approaches have demonstrated promising potential in planning multi-contact tasks, such as using reinforcement learning to generate velocity commands and contact sequences for quadruped locomotion [173, 174]. These

learning-based planners are mostly modular, making it possible to form a hierarchical architecture with model-based planners and controllers at the low level. Compared with traditional optimization-based or heuristics-based approaches, learning-based elements enhance the computation efficiency in multi-contact planning. For example, [175] learns to predict the evolution of centroidal dynamics and contact sequence for dynamic humanoid loco-manipulation in under 0.1 s, achieving a computation speed boost 300 times faster compared to traditional optimization-based methods.

In addition, learning-based approaches can assist contact prediction, which allows additional information for contact (re)planning in real time. Precise contacts are often hard to obtain from motion capture data, making it challenging to learn directly from data. To synthesize plausible motion, a naive supervised learning approach often leads to objects moving without any contact or significant penetration between the predicted human body and the objects. [176] introduces contact correction and predicts motions relative to the contacts predicted in human-object interaction tasks. [177] separates the contact prediction and whole-body motion prediction by first predicting the contact positions with a large, dynamic object, which are then used as constraints to synthesize whole-body human motion. These models have the potential to serve as a loco-manipulation planner for humanoid robots. [178] learns to find from the video scenes the affordance, *i.e.*, potential contact points, for household object manipulation tasks. These contacts can be used as heuristics for subsequent motion planning.

### D. Challenges in Multi-contact Planning

Despite the advancements in multi-contact planning, several key challenges remain. First, computational efficiency remains a fundamental bottleneck. Contact planning inherently involves a combinatorial explosion in possible contact mode sequences, which poses major scalability issues. Optimization-based methods, such as CITO, often suffer from local minima and sensitivity to initialization. Search-based methods, though capable of global exploration, are computationally expensive and unsuitable for online replanning. Learning-based approaches, though computationally efficient, struggle with adaptivity and precision in dynamic contact-rich control. A promising direction lies in combining these paradigms to balance global reasoning, dynamic feasibility, and real-time performance. Second, current approaches rely on simplified point contact models, which fail to capture the complexity of real-world loco-manipulation. In practice, especially with full-body contact and manipulation, robots frequently engage in patch contacts that are complex to model with model-based control and planning frameworks. Extending contact modeling beyond point approximations is essential to better align with real physical interactions. Third, most existing methods work around fixed contact mode sequences, which diverge from the unpredictable and rapidly changing nature of real-world contact interactions. Unexpected contact events, which are often hard to observe due to visual occlusion and lack of whole-body tactile sensing, pose serious challenges for

accurate contact estimation, control, and online planning. This raises a fundamental question: is precise contact planning always necessary? To enable more adaptable loco-manipulation behaviors in uncertain environments, future research could explore planning, control, and learning strategies that are robust or even agnostic to small, transient changes in contact.

**Key Takeaways:** While significant progress has been made in humanoid multi-contact planning, future work should focus on developing more integrated approaches that combine the strengths of search-based, optimization-based, and learning-based methods. Specifically, addressing the computational complexity of CIP and improving real-time performance will be key. Future directions could explore hybrid approaches that incorporate efficient contact sequence generation/contact dynamics, apply contact-implicit constraints in real-time, and achieve learning-based contact prediction to enhance robustness and adaptability in complex loco-manipulation tasks. The readers are recommended to further read the survey on humanoid multi-contact planning [143].

## V. MODEL PREDICTIVE CONTROL FOR LOCO-MANIPULATION

Optimization-based Model Predictive Control (MPC) has advanced significantly in robotics. The advantages of its flexibility to define versatile motion objectives, rigorous mathematical formulations, and widely available solvers establish MPC as one of the most popular approaches to trajectory planning for locomotion and manipulation.

A uniform optimization formulation of the loco-manipulation planning problem seeks an optimal state trajectory and control input over a finite horizon in the future. MPC is often formulated as an Optimal Control Problem (OCP):

$$\min_{\mathbf{x}(\cdot), \mathbf{u}(\cdot), \boldsymbol{\lambda}(\cdot)} \mathcal{L}(\mathbf{x}(\cdot), \mathbf{u}(\cdot), \boldsymbol{\lambda}(\cdot)) \quad (1)$$

$$\text{s.t. } \dot{\mathbf{x}}_k = f(\mathbf{x}_k, \mathbf{u}_k, \boldsymbol{\lambda}_k) \quad (2)$$

$$h_{\text{task}}(\mathbf{x}_k, \boldsymbol{\lambda}_k, \mathbf{u}_k) = 0 \quad (3)$$

$$g_{\text{task}}(\mathbf{x}_k, \boldsymbol{\lambda}_k, \mathbf{u}_k) \geq 0 \quad (4)$$

where  $\mathbf{x}(\cdot)$ ,  $\mathbf{u}(\cdot)$ ,  $\boldsymbol{\lambda}(\cdot)$  are the trajectories of the states, control inputs, and constraint forces, respectively.  $\mathcal{L}(\cdot)$  is the cost function. The dynamics is represented in (2).  $h_{\text{task}}$  and  $g_{\text{task}}$  are other tasks represented as equality and inequality constraints.  $h_{\text{task}}$  are holonomic constraint tasks to be enforced strictly (e.g., a contact-explicit formulation [179]), and  $g_{\text{task}}$  are unilateral constraints to encode set-valued tasks (e.g., joint limits, non-sliding contact with fiction cones, etc.).

Depending on the choice of dynamics models (2), costs, and constraints, the OCP formulation is commonly transformed as a linear convex MPC (e.g., [2, 179]) or a Nonlinear MPC (NMPC) (e.g., [180, 181]). Table III summarizes recent MPC-based works on humanoid robots in loco-manipulation tasks.

### A. Simplified Models

In pursuit of high-frequency online planning with lightweight computation for motion control, simplified dynamics models, or reduced-order models (ROMs) are often

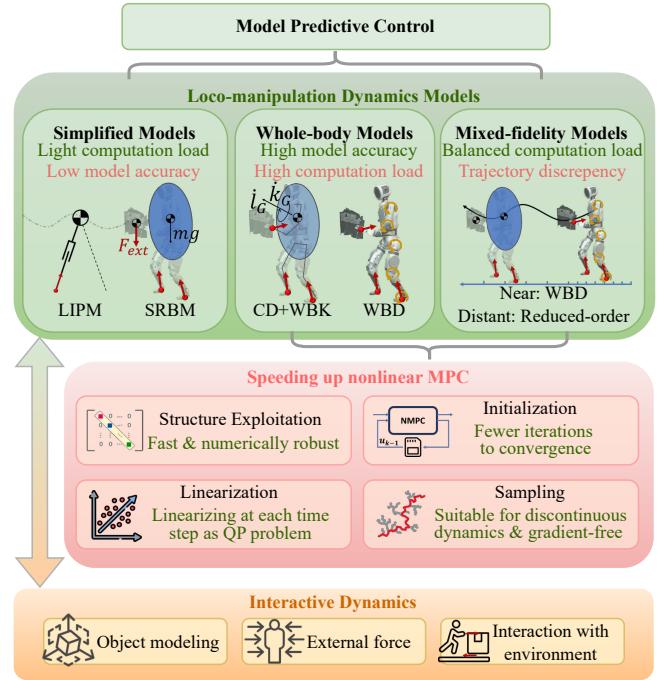


Fig. 7. An illustration of model predictive control in humanoid robotics, showcasing three primary categories of dynamics modeling in loco-manipulation: (i) simplified dynamics, (ii) nonlinear dynamics, and (iii) Mixed-fidelity dynamics. The figure also highlights the consideration of interactive dynamics modeling with environments and/or objects for loco-manipulation tasks. Additionally, four common approaches to speeding up/simplifying NMPC are summarized.

employed in MPC. For example, the Single Rigid Body Model (SRBM) can be linearized by providing explicit foot position sequence reference and be formulated in a linear convex MPC [182]. Using SRBM, [179] realizes dynamic aerobic behaviors on the MIT humanoid. Extending to humanoid loco-manipulation, [2] achieves object-carrying and rough terrain locomotion by simplifying the interaction dynamics as external gravitational forces applied to robot CoM.

On the other hand, the Linear Inverted Pendulum Model (LIPM) has been a popular choice as a linearized dynamics model for humanoid locomotion [27] and in multi-contact [189]. Extending LIPM to loco-manipulation tasks is achieved through teleoperation [185]. However, such a model inherently lacks the ability to address contact interactions and loco-manipulation dynamics, necessitating a lower-level whole-body control for balancing and manipulation tasks.

### B. Whole-body Models

While simplified dynamics models offer computation efficiency, the simplification assumptions cause a detriment to the model's accuracy and limit the capability of whole-body motion planning. Conversely, whole-body models are more accurate representations of the robot dynamics and are better suited for planning versatile motions and interactions with objects and the environment. NMPC comes into play when constraints or cost functions become nonlinear, for example, dynamics constraints formed by kinodynamics and Whole-Body Dynamics (WBD).

TABLE III  
RECENT MPC APPROACHES ON HUMANOID LOCO-MANIPULATION

Paper	Year	Robot Model*	Interaction Modeling Method	Locomotion ( $L$ ) and Manipulation ( $M$ ) Planning	MPC Frequency	Solving Method
[183]	2014	SRBM	Optimizing external wrench(es) at contact(s)	Unified	20 Hz	QP
[2]	2023	SRBM	Predefined external force	Unified	300 Hz	QP
[184]	2024	SRBM	Negligible object dynamics	Separated: $L$ : MPC; $M$ : Keyframe interpolation	—	QP
[185]	2019	LIPM	Negligible object dynamics	Separated: $L$ : MPC; $M$ : Teleoperation and retargeting	—	QP
[130]	2016	CD	Optimizing external wrench(es) at contact(s)	Unified	offline	SQP
[33]	2019	CD	Predefined external wrench	Unified	—	QP
[186]	2014	CD	Optimizing external wrench(es) at contact(s)	Unified	5 Hz/offline	SQP
[187]	2020	CD	Optimizing external wrench(es) at contact(s)	Unified	10 Hz	Interior-point
[181]	2023	CD	Estimated as external wrench through sensors	Unified	5 Hz	Interior-point
[188]	2021	WBD	Optimizing external wrench(es) at contact(s)	Unified	100 Hz	DDP

\* SRBM - Single rigid-body model; LIPM - Linear inverted pendulum model; CD - Centroidal dynamics; WBD - Whole-body dynamics.

In the context of humanoid motion planning, kinodynamic constraints are often referred to as the combination of Centroidal Dynamics (CD) and whole-body kinematics (WBK) constraints [186], where CD is derived from the total momenta of the system, and captures the effect of full-body inertia of a multi-linkage dynamics system [190]. For example, achieving consensus between CD and full-body kinematics in one Trajectory Optimization (TO) generates versatile humanoid motions [186].

Leveraging joint-space WBD (eqn. (5)) in MPC has gained popularity, particularly for free-floating articulated robots like humanoids. While WBD allows flexible contact modeling, its high nonlinearity and nonconvexity impose significant computational burdens on WBD-based Nonlinear Programs (NLP) or NMPC, making real-time planning challenging. This issue is especially critical for high-DOF humanoids in loco-manipulation tasks, such as payload transport, which require additional modeling of object dynamics and safety constraints. Thus, this section surveys methods to accelerate NMPC while preserving solution accuracy.

### C. Mixed-fidelity Models

Instead of using full joint-space dynamics across the entire horizon of an MPC, mixing multiple models of varying abstraction levels demonstrates improved performance and efficiency.

Cascaded-fidelity models (*a.k.a.* hierarchical dynamics) use different models to govern segments of the horizon [191, 192, 193]. These methods typically employ high-fidelity (*e.g.*, full-order) models in near horizons and low-fidelity (*e.g.*, simple) models for distant horizons, thus maintaining the solution accuracy in near horizons while solving the myopic issue by allowing a longer horizon using simple models. This approach could be suitable in loco-manipulation tasks as it would either simplify interaction dynamics as simple external forces or impose the object dynamics as part of CD in the far horizons to allow improved real-time computation compared to full dynamics models.

Another method is to overlap different dynamics models across their horizons. In such cases, achieving a consensus between these overlapped models is necessary. To solve problems with such mixed-fidelity models, [194] decomposes a single TO that incorporates both dynamics into two subproblems and

TABLE IV  
NMPC SPEED-UP STRATEGIES:  
ACCURACY, EFFICIENCY, AND LIMITATIONS

Method	Accuracy	Efficiency	Limitation and Challenge
<i>i. Dynamics Modeling</i>			
Simplified model (Sec. V-A)	low	high	Possible infeasible motion trajectory due to simplification
Whole-body model (Sec. V-B)	very high	very low	Requires very accurate dynamics modeling of robot and object
Mixed-fidelity model (Sec. V-C)	medium	medium	Motion trajectory discrepancy due to model mismatch
<i>ii. NLP Structure Exploitation</i>			
<b>Backward recursion:</b> DDP [196] iLQR [197]	medium low	high very high	Poor constraint-handling
<b>Derivative calculation:</b>			
Auto-Diff [198] Hypergraph [199]	very high high	low medium	May lead to high memory usage
<b>Exploring sparsity:</b> FATROP [200] AdaptiveNLP [201]			
FATROP [200] AdaptiveNLP [201]	very high high	high high	Requires specific problem structure Less effective in small-scale NLPs
<i>iii. Linearization</i>			
SQP [202]	high	medium	Scalability to high-order
Successive Linearization [203]	medium	high	Requires good initial guess
<b>QP Speed-up:</b>			
Condensed [204] ReLU-QP [205]	very high high	medium very high	No state trajectory output Only time-invariant states matrices
<i>iv. Warm Start</i>			
Gait Library [206] Memory of Motion [188]	high	very high low	Constrained by onboard storage Requires large dataset coverage

Performance scale (low to high): very low → low → medium → high → very high.

then alternates between the two to achieve consensus. Similarly, [195] alternates between the CD and WBK subproblems. Overall, model simplification over MPC horizons will remain an effective approach [9]. On the other hand, mixed-fidelity models demonstrate superior capability but require careful consideration of combined models.

### D. NMPC Speed-up

In addition to tailoring dynamics modeling strategies for specific loco-manipulation setups, solving NMPC can be accelerated using several key methods. Below, we summarize these methods and provide Table IV to compare their solution accuracy, real-time efficiency, and application limitations.

**NMPC Speed-up via Structure Exploitation:** NMPC problems often involve complex dynamics and constraints that can be computationally intensive to solve. Exploiting the structure within these problems can significantly enhance their

solvability and efficiency, such as extracting variables that directly interact with each other, identifying repetitive and symmetric structures, and arranging block-diagonal structures. One of the most common approaches to solving NMPC is direct methods, which transform the NMPC into a NLP with the complexity of  $O(N^3)$ , where  $N$  is the problem size [207]. Some direct methods, such as direct multiple shooting and direct collocation, result in sparse NLPs, whose computation complexity can be reduced to  $O(N)$  [208]. Another approach to solve NMPC is single-shooting methods, such as DDP [196] and its variant, Iterative Linear Quadratic Regulator (iLQR) [197], which only retains the first-order derivative approximation of dynamics and exhibits a linear increase in computation over the horizon [209]. With proper exploitation of the sparsity structure through the hypergraph approach, [199] shows improvement of the nonlinear solver in computation efficiency. Furthermore, AdaptiveNLP leverages the previous NLP structure to significantly reduce the overhead and update time for constructing the current NLP [201], which is particularly suitable for humanoid robot NMPCs with static kinematics and actuation constraints.

**NMPC Speed-up via Linearization:** Another way to tackle the computational burdens of NMPC is through successive linearization, which involves linearization at every timestep around the nominal system state and control input. The linearized dynamics become piece-wise affine, which can be formulated in a large, sparse Quadratic Program (QP) and can be solved online [203, 210]. SQP solves NLP problems by iteratively linearizing them into QPs and computing search directions at each step. While breaking down an NLP into QPs is viable, it requires efficient and scalable QP solvers. Aiming for high efficiency, ReLU-QP [205], a GPU-accelerated QP solver, has improved linear convex MPC real-time control frequency in high dimensional space balancing tasks from the original 206 Hz to up to 2600 Hz. While providing reasonably efficient real-time deployment, linearization techniques highly rely on linearization assumptions, which compromise model fidelity (accuracy), leading to motion errors compared to full nonlinear models. However, trading accuracy for speed is often preferred in humanoid robot control, as controllers may struggle to precisely track full-order trajectories by NMPC, making highly accurate solutions not practically beneficial.

**NMPC Speed-up via Warm Start:** The real-time requirement motivates many researchers to seek a more effective initialization. One simple yet effective approach is to warm start with the solution from the previous iteration. Another common approach is to offload the computation burden from online to offline, *e.g.*, the *gait library* [206]. It can be regarded as a specific type of warm-start technique and requires only a cheap online interpolation among the gaits to obtain an approximately optimal full-body trajectory. Similarly, [188] uses *memory of motions* to warm-start an MPC and overcome the sensitivity of initial conditions. However, the key challenge lies in the management of massive trajectories with limited storage. In Sec. VII-E, we discuss a solution from the learning community: learning compact models to distill from large-scale offline trajectories.

**NMPC Speed-up via Sampling:** Real-time sampling-based

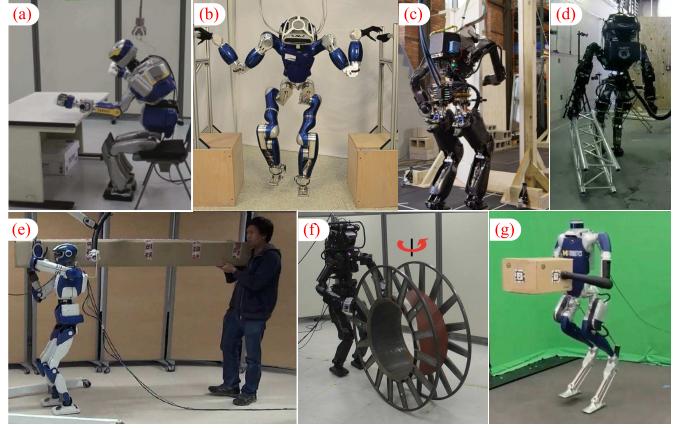


Fig. 8. Loco-manipulation skills from model-based methods. (a) Getting up from a chair [140]. (b) Multi-contact balancing [3]. (c) Door traversal [5]. (d) Transporting a bulky beam [130]. (e) Collaborative carrying [33]. (f) Rolling a bobbin [91]. (g) Box loco-manipulation [217].

planning, such as Model Predictive Path Integral (MPPI) control [211], is a simple and effective scheme. However, extending MPPI to high-dimensional loco-manipulation tasks presents significant computational challenges, especially within contact-implicit settings. Two primary techniques have enabled the recent success of MPPI: reducing the search space and leveraging parallelization in modern simulators. To limit the search space, researchers use suboptimal planners to guide the search, apply constraints, and employ spline control points to reduce the number of planning knot points [212]. Furthermore, advancements in sampling speed also facilitate real-time planning. For instance, MuJuCo MPC (MJPC) [213] leverages the established parallelization capabilities of MuJuCo [214] on multi-core CPUs. Additionally, modern simulators such as IsaacLab [215] and MuJuCo can roll out thousands of samples on GPUs, which allows additional randomization for robust control [216].

### E. Environment and Object Interaction Models for Loco-manipulation

In this subsection, we survey motion planning algorithms for loco-manipulation tasks that involve interactions with the environment and/or objects with large weights and sizes, specifically in the context of loco-manipulation MPC. Assuming the sequence of contact modes is defined through contact planning methods outlined in Sec. IV, the loco-manipulation MPC algorithms find a feasible trajectory that leads to a viable state over a horizon, while satisfying the dynamics constraint and contact stability constraints. The existing loco-manipulation MPC algorithms differentiate the interaction with a fixed environment and a manipulated object.

**1) Interaction with a static environment:** An environment, including static surfaces such as the ground and walls, provides contact forces that contribute to the robot's stability and enable interactive tasks such as walking and pushing. An example of a static environment is in Fig. 8 (a) and (b). Since the environment is static, the robot does not need to consider the environment's state or stability during planning. Instead, the robot is often required to deal with acyclic contact patterns

and non-coplanar contacts given specific environment geometries. This challenging problem is referred to as multi-contact planning (MCP) [95, 143, 218]. MCP is a widely studied area that involves both contact planning and motion planning. Since contact planning has been discussed in Sec. IV, in this subsection, we focus on the motion planning aspect of MCP, specifically in terms of real-time multi-contact MPC. Given a sequence of contact modes, MCP aims to find feasible motion and contact wrenches of all contacts.

Multi-contact MPC for humanoid robots can be solved by optimization-based methods [186, 187]. Among these methods, CD is the most common model due to its accurate representation of contact forces and the system’s centroidal momentum. Despite the model’s accuracy, CD contains a nonlinear term derived from the cross-product between the state (CoM position) and control (contact wrench), posing a challenge to trajectory optimization. Using multi-contact MPC as a motion planning technique also has limited dynamic locomotion capabilities because it treats arms and legs uniformly as general contacts, rendering it less efficient at handling frequent contact switches compared to pure locomotion models such as the LIPM. Although MPC has the ability to plan contact with any surface of the robot, the regularization of the planned contact forces usually requires accurate joint torque sensing or whole-body tactile sensing (Sec. III-C), which still has significant space to explore and presents great potential for rich and safe environment interaction.

2) *Modeling interaction with a manipulated object*: In the context of loco-manipulation MPC for humanoid robotics, modeling strategies for manipulated objects represent a crucial aspect and an area of ongoing research, alongside contact planning. An object can be a free-floating body (*e.g.* a box), a fixed-base articulated mechanism (*e.g.* a door or drawer), or actuated joints (*e.g.* another robot) [219], as shown in Fig. 8 (c-g). Unlike interactions with a static environment, the contact force exerted by an object depends not only on the robot’s joint torque but also on the object’s mass and inertia. As a result, interacting with objects in loco-manipulation tasks introduces significant complexity. Planning such tasks typically requires accurate knowledge of the object’s state and physical properties, especially when handling heavy, irregular, or dynamically moving objects. As a result, loco-manipulation tasks become significantly more complex, often requiring precise knowledge of an object’s state and physical properties—especially when dealing with heavy, irregular, or dynamic objects.

To address unknown object states and properties, adaptive control and online estimation techniques have been proposed to enhance robustness against dynamic effects and external loads. For example, [220] compensates for residual dynamics without requiring predefined object parameters, while [221] estimates reflected inertia to handle changing loads. Nozawa et al. [222] estimates object mass to optimize manipulation strategies for bulky items. Chappellet et al. [93] leverages wide-angle camera tracking to aid in large object handling. However, integrating these methods with MPC-based approaches introduces challenges, such as predicting object dynamics over preview horizons and managing increased computational load. Bang et al. [223] takes a step toward simplifying

centroidal momentum evolution through supervised learning, preserving a convex CD-based MPC formulation for humanoid locomotion.

Given the diversity of tasks, creating a unified model for the robot-object system is essential. We introduce two common approaches to incorporate object dynamics into the MPC-based planning process [189].

The first approach models the manipulated object as external wrenches and plans the control input to compensate for them [2, 79, 187]. This approach offers a flexible solution, as it integrates well with the MPC regardless of the linearity of the model, which treats all contacts as external wrenches. However, the contact wrench needs to be predefined, for example, to compensate for the gravity of the object or to exert a user-specified pushing force. Obtaining accurate contact wrenches for dynamic tasks like swinging a baseball bat is already inherently challenging, especially when considering their evolution over the entire prediction horizon in MPC. Static/quasi-static assumptions are usually made to neglect the dynamics of the object, resulting in less dynamic loco-manipulation motions. Another aspect to note is that contact wrenches can be applied at the contact location [91] or at the robot’s CoM [2, 33]. In the former setting, the object affects the contact wrenches for both self-balance and object manipulation. In the latter setting, the object affects only the contact wrenches that are responsible for balance, and contact wrenches for object manipulation require additional regulation. For example, the loco-manipulation MPC approach in [2] adjusts the foot contact wrenches to the weight of an object applied to the robot CoM and additionally regulates the object position via hand contact wrenches with a separate controller. Unlike MCP, such loco-manipulation MPC prioritizes mobility over manipulation and typically employs specialized locomotion models, such as a linear inverted pendulum model (LIPM). These models introduce additional assumptions for bipedal locomotion, such as assigning foot contact for locomotion and hand contact for manipulation, maintaining body height, and conserving angular momentum, making them computationally efficient in an online MPC setting albeit being less general.

The second planning approach incorporates the object’s dynamics directly into the robot’s dynamic equation of motion, creating a unified robot-object dynamic system [160, 189]. This approach eliminates the static/quasi-static assumption from the first approach and leverages the time-varying robot-object dynamics in MPC to achieve more dynamic and adaptive loco-manipulation behaviors. In such planning problems, the interaction wrenches are usually treated as control variables, and contact stability constraint on the interaction wrenches is enforced to securely attach the object to the robot. The planner generates the combined motions for both the robot and the object, leading to their desired states. Compared to modeling objects as external wrenches, this method requires a perfect knowledge of the object’s state, which is more challenging from the sensing perspective.

3) *Interaction with a dynamic environment or deformable objects*: Dynamically changing environments, such as those with a moving surface [224] or with physical human interactions [33], introduce additional challenges to loco-

manipulation planning and control. Similar to a dynamic object manipulation problem, the interaction model between the robot and the dynamic environment is also time-varying. Although one can infuse the dynamics of the object with the robot model to form unified dynamics, it is impractical to model the dynamics of the environment numerically in most cases. Therefore, in an MPC setting, the planner may require sensor feedback to predict the movement of the environment and replan the loco-manipulation motion adaptively [224]. For example, interacting with an environment that involves humans requires the anticipation of human intentions for collaborative manipulation such as lifting payload [33]; see a more challenging recent achievement in direct human-humanoid physical interaction [225]. For such tasks, the interaction force is an important way of communicating intentions, which can be measured as force feedback signals to trigger robot movements. However, the evolution of such sensed forces can not be well-predicted beyond the current timestep for MPC to leverage, suggesting further static/quasi-static assumptions are required. Otherwise, the robot can only treat the dynamic environment as a disturbance and counteract it through reactive control (*e.g.*, whole-body control). Given the challenge of dealing with the changing environment, loco-manipulation in a dynamic environment is largely unexplored.

In addition to rigid objects with regular geometries, deformable objects are ubiquitous in our real world, such as those in caregiving or housekeeping scenarios. Modeling the dynamics of these objects requires a deep understanding of their physical properties and behaviors, such as flexibility, elasticity, and deformation under force. Consequently, simplifications tailored to specific problems and applications are often necessary [226, 227]. For example, to plan the manipulation of a deformable belt, [228] simplifies the motion of the belt by representing only its tail movement in a 2D plane. However, to fully exploit the object’s deforming property for effective loco-manipulation, integrating accurate deformable objects [229] into robot models is essential. Although this area is relatively underexplored for humanoid loco-manipulation, such integration opens up significant opportunities beyond basic pick-and-place operations, enabling robots to tackle more intricate and delicate tasks.

*Key Takeaways for Model Predictive Control:* With the advanced capabilities of gradient-based numerical optimization in motion planning, MPC is gaining popularity in humanoid loco-manipulation, showcasing numerous variations in recent years of literature. The essence of this method lies in making reasonable choices regarding the dynamics model, constraint, task definitions, and computation requirements. These choices often require expert design and tuning to trade-offs among task versatility, solution feasibility, and optimality. By identifying the computation intensities and proper dynamics representation of the loco-manipulation tasks, one can offset the computation load by introducing simplified models and relaxed constraints in MPC. In addition, the MPC efficiency can greatly benefit from proper solver choices, an evolving area presenting opportunities for research on both solver-level and problem-formulation-level innovations. For further reading, we recommend the survey on MPC for legged and

humanoid robots [14]. Additionally, loco-manipulation tasks present further challenges due to the complexity of dynamic interactions with both the environment and objects, which leaves open questions on how to choose and formulate the interactive dynamics effectively based on the specific task requirements in an MPC setting.

## VI. WHOLE-BODY CONTROL

Whole-Body Control (WBC) represents a body of controllers that generate joint torques, constraint forces, and generalized accelerations to achieve a given set of desired dynamic tasks [7]. Three common cases necessitate a computationally efficient whole-body controller that can track desired trajectories and send torque commands to a physical robot. (i) The desired trajectory is computed based on a reduced-order model. Such a trajectory encodes only an important subset of the robot’s full-body motion (*e.g.*, desired CoM and end-effector trajectories in operational space [230]) and does not contain information for all joints. (ii) The trajectories are planned with a full-order model but are too computationally heavy [206] to be used in real-time, particularly for humanoids in loco-manipulation scenarios. (iii) Environmental uncertainties and planning inaccuracies induce disturbances that require robust WBCs to compensate [231]. Therefore, the WBC has been widely used in the humanoid community.

A WBC is different from an MPC in two aspects. (i) The MPC often solves a receding horizon problem, whereas the WBC solves an instantaneous control problem (*i.e.*, only for the current timestep). (ii) The model adopted by MPC may vary based on the user choice, as discussed in Sec. V, but WBC usually employs full-order Euler-Lagrangian dynamics that express equation (2) as

$$M\ddot{\mathbf{q}} - J^T \boldsymbol{\lambda} - S^T \boldsymbol{\tau} = -C\dot{\mathbf{q}} - G, \quad (5)$$

where the decision variables  $\mathbf{X} = [\ddot{\mathbf{q}}, \boldsymbol{\lambda}, \boldsymbol{\tau}]^T$  are generalized accelerations, external forces, and joint torques, respectively.  $M, C, G$  are the spatial inertia matrix, bias terms (*i.e.*, centrifugal and Coriolis forces), and the gravity term, respectively.  $J$  is the Jacobian, and  $S$  is the selection matrix. Given the selection of decision variables  $\mathbf{X}$  above, (5) becomes linear, which enables the WBC to be computed in real time.

In a humanoid robot with a floating base, the rank of  $S$  is smaller than the dimension of the generalized position  $\mathbf{q}$ , meaning the system is underactuated. This requires physical contact with the environment to achieve balance, mobility, and manipulation. The contact constraint is described using contact Jacobian  $J_c$ :

$$J_c\dot{\mathbf{q}} = 0 \Rightarrow J_c\ddot{\mathbf{q}} + \dot{J}_c\dot{\mathbf{q}} = 0. \quad (6)$$

These underactuated and contact-constrained dynamics (5), (6) represent the main components for solving the WBC for humanoid robots.

### A. WBC Dynamic Tasks

A dynamic task vector  $e_i$  can be expressed as a linear equation with respect to decision variables:

$$e_i = A_i(\mathbf{q}, \dot{\mathbf{q}}, t) \begin{pmatrix} \ddot{\mathbf{q}} \\ \boldsymbol{\lambda} \\ \boldsymbol{\tau} \end{pmatrix} - \mathbf{b}_i(\mathbf{q}, \dot{\mathbf{q}}, t), \quad (7)$$

where  $t$  is the time. Dynamic tasks can be equality constraints ( $e_i = 0$ ), inequality constraints ( $e_i \leq 0$ ), or cost terms ( $|e_i|^2$ ). The main idea of WBC is that the linear equation (7) is sufficient to describe a universal set of locomotion and manipulation tasks.

Although the appropriate set of WBC tasks depends on factors ranging from robot morphology to available hardware sensing, we highlight some of the common tasks for loco-manipulation. A task for tracking reference joint-space accelerations  $\ddot{\mathbf{q}}^d$  can be formed by setting  $A_i$  to a selection matrix and setting  $\mathbf{b}_i = \ddot{\mathbf{q}}^d$ . Similarly, a task for tracking a desired operational-space acceleration is derived through the end-effector's Jacobian [232]. A task for tracking a desired centroidal momentum rate  $\dot{h}^d$  can be formed by differentiating the centroidal momentum  $h$  [233]. Other potential WBC tasks include capture point [234], reference reaction forces [179] and collision avoidance [235]. The source of these dynamic tasks varies and may be predefined, computed online (e.g., from an MPC), or commanded through teleoperation.

MPC is commonly used to provide WBC with dynamic tasks in operational space. For example, an SRBM-based MPC [230, 236] outputs the centroidal trajectories and end-effector trajectories as dynamic tasks in operational space. These operational-space tasks can also be converted to joint accelerations and thus become joint-space tasks. For instance, whole-body inverse kinematics [237] is a common approach for this conversion. Additionally, Riemannian motion policy [235] and kino-dynamics fabric [217] can construct diverse joint acceleration from a hierarchy of primitive motions.

Teleoperation provides an interactive way to generate dynamic tasks such as the robot's posture, walking direction, and grasp targets [238]. WBC setpoints are often mapped to a visual interface, enabling an operator to modify controller setpoints on the fly. This mapping may be direct [239] or retargeted in order to account for the robot's morphology [240] or to ensure the feasibility of the commanded motion [158, 241]. Virtual-Reality (VR) interfaces enable spatially mapping handheld controllers to WBC poses. This approach has been deployed in various loco-manipulation scenarios, including doorway traversal, object grasping, and pushing tasks [242, 243]. Haptic feedback can inform an operator of the WBC state through various modalities, such as force feedback indicating CoM stability margin [244] and vibrating gloves indicating contact during manipulation [129]. Mapping to dynamic WBC setpoints, such as the capture point, has also been demonstrated and can account for variation in the natural walking frequency of the operator and robot [245, 246]. As shown in Fig. 9, WBC approaches can be categorized based on closed-form or optimization-based approaches to achieve a desired list of dynamic tasks.

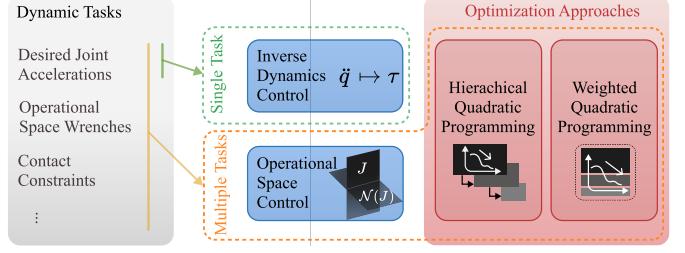


Fig. 9. Whole-body control is categorized into closed-form approaches and optimization-based approaches. Both approaches can incorporate multiple dynamic tasks and resolve task conflicts.

### B. WBC in Closed Form

An inverse dynamics controller is among the early works that address the WBC problem in closed form. In particular, it solves a single dynamic task: achieving the desired generalized acceleration  $\ddot{\mathbf{q}} = \ddot{\mathbf{q}}^d$ . The closed-form torque  $\boldsymbol{\tau}$  can be solved from (5) if we can measure all constraint forces  $\boldsymbol{\lambda}$ . However,  $\boldsymbol{\lambda}$  is usually unattainable due to the lack of sensing and estimation capability. To derive torque analytically, several methods [247, 248, 249] project the system dynamics (5) into a constraint-free manifold, establishing an equation between only  $\ddot{\mathbf{q}}$  and  $\boldsymbol{\tau}$ . [247] projects (5) into the actuated joint space, resulting in an equation of constrained dynamics independent of the constraint forces  $\boldsymbol{\lambda}$ . [248] uses orthogonal decomposition of the constraint Jacobian  $J$  to project (5) into the nullspace of the constraints. Although different projection techniques are used, [249] shows that these projections produce equivalent torque results. However, from the perspective of solving speed and numerical stability, orthogonal decomposition [248] is favored as it is faster and does not require the inversion of the inertia matrix.

Other than the task of tracking generalized acceleration  $\ddot{\mathbf{q}} = \ddot{\mathbf{q}}^d$ , a set of operational-space tasks and constraints can be achieved given the redundancy in the degrees of freedom of humanoid robots. As a multi-task example, a humanoid robot is often tasked to generate interaction forces with a low priority while maintaining the whole-body balance with a high priority. Operational-Space Control (OSC), *a.k.a* task-space control, achieves multiple dynamic tasks by prioritizing tasks in a hierarchy [247, 250]. Tasks with low priorities are solved within the null space of the high-priority tasks, enforcing that hierarchies are strictly maintained among tasks. Such a task hierarchy is also named the stack of tasks [251].

Overall, closed-form approaches are computationally efficient and are straightforward to implement. However, they have difficulty in incorporating inequality tasks, such as joint limit and obstacle avoidance. Although this issue can be addressed within the closed-form approach, such as using a smooth operator [251], much of the community uses optimization-based methods that address this issue efficiently.

### C. WBC through Optimization

In contrast with closed-form approaches, there have been a variety of studies formulating the WBC as an optimization problem. These optimization-based methods enhance the flex-

ability of WBC, enabling the modular addition and removal of dynamic tasks [252, 253], including inequality tasks.

A salient feature of optimization-based WBC is the resolution of conflicting dynamic tasks through two prioritization schemes: strict task hierarchy [254] or soft task weighting [130, 233, 234, 255]. Due to the linear property of both the dynamics equation (5) and dynamic tasks (7), optimization-based WBC is often formulated as a Quadratic Program (QP), which can be solved efficiently to a global optimum and enjoys a wide range of solver selections.

A strict task hierarchy can be ensured through a cascaded hierarchical QP. This method sequentially solves a series of QP subproblems with tasks from high to low priorities; lower-priority QPs produce solutions within the combined null space of all preceding QPs [254]. The sequential solve of QP terminates either when it successfully solves all subproblems or when it encounters an infeasible subproblem and thus skips all the remaining low-priority tasks [256]. Hierarchical QP is essentially equivalent to the closed-form stack-of-tasks approach, with the benefit of incorporating inequality constraints more naturally. However, solving multiple QP subproblems imposes significant computation and memory burdens. Additionally, the hierarchical QP inherits a common issue of OSC: the task Jacobians becomes rank-deficient when approaching singularities, which induce large unstable movements [257].

In contrast to hierarchical QP, which solves tasks in a strict order, weighted QP resolves task conflicts by treating dynamic tasks as *soft constraints* within the cost functions, with weights indicating their relative priority. Therefore, weighted QP can be regarded as a special case of hierarchical QP with only one hierarchy or vice-versa as documented in [258]. Such a setup enjoys the benefit of solving only a single optimization, which is faster than a hierarchical QP and can be further accelerated by exploiting sparsity and warm-start. However, tuning weight parameters can be burdensome for a large number of tasks, and can lead to instability [259]. Even with well-tuned parameters, the loss of strict task priorities means low-priority tasks can interfere with high-priority ones. Nevertheless, weighted QP is still widely applied in many robotics studies due to its easy setup and computation efficiency compared with hierarchical QP. For example, many of the weighted QP methods were designed during the DARPA Robotics Challenge [130, 233, 234, 260].

#### D. WBC for Loco-manipulation

WBC for loco-manipulation aims to achieve the desired motion while maintaining instantaneous *balance* and *contact stability*. Given the desired motion and contact sequence, loco-manipulation control can be categorized into two folds. (i) When all interactions with the environment and objects are static or quasi-static, they can be modeled as external wrenches. In this case, the WBC solves a robot balancing problem with the external wrenches as dynamic tasks. (ii) When the manipulated object has a substantial mass or is dynamically moving, such as carrying a heavy box, the object becomes an integral part of the robot-object system. Therefore, the WBC must account for the balance of both the robot and the dynamic object.

1) *Interaction as an External Wrench*: In this first category, a subset of contacts is responsible for interacting with the environment or objects to apply a desired wrench. This desired wrench can be specified by a user or derived from the estimated object weight. Considering the desired wrench from the interaction, the remaining contacts maintain the system balance using three distinct strategies.

The first strategy involves simultaneously optimizing contact wrench, joint acceleration, and joint torque using the robot's full-body dynamics, as detailed in Sec. VI-C. In this approach, the desired wrench for interaction is a dynamic task within the WBC. The WBC must also satisfy the dynamics constraint, contact stability constraint, and balance stability constraint. The contact stability constraint enforces that the resultant contact wrench lies inside the contact wrench cone (CWC) [261], maintaining firm and stable contact. The balance stability constraint designs a desired rate of centroidal momentum, often based on feedback in the CoM position and body orientation [190]. In the presence of state deviations, the balance stability results in a redistribution of the contact wrenches or a movement of the centroidal state to counteract and restore stability [91].

The second strategy, known as pre-optimization [3, 262], involves two stages in sequence. First, it determines the optimal distribution of contact wrenches based on the desired rate of centroidal momentum derived from the balance stability of CD. The second stage computes the joint torques needed to realize the contact wrenches using inverse dynamics of full-body dynamics. Note that, deriving the desired rate of centroidal momentum in the first stage is particularly challenging due to the non-holonomy [263] of angular momentum, *i.e.*, the kinetic momentum of rotation is not directly related to the orientation of body links. As a result, the body orientation requires additional regulation (*e.g.*, joint-level postural feedback [3]) beyond simple feedback on angular momentum.

To address the non-holonomy issue, the third strategy uses post-optimization [264]. The main idea is to treat the floating-base robot as a fixed-based system when calculating joint torques. The underactuated portion of the obtained torque is then mapped to contact wrenches through an optimal distribution problem. This method avoids the challenge of specifying the momentum of rotation in the pre-optimization strategy.

2) *Interaction as a Unified Robot-Object Model*: A unified robot-object system can leverage the additional object to regulate the robot's dynamics. This yields more dynamically feasible behavior when carrying heavy or dynamically-moving objects. The unified model incorporates each manipulated object as an additional "robot" – either a passive object or a real robot – and connects the robot and object via action-reaction force pairs [265]. The balance stability must consider the combined CoM and inertia of the robot-object system [219]. Additionally, the contact stability between the robot and the object is maintained to ensure that the object remains securely attached. While direct control of the interaction forces is feasible, adaptive force control that regulates the relative position between the object and the robot offers greater robustness. This approach mitigates the impact of inevitable inaccuracies in modeling inertia parameters and stiffness properties [189].

**Key Takeaways for Whole-Body Control:** The core of whole-body control lies in addressing an inverse dynamics problem to produce joint-level torque control. However, this problem is challenging due to the underactuation and contact-constrained nature of humanoid robots. Closed-form approaches such as inverse dynamics control and operational-space control are computationally efficient. Therefore, they have been traditionally prevalent. On the other hand, optimization-based strategies, particularly quadratic programs, are increasingly favored as they adapt more effectively to a wide range of task specifications and offer more reliable solutions. Undoubtedly, both lines of WBC research have significantly advanced the progress of humanoid robot control over the past two decades. In the near term, optimization-based WBC will continue to be a popular choice for low-level control to achieve high-level loco-manipulation tasks. We also see neural WBC [266, 267, 268] gaining popularity, as we will discuss in the following section. For further reading, we recommend the survey on optimization-based WBC for legged robots [9] and the chapter on closed-form WBC techniques for humanoid robots [7].

### E. Challenges in Numerical Optimization

Robotic planning and control techniques that are formalized as numerical optimization problems heavily rely on advances in applied discrete mathematics and optimization theory. These advancements address challenges such as non-convexity, numerical robustness, and real-time resolution performance. However, a plateau may have been reached in the transfer of these techniques to the field of robotics at large – humanoids specifically – despite exploiting the unique properties of robotic models. These properties can enhance the efficiency of optimization problem formulation and its resolution by tailoring them to specific applications. However, as evidenced by the formulation OCP, including MPC (Sec. V) and WBC (Sec. VI), inherent physical uncertainties can disrupt closed-loop performance. Extending these formulations to loco-manipulation primarily involves (i) augmenting models to incorporate loco-manipulated counterparts and (ii) refining contact models formulations to account for various interactions (e.g., impact [269], rolling, deforming). However, these extensions risk overcomplicating the problem, potentially hindering effective formalization, even if the resulting formulations are sparse.

Contact-explicit optimization formulations [6] are generally preferred due to their faster convergence and simplified formulation. However, they still suffer from the *curse of dimensionality*, which limits the complexity of the considered problem, the length of the preview window, and the discretization granularity of the decision variables. While increasing the resolution of the discretization and the length of the preview horizon increases the quality and stability of the results, the corresponding decrease in solution speed can, in some cases, cause an overall reduction in system stability. However, as computational power increases, this lowers the significance of these challenges.

Importantly, however, the contact-explicit formulations have the significant limitation of requiring the user to determine

the contact mode sequence for the problem, which generally limits the ability to generate complex motions. Alternatively, contact-implicit formulations introduce complementarity conditions to eliminate the strict dependence on the contact mode sequence [162, 167, 270]. However, contact complementarity conditions are nonsmooth, introducing severe computational challenges. Generally, this is tackled via regularizing the complementarity problem, e.g., [170], which approximates the constraint with a continuous affine function. Even with this linearized approximation, contact-implicit approaches struggle to scale to the high dimensionality of humanoid robots due to excessive computation and numerical difficulty. However, improved solver formulations, either as first-order solvers such as Alternating Direction Method of Multipliers (ADMM) [271] or specialized methods [168] for handling the complementarity conditions, have the potential to increase the tractability of real-time contact-implicit optimization.

Possibly more importantly, all these approaches, however, still have only *local* optimality guarantees. The level of nonconvexity or nonlinearity of the problem influences the solution quality, which often results in poor local minima or sub-optimal solutions. When the structure of the problem requires deviations from local candidate solutions, a feasible solution, even one that is only locally optimal, may never be found. This feasibility issue remains an open challenge for nonconvex nonlinear optimization.

Additionally, optimization approaches are almost always deterministic in nature, failing to capture the stochasticity in the state estimates and future contact events. This lack of consideration of uncertainty along the optimization horizon is critical, as it leads to limited robustness in translation to the real world. Addressing this uncertainty and lack of global optimality has led to the combination of search techniques with traditional trajectory optimization. For example, Model Predictive Path Integral (MPPI) [211] samples a variety of random control signals and their resulting state trajectories to determine the best action to take. Alternative contacts and objectives can also be sampled with contact-implicit approaches to help avoid local minima and find the globally optimal solution to accomplish a task [272]. Although both of these approaches heavily leverage computational parallelization for expediency, parallelizing the underlying optimization algorithms explicitly designed for trajectory optimization is gaining increasing prevalence, both on the CPU [273] and GPU [205, 274], as discussed in Sec. V. Despite this, the inclusion of the uncertainty into the optimization problem, whether through sampling [275], smoothing [276], or the introduction of stochasticity [277], is a promising approach for mitigating some of these challenges.

Despite the gains in these algorithms for considering the full system dynamics, the robustness of the mathematical solution when performing numerical optimization has been a concern due to the infeasibility of complex optimization problems, regardless of the computation speed improvement. As discussed in Sec. VI WBC, arbitrating techniques address infeasibility by (i) relaxing the hard constraint to the soft constraint by combining them in the cost with a weighted sum and (ii) prioritizing the constraints by achieving the important ones

first. However, how to design a smart solver to automatically resolve this issue and provide numerical robustness is still an open question. In addition, weight-tuning in high-dimensional problems with complex objectives is nontrivial, highly task-dependent, and can lead to instability [259]. Researchers have made initial steps to apply auto-tuning techniques to streamline the tuning process in Optimal Control Problem (OCP) designs for humanoid robots [278, 279]. Until these are solved, assigning definitive but non-violating constraints and designing objective functions while maintaining global versatility still depend on expert knowledge.

## VII. LEARNING LOCO-MANIPULATION SKILLS

Robot skill refers to the ability to use its own perception, planning, and control capabilities to complete specified tasks autonomously [280]. Among a variety of robot skills, loco-manipulation is highly valuable for augmenting and complementing human capabilities. Traditionally, loco-manipulation skills are developed from human designer knowledge, distilled into pre-programmed planners or controllers. In contrast, learning-based methods leverage computation and data. Although learning skills require collecting extensive data from either autonomous exploration or expert guidance, this approach is powerful as it tends to yield novel behaviors that are difficult to encode from human knowledge.

This section reviews learning-based approaches that explore two main directions: (i) enhancing a specific skill in terms of agility, robustness, and safety, and (ii) broadening the overall skill set of robots, revolving around two pivotal goals: versatility and generalization. Versatility refers to the capability of a single framework or policy to master *multiple* skills, whereas generalization involves adapting existing skills to new, out-of-distribution tasks and environments. Versatility can be viewed as a prerequisite for generalization. For a model to generalize across a wide range of tasks, it must possess a broad and versatile set of skills that can be applied in novel contexts. However, possessing versatile skills alone does not guarantee generalization—the model must also be capable of correctly prompting and adapting those skills to meet novel objectives. For instance, in the context of skill learning, a versatile policy can track multiple trajectories within the training dataset, while a generalizable policy should be able to track motions beyond those seen in the dataset, thereby demonstrating out-of-distribution capabilities.

Among learning-based methods, Reinforcement Learning (RL) without demonstration and learning from demonstration, also known as Imitation Learning (IL), have shown remarkable proficiency for robotic skill learning. RL has been successful in coordinating complex full-body motions for humanoid robots, including dancing [266, 281], agile soccer maneuvers [282], and robust locomotion [38, 42]. However, RL policies are often fine-tuned for specific tasks within specific environments. This limitation largely stems from the reward function being narrowly tailored to a specific task, and the policy only capable of learning from the same or similar environments. In contrast, IL addresses this problem by leveraging large datasets of demonstrations [40, 283]. Recent advancements in IL have

demonstrated promising results for scaling to a large number of skills [10], showing potential for solving complex multi-skill tasks.

For the basics of RL and IL, we refer readers to the survey paper [16]. In this section, we discuss these methods for learning humanoid loco-manipulation skills. As shown in Fig. 10, we introduce RL in Sec. VII-A and IL in Sec. VII-B and Sec. VII-C. Thereafter, we discuss the benefits of combining model-based and learning-based methods in Sec. VII-D. Finally, we discuss methods for learning versatile skills with a single policy in Sec. VII-E.

### A. Skill Learning: Reinforcement Learning from Scratch

RL, enhanced by modern deep learning toolchains and algorithms, has garnered significant attention in the field of robotics over the past decade. RL promises an effective way to learn motor skills by rewarding desirable behaviors and penalizing undesired behaviors, with minimal or no supervision during training. The end-to-end RL policies translate raw sensory input to actuation and are executable in real time.

RL provides distinct benefits but also comes with its own challenges. Compared with model-based methods, numerous RL methods are model-free (see Sec. VII-E for model-based RL), eliminating the need for accurate dynamics. Furthermore, RL does not require demonstration data, making its training setup straightforward. However, it often requires meticulous design of reward functions to shape the policy's behavior. In addition, deploying a learned policy on robot hardware often encounters a sim-to-real gap, a well-known issue induced by the inaccurate physical model used by the simulator. Policy learning from scratch often requires extensive and time-consuming interactions with the environment without a guarantee for task completion. For example, popular RL algorithms such as Proximal Policy Optimization (PPO) [284] and Soft Actor-Critic (SAC) [285] fail in most humanoid loco-manipulation tasks [286], partly due to the complexity of these tasks, the sample inefficiency, and the sparse reward design.

1) *Comparison Between RL Algorithms*: PPO [284], and similarly other on-policy reinforcement learning algorithms such as A3C [287], is sample inefficient, but the policy update is fast and scalable. PPO is widely used in sim-to-real settings where simulation data is relatively cheap to obtain, *e.g.*, in a parallelized simulation environment such as Isaac Gym [215]. SAC [285], and similarly, Twin Delayed Deep Deterministic (TD3) [288] and Deep Deterministic Policy Gradient (DDPG) [289], are sample-efficient off-policy algorithms, but slower to train due to higher computational costs for the policy updates. In addition, Q-learning-based methods such as SAC can introduce model bias during the training process since they estimate gradients using a learned Q-function, whereas policy gradient methods such as PPO have the advantage of directly estimating the gradients of the learning objective. Off-policy algorithms are widely used when data collection is expensive, *e.g.*, when the data can only be obtained from a real robot. The recent literature on RL for humanoid robots relies mainly on simulation data, and hence PPO is widely employed. For other RL settings, such as real-world RL, we envision that

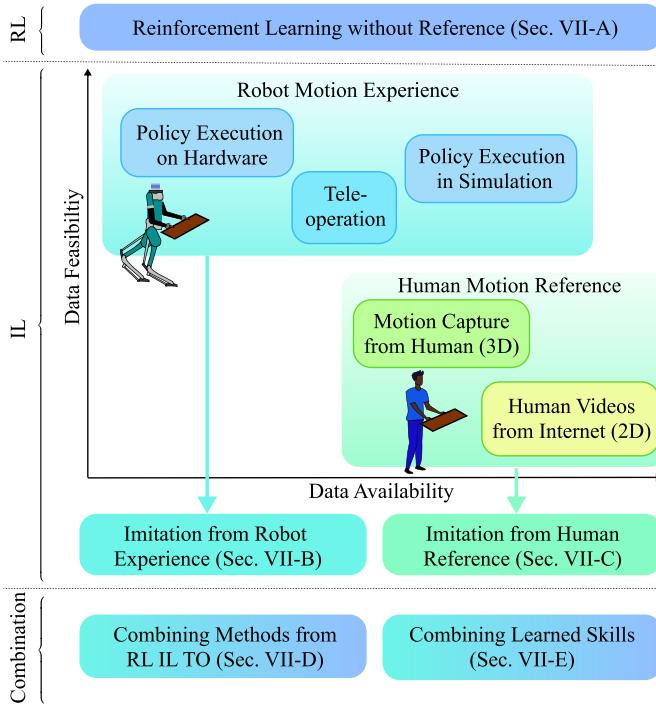


Fig. 10. The organization of approaches for skill learning. RL does not require reference data in a standard setup. IL makes use of four different data sources for efficient learning. The larger the morphology gap or sim-to-real gap is, the more challenging it becomes to learn effectively from these data.

more off-policy RL algorithms and novel RL algorithms will play an important role in learning humanoid skills.

2) *Improving Learning Efficiency*: Several approaches have been proposed to improve learning efficiency. Curriculum learning expedites training by allowing the policy to achieve simple tasks in the early stage of training and then progressively increasing the task difficulty and complexity [290]. Another approach is to promote exploration. Researchers use curiosity mechanisms, which encourage visiting unexplored states, to intrinsically motivate learning without an explicit reward design [291]. This has been shown to overcome the sparse reward setting and achieve complex loco-manipulation behaviors such as door opening. [281] also incorporates curiosity-based rewards to learn versatile loco-manipulation skills without any motion priors. Lastly, substituting reward terms with constraints in a constrained RL framework can significantly simplify reward tuning while achieving state-of-the-art locomotion performance [292].

3) *Addressing the Sim-to-real Gap*: Sim-to-real is another formidable challenge in RL. Nevertheless, RL has been successfully applied to various areas of robotics, notably in quadrupeds [293], where the sim-to-real gap has been consistently overcome. The success of quadrupeds hinges on new investments in infrastructure for affordable hardware and highly parallel physics engines, spearheaded by key players in the robotics field. It is also important to note that quadrupeds benefit from an inherently stable dynamic system similar to manipulators while operating in less complex environments compared to typical loco-manipulation tasks.

In contrast, humanoid loco-manipulation faces steeper sim-

to-real challenges. Humanoid robots possess higher DoFs and unstable dynamics, where the center of mass constantly moves out of the support polygon. Therefore, learning whole-body balance control is sensitive to parameters in a physics simulation, underlining the sim-to-real gap due to differences in dynamics between the virtual and real worlds. Additionally, humanoids are expected to perform human-level manipulation tasks where the complex environment and the differences in observation space, such as visual appearances, aggravate the sim-to-real-gap.

To address the sim-to-real challenge, a diverse set of mainstream approaches have been explored for humanoid robots. Domain Randomization (DR) is among the most widely adopted approaches, which varies the properties of a robot model, such as mass, friction, and actuator dynamics, to train a generalized policy robust in the real world. Many humanoid works [42, 282] achieve sim-to-real transfer through DR. While DR is straightforward to set up, policy training is sensitive to the parameter randomization range, inducing laborious tuning: a larger range is challenging for the policy to fit all physical parameters (*i.e.*, fail to learn), and a smaller range does not cover the full spectrum of parameters that can occur in the physical world (*i.e.*, fail to transfer).

System Identification (SI) is another popular approach to enhancing model fidelity by approximating the system's input-output behavior from real-world data. Real-to-sim techniques use optimization [294] or search [295] to obtain simulation parameters that can best explain the real robot trajectories collected from policy execution. Learning to align simulation behavior with policy rollouts in the real world has also shown success in sim-to-real transfer [296]. However, it is challenging to collect real-world data covering the full space of states and actions, particularly for versatile, safety-critical tasks.

While SI uses real-world data to obtain an accurate model, Domain Adaptation (DA) uses real-world data directly to fine-tune a simulator-trained policy. In DA, the parameter distribution in simulation is defined as the source domain, and parameter distribution in the real world is defined as the target domain. The fine-tuning effectively transfers the policy from the source domain to the target domain. For example, Sim-to-Lab-to-Real [297] develops a two-stage transfer: pre-training in simulation and fine-tuning in the real world. Although only limited hardware data is needed for fine-tuning, safety is still a major concern. Safety filters are often deployed to prevent dangerous movements when collecting real-world data [298].

Despite these efforts to address the sim-to-real gap, a systematic solution remains elusive, as the aforementioned approaches are often case-specific. Against this backdrop, developing physics engines that facilitate real-to-sim construction and sim-to-real transfer is crucial. However, physics engines are constrained by real-time requirements and continue to face challenges in accurately solving contact dynamics. For instance, MuJoCo [214] employs a soft contact model that can lead to unphysical penetration artifacts, while PhysX [215] approximates friction cones using low-fidelity pyramids that widen the sim-to-real gap. Although existing solutions inevitably involve trade-offs between accuracy and efficiency, advancements in contact modeling will significantly narrow

TABLE V  
SKILL LEARNING METHODS BASED ON DATA SOURCE

Methods and Data	Pros and Cons	Algorithms
RL Without Reference	✓ novel behavior ✗ reward tuning	PPO [284], SAC [285]
IL Robot Execution	✓ annotated data ✓ dynamically feasible ✗ scarce ✗ limited diversity	Diffusion [283], IRL [300]
IL Teleoperation	✓ multimodal behavior ✗ rare full-body motion	BC-RNN [89], ACT [268, 301]
IL Motion Capture	✓ accurate kinematics ✗ small dataset ✗ limited outdoor data ✗ proprioception-only	RL motion imitation [302], GAIL [303], AMP [304, 305]
IL Human Video	✓ diverse abundant data ✗ non-physical motion ✗ proprioception-only	RL motion imitation [306], GAIL [307], OKAMI [308]

the sim-to-real gap.

*Key Takeaways:* RL provides an effective way to learn novel behaviors for humanoid loco-manipulation. However, in practice, the success of RL often relies on informative representations for both observation and action, extensive reward engineering, curriculum learning design, and a vast amount of trial-and-error experiences to estimate gradients. Consequently, using RL to train robots is almost never practical in the real world, at least for the current stage of development. Therefore, RL policies are predominantly trained in simulations. This makes the sim-to-real gap the Achilles' heel of RL, significantly dampening its initial promise. Compared with quadruped robots, the sim-to-real gap is particularly challenging for humanoid robots with high DoFs executing complex loco-manipulation tasks. This is why IL, leveraging limited but in-domain real-world data, has gained popularity over pure RL without demonstration data. For further reading in RL, we recommend the survey on learning-based legged locomotion [16] and the practical lessons for training robotic RL agent [299].

### B. Skill Learning: Imitation from Robot Experience

Imitation Learning (IL) is an umbrella term that represents a class of algorithms, including supervised, unsupervised, and reinforcement learning, that train policies from expert demonstrations. IL is particularly effective for complex tasks that are difficult to specify explicitly. Three essential steps exist in IL [18]. The first step is to capture the expert demonstration. The next step involves retargeting, where these demonstrations are mapped to the robot motions. If the captured motion comes from the same robot, such as from teleoperation, the retargeting step is unnecessary. The final step is policy training using the retargeted data.

We discuss four possible sources of demonstrations for humanoid robots: (i) policy execution, (ii) teleoperation, (iii) motion capture, and (iv) human videos, as illustrated in Fig. 10 and Table V. We group these data sources into two categories: the robot experience data, which represents those directly obtained from the robots through policy execution or teleoperation, and the human data, which includes human motion captures and videos of human activities obtained from the Internet. Robot experiences exhibit smaller morphological

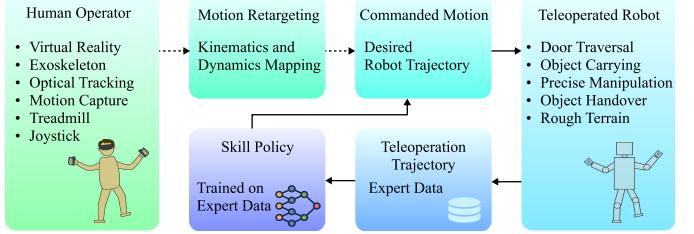


Fig. 11. Control flow for learning from teleoperated demonstrations. Expert data is created from teleoperated trajectories (dashed lines), which in turn is used to train and deploy an autonomous skill policy (solid lines).

discrepancies and are directly applicable to policy learning but are typically scarce. Conversely, human data are more abundant but present significant morphological differences to humanoid robots.

1) *Obtaining Robot Experience Data:* A reliable way to obtain robot experience data is to execute existing policies, either model-based or learning-based. However, collecting data on a physical robot requires a laborious setup of the environment and raises significant safety concerns. Therefore, conducting these executions in simulation is a more efficient approach, although the fidelity of the simulator will inevitably cause a sim-to-real gap.

Teleoperation is one of the most common ways to directly capture robot data commanded by human experts. A main advantage of teleoperation is its ability to provide smooth, natural, and precise trajectories for a wide range of tasks. Fig. 11 outlines the control flow of teleoperation data used as a source of policy training. The first step of this process is generating the demonstration through teleoperation, represented by the top path of the control flow (dashed lines). Motion retargeting maps human measurements from the teleoperation device to the desired trajectories in the robot domain. Robot execution data collected from teleoperation can be used to train autonomous policies (solid lines) that directly command the robot's motion without human intervention.

However, teleoperation for the *full-body* of a humanoid has a number of limitations. Firstly, a majority of teleoperation systems capture only manipulation skills [10]. Generally, Virtual Reality (VR)-based teleoperation schemes cannot sense the operator's gait and are restricted to simply commanding walking speeds and directions via joysticks [242, 309, 310]. Full-body sensing, including human gaits, often requires additional equipment such as IMU suits [240, 311] or exoskeletons [245]. These devices may be prohibitively expensive, bulky, complicated to maintain, and lack transparency and user-friendliness. Secondly, while teleoperation can generate versatile training data for a wide range of tasks, the utility of this data may be limited if the robot's kinematics fail to seamlessly retarget the reference human motion. Additionally, retargeting dynamic tasks such as walking or pushing an object relies heavily on the dynamic model of the human demonstrator [240] and requires meeting high synchronicity [311] and rich sensory feedback.

Although both teleoperation data and policy execution data are recorded from robots, they exhibit distinct characteristics. During teleoperation, human instructors tend to provide diverse demonstrations even for the same task. Therefore, teleop-

eration data are often multimodal; that is, given a specific task, there exists a distribution of plausible ways to accomplish the task. In contrast, data from executing a single policy are often unimodal; that is, given an input, the output is often fixed. Different policy learning approaches have been proposed to address these multimodal and unimodal data features.

### 2) Approaches to Learning from Robot Experience Data:

From unimodal policy execution data, which contains paired observations and actions, IL approaches are often used for policy distillation. Behavior Cloning (BC) casts IL as a supervised learning problem, which remains one of the most straightforward approaches for robot skill learning [312, 313]. Another IL technique is Inverse Reinforcement Learning (IRL), which reconstructs rewards from the data in addition to training an RL policy. The IRL study in [300] infers a generalizable reward of the expert demonstration for bipedal locomotion and then uses it to train an RL policy in unseen terrains.

To capture the data multimodality and produce diverse future actions, the Action Chunking Transformer (ACT) [268, 301] is adopted to handle distribution shifts due to the compounding error inherent in naive BC. Diffusion policy [283], a BC method, shows the ability to acquire multimodal locomotion skills by learning from a large dataset collected from multiple expert policies. However, obtaining these skills requires large-scale versatile data, which motivates the scaling of data collection via teleoperation in industrial companies, *e.g.*, Tesla and Toyota Research [10].

**Key Takeaways:** Although collecting high-quality data demands considerable effort and resources, IL from robot experience remains a reliable method for attaining skills with expert-level performance. Industrial companies and research labs are increasingly focusing on scaling data collection to develop a broader range of diverse policies through IL. Especially, teleoperation is one of the most popular ways to collect humanoid robot experiences nowadays. For further reading on collecting robot experience data, we recommend the survey on humanoid robot teleoperation [314]. We also find the survey on imitation learning of humanoid bipedal locomotion [18] a decent summary.

## C. Skill Learning: Imitation from Human Data

While robot experiences can serve as a reliable data source, collecting loco-manipulation data directly from robots remains a formidable challenge. Collecting teleoperation data, even though it is one of the most commonly used approaches, is costly and tedious to scale. Additionally, gathering robot data by deploying existing model-based methods or trained policies presents additional difficulties. Deploying these methods on hardware raises safety concerns. On the other hand, data collected from a simulator or synthesized from a model suffers from sim-to-real gaps. Furthermore, model-based methods based on human knowledge (*e.g.*, dynamics models, heuristic trajectories) generate consistent but similar behaviors, leading to limited data diversity.

Learning from a large, diverse corpus of human data mitigates these challenges, as recording human data is more accessible and scalable due to the reflexive usage of loco-

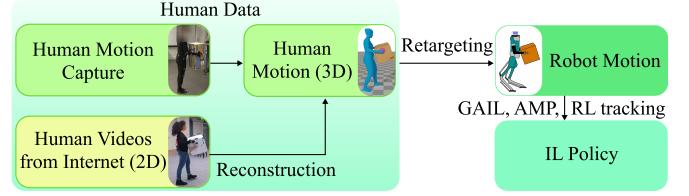


Fig. 12. A pipeline for learning from human data. The 3D human motion data can be recorded from motion capture systems or reconstructed from 2D human videos. Robot motion is retargeted from 3D human motion for imitation learning.



Fig. 13. Motion capture of human data. (a) The human interacts with household items [177]. (b) The human interacts with items with whole-body manipulation [315].

manipulation skills by humans in their daily lives. Furthermore, training policies to imitate human data can simplify the synthesis of loco-manipulation behaviors. Recent research efforts in 3D human motion data archival have surged in the vision and computer graphics communities. As shown in Fig. 12, there are currently two primary approaches to acquiring 3D human motion data: (i) recording directly from motion capture systems and (ii) reconstructing from 2D videos.

**1) Obtaining Human Data:** Various tracking systems are used to obtain human motion data. As shown in Fig. 13, [177, 315] captures humans interacting with various objects while moving around. The following datasets, CMU [316], SFU [317], LAFAAN [318], and AMASS [319], are commonly used because they provide a wide variety of human motions. However, motion capture data often require heavily instrumented environments and actors, making them expensive to scale, and indoor lab settings provide little exposure to outdoor activities.

Alternatively, videos and images obtained from the Internet offer a rich and diverse source of human motion data, including athletic performances, artistic dances, or daily chores. However, motion extracted from internet data is usually of lower quality, containing noise, jittery, and non-physical artifacts due to occlusion and motion blurs. Therefore, the reconstruction of accurate 3D human poses from 2D data remains an active research topic in the computer vision community.

Animation is also a widely explored approach to obtaining human motion data. Although the approach is effective in designing expressive motions for virtual human characters, the process often requires the use of sophisticated animation tools by professional animators [320], which makes it less scalable than motion capture or Internet videos. To address this limitation, researchers have been pursuing *motion generation*, leveraging human data to generate diverse and realistic motions. Recent advances in generative models, particularly diffusion-based approaches for video generation such as Sora [321] and Stable Video Diffusion [322], have significantly propelled progress in motion generation. A comprehensive survey [323] highlights the impressive capability of these models to generate realistic human motions. However, our focus here is training IL policies to generate *physically plausible* actions that achieve the motions demonstrated in human data. The training pipeline is shown in Fig. 12.

2) *Challenges in Learning from Human Data*: Using offline human motion data (from any source) to train humanoids inevitably creates an embodiment gap both in observation and action spaces due to different body proportions, joint configurations, and mass distributions between most humanoids and humans. Closing this embodiment gap requires retargeting, which involves mapping the motion collected from a source skeletal model to a target robot model. Previous work in computer graphics and robotics has explored various retargeting strategies, such as joint-space [240, 268, 324] and task-space correspondence [301], contact points [219], fingertips [325], gait synchronization [245, 311], and motion feasibility filter [158, 267]. Developing systematic solutions for retargeting the entire human body, including dexterous hands, remains a critical topic for advancing humanoids.

Another correspondence problem is that the human data are only proprioception-based, which lacks sensory input and action output. Notably, these human data lack tactile or force measurement from interactions, which limits the capability of learning for loco-manipulation with rich physical interaction. To solve this issue, IL trains control policies that track references within a physical simulator. Specifically, IL control policies accept state-only references instead of state-action pairs and output control signals. The physical simulator plays a key role in providing sensory input and validating the physical feasibility of the policy action.

3) *Approaches to Learning from Human Data*: Examples of learned human-like motions include walking using Generative Adversarial Imitation Learning (GAIL) [303, 307] and Adversarial Motion Prior (AMP) [304, 305, 326], which is an extension of the GAIL. Recently, RL-based motion imitation using motion capture data has successfully transferred to humanoid robots [267, 268, 324, 327]. However, most of these works achieve only relatively conservative loco-manipulation behaviors. Highly agile behaviors, as shown in DeepMimic [328], still exist only in simulators, and similar capabilities have yet to be replicated with real robots.

Human motion imitation can also achieve robust policies capable of rich interactions with objects in unstructured environments. For example, [306, 329] demonstrates the learning of full-body gymnastic skills for humanoids in physics

simulation, utilizing video reconstruction data and RL-based motion imitation. [330, 331] learns loco-manipulation policies enabling simulated humanoid characters to carry boxes. Furthermore, [302] mimics motion capture data to achieve tasks such as playing basketball and grasping objects. However, many of these approaches rely on privileged information, such as the ground truth of object and ego poses, from the simulator, which limits their extension to real-world hardware.

*Key Takeaways:* Transferring skills from humans, especially through Internet-scale datasets, unlocks a broad range of loco-manipulation capabilities for humanoid robots. Although learning from human data holds great promise, a large body of what has been achieved in simulation has yet to be realized in real-world robots. Developing affordable and capable humanoid robots and high-fidelity simulators of real-world loco-manipulation scenarios can accelerate progress in this research direction. As we witness an increasing accessibility of humanoid robots in the market, we foresee that imitation from human data will enable humanoid robots to acquire a vast and diverse skill set. This skill set is fundamental to building humanoid foundation models discussed in Sec. VIII.

#### D. Skill Learning: Hybrid Methods

Approaches that combine learning-based methods (IL and RL) and model-based methods are illustrated in Fig. 14.

1) *Combining Pure RL and IL*: The combination of IL and pure RL without demonstration data has led to effective sim-to-real transfers. A two-stage teacher-student paradigm is widely adopted [40, 42, 332]. In these works, a teacher policy is first trained from simulated privileged observation using pure RL. Then, a student policy clones the behavior of the teacher, achieving a similar performance using only partial observations. The trained student policy is readily deployable on hardware with onboard observations. Another two-stage paradigm [333] reverses the order of the two policies. It first uses IL to pre-train an imitation policy from expert data and then uses an RL policy to fine-tune the imitation policy, achieving performance beyond the IL expert and adapting to different environments or tasks [334].

2) *Learning to Track Trajectory Reference*: Combining model-based methods with learning-based methods is extensively explored. Numerous works on quadruped robots [335, 336] use MPC to generate reference motions and use them as imitation rewards for RL-based motion imitation. However, calculating MPC online can prolong the training time and occasionally fail to find a feasible solution. Online RL that tracks *offline-generated* trajectories avoids this problem and has achieved effective RL policies for versatile bipedal locomotion [38, 337, 338]. Supervised learning is also used to mimic TO-generated offline motions. For example, [339] allows humanoid robots to achieve brachiation on monkey bars. In addition to enhancing learning methods with model-informed trajectories, a learned policy can also suggest effective warm starts or hyperparameters for model-based methods, thus significantly reducing the iterations required for convergence to an optimal trajectory.

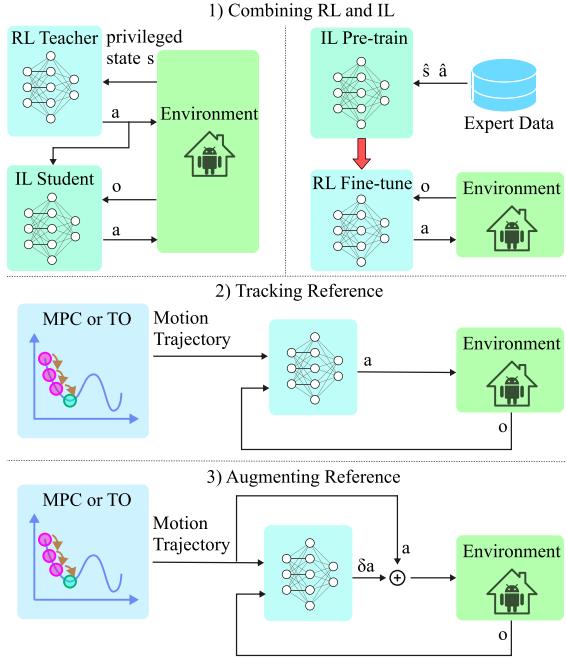


Fig. 14. The frameworks of combining model-based methods and model-free methods for skill learning.

**3) Learning to Augment Trajectory Reference:** Rather than tracking the reference trajectory, augmenting the reference with a residual is another popular approach. Early work on dynamic movement primitives [340] modifies the reference trajectory by learning a task-specific force output, which achieves a humanoid racket swing. A milestone in trajectory augmentation is the demonstration of bipedal locomotion in [341]. Since then, there have been more studies on trajectory augmentation with successful dynamic locomotion [342, 343]. Recently, loco-manipulation skills are also achieved through augmenting TO-based reference trajectory [344]. Beyond joint-level reference, a policy can also augment the task-space reference. For example, [345] learn a locomotion policy to adapt the foot placement reference derived from inverted pendulum models.

Both imitation and augmentation of the reference trajectory share benefits and drawbacks. As a benefit, references expedite the learning process and provide an effective way of learning complex skills. However, both methods rely on predefined trajectories and, therefore, have limited potential to learn emergent, diverse behaviors.

**Key Takeaways:** This subsection summarizes hybrid methods with successful humanoid hardware deployments. Hybrid methods that combine learning methods with and without demonstration are effective for learning complex and robust behaviors. We also show that model-based methods are not in conflict with learning-based methods, as they can provide guidance for efficient learning. Overall, hybrid methods that combine learning-based and model-based approaches have achieved superior efficiency, versatility, and performance in humanoid tasks that outperform single methods alone.

### E. Representation and Composition for Versatile Skills

A good representation of skills makes it easy to compose tasks. In general, a skill can be represented explicitly as a state-action trajectory that accomplishes a task or implicitly as a network structure and its learned weights. In this subsection, we explore implicit representations for skill composition to achieve versatile and multi-skill tasks.

Recent learning-based approaches enable smooth transitions between multiple skills. Among them, Mixture of Experts (MOE) is widely used. MOE employs a hierarchical architecture; it first trains multiple distinct skills, implicitly encoded in low-level expert policies, and then learns a high-level policy to select [346] or blend [47] these expert networks. This architecture allows for smooth transitions between skills and facilitates the completion of diverse tasks. However, MOE encounters expert imbalance issues favoring certain experts while degrading others, which could diminish the diversity provided by the experts. Instead of obtaining and then blending *multiple* policies, structured representations improve memory efficiency and allow a *single* policy to achieve multiple tasks. Next, we introduce three well-received structured representations: motion representation, goal representation, and state transition representation, all shown in Fig. 15.

**1) Motion Representation:** Motion representation extracts the essential features and temporal dependencies of high-dimensional long-horizon motions [347]. Specifically, motion representation encodes high-dimensional motions in a low-dimensional latent space. Such latent-space representations are commonly learned in an unsupervised manner using generative models such as Variational Autoencoders (VAEs) [348] and Generative Adversarial Networks (GANs) [349]. Compared to VAEs, GANs have greater potential to generate realistic motions following the reference data distribution, but they are often difficult to train. The result of learning motion representation is often a model that can synthesize versatile motions given latent codes. Therefore, the generative model can be reused for new downstream tasks by pairing it with a high-level task-specific policy. For example, several studies [350, 351, 352] learn a high-level RL policy to efficiently use a reduced-dimensional latent space motion representation and allow simulated humanoid robots to follow a set of user-commanded tasks.

**2) Goal Representation:** Another approach to learning a *single* policy for multiple tasks is through the representation of goals. The goal is typically represented as a feature vector, which can be encoded from an image of the scene in its final state, a natural language instruction, or a desired state from observing human demonstrations. This goal representation is often paired with Goal-Conditioned Policies (GCPs) [353]. Unlike standard RL policies that achieve only one task, GCPs achieve multiple tasks within a single general policy conditioned on different goals. GCPs have demonstrated versatile humanoid skills using IL, such as diffusion-model based BC [283] and RL with imitation objectives [267, 324, 354].

**3) State Transition Representation:** The latent space can also represent the transition dynamics of a Markov Decision Process (MDP). In this representation, data collected from the MDP are used to train a dynamics model that predicts the

transition probability between abstract, compressed representations of MDP states. This learned dynamics model is often referred to as a world model [355]. Sampling from a world model yields imaginary data, which can be achieved efficiently in parallel. By leveraging the imaginary data, Model-Based RL (MBRL) achieves greater efficiency compared to a typical model-free setting, where interaction data must be obtained from a simulated or real environment. MBRL has shown success in agile motor skills on humanoids [356, 357]. Another approach, TD-MPC2 [358], uses the world model in an MPC fashion, planning actions that lead to imaginary trajectories with high scores. Beyond data efficiency, the world model can mitigate the sim-to-real gap by fine-tuning a small batch of real-world data [359].

**Key Takeaways:** Enabling robots to accomplish versatile skills and multiple tasks is one of the main trends in robot skill learning. Whereas obtaining and blending single-skill policies is widely explored, more recent methods put efforts into achieving multiple tasks in a *single* policy. This requires a more structured representation of the skill motion, the task goal, and/or the environment dynamics. Latent space models, goal-conditioned policies, and world models are promising approaches toward this objective. However, many of these methods are still limited to the computer graphics community and yet to be implemented on humanoid hardware.

#### F. Learning for Humanoid Loco-manipulation

Lco-manipulation skills are challenging for learning-based methods as they often struggle with achieving physically stable contact or precise contact forces. While many learning-based approaches, such as CooHOI [360], demonstrate humanoid loco-manipulation skills in simulation, the physical interactions with external environments or objects are often oversimplified. As a result, only a few studies have demonstrated sim-to-real transfer for loco-manipulation skills. A large part of these studies rely on sim-to-real RL, with a few examples shown in Fig. 16. However, sim-to-real RL for loco-manipulation typically involves complicated reward designs that are fine-tuned for specific environments and tasks [361, 362], see Sec. VII-A. To enable loco-manipulation tasks, these methods define the contact sequence either implicitly using reference trajectory [344, 361] or explicitly via reward design [281], thus increasing the success rate of sim-to-real transfer. To address the uncertainty of the mass and other properties of the manipulated object, most RL approaches rely on domain randomization to provide robustness to object parameters [42].

A hierarchical RL structure that manages distinct skills is a viable strategy for achieving various loco-manipulation tasks such as falling recovery and ball kicking [282]. However, hierarchical RL has scalability issues, see Sec. VII-E. On the other hand, IL methods, particularly those employing RL with motion imitation, have enabled humanoid loco-manipulation via teleoperation, as outlined in Sec. VII-B. Although teleoperation is not autonomous, it is a crucial intermediate step to collect humanoid data. This has led to promising developments in *autonomous* loco-manipulation skills [267, 268],

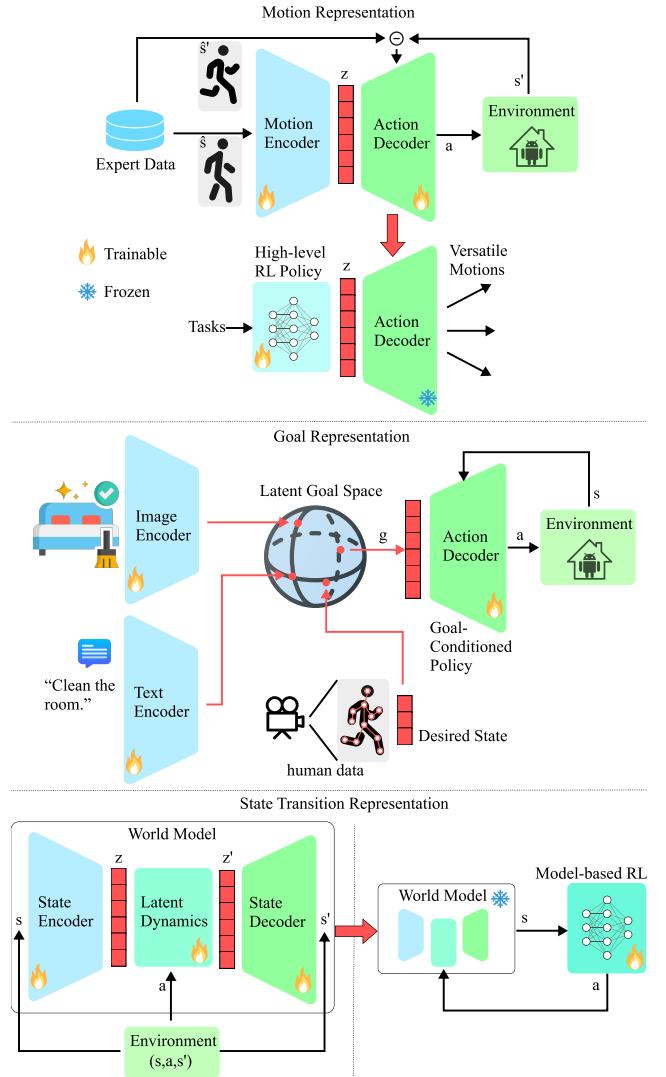


Fig. 15. Implicit skill representations. (a) Represent motions in latent space. (b) Represent goals in latent space and instruct a goal-conditioned policy for execution. (c) Represent state transitions with a world model and then use the learned world model for sample efficient training.

with the potential for further expansion to more diverse loco-manipulation tasks.

**Key Takeaways:** Although learning-based methods for humanoid loco-manipulation are less developed than model-based methods, their significance should not be overlooked. Learning-based methods are potentially more robust, as they can adapt to unstructured scenarios that model-based methods struggle to explicitly address, such as recovering from a fall in an arbitrary configuration [363]. Moreover, learning-based methods can find emergent behaviors that are challenging for model-based methods [291]. The recent success of quadrupedal loco-manipulation [81, 85, 86] shows significant promise for humanoids. However, transferring these algorithms from quadruped to humanoid is challenging because of the more complex dynamics of humanoids, which require enhanced safety measures and precise balance control.

### G. Challenge in Lack of Benchmarks for Loco-manipulation

The development of humanoid loco-manipulation skills is still in its infancy compared to more established tasks such as humanoid locomotion [364] and tabletop manipulation [365]. To advance this emerging field, there is a pressing need for large-scale, systematic benchmarks specifically designed for humanoid loco-manipulation, encompassing both simulation and real-world environments. Benchmarks that define loco-manipulation tasks and evaluation metrics can significantly accelerate research in this field.

For benchmarks in simulation, HumanoidBench [286] and Mimicking-Bench [366] offer a list of loco-manipulation tasks, such as sitting on a piece of furniture and wiping a window. SMPLOlympics [367] presents a collection of sports as tasks for simulated humanoids. These benchmarks provide metrics for evaluating the performance of humanoid algorithms and help verify the reproducibility within MuJoCo [214] and Isaac Gym [215]. However, these tasks represent only a subset of human skills, and the evaluation metrics remain highly task-specific. Expanding the benchmarks to include a broader range of real-world tasks, along with more standardized metrics, would enhance their utility and practicality. Moreover, the sim-to-real gap continues to limit the practical value and reliability of these simulation benchmarks: it remains uncertain how well performance in simulation translates to real-world deployment, although promising results have been demonstrated in zero-shot sim-to-real transfer.

For hardware benchmarks, the development of affordable and capable humanoid robots as standardized platforms for hardware evaluation can significantly accelerate the research. Research efforts on open-source humanoid hardware and software, such as Hector [2], the MIT Humanoid [179], the Berkeley Humanoid [368], and the Duke Humanoid [46] represent valuable contributions. Additionally, rapid prototyping of diverse humanoid robots facilitates cross-embodiment generalization, while mass production of standardized robots streamlines the deployment of benchmarks, such as through the Robotic Grand Factory initiatives [369]. Complementing these efforts, the construction of remote-accessible real-world environments, as exemplified by the DARPA Robotics Challenge, Robotarium [370], and TriFinger [371], enables researchers to focus on advancing algorithms. Ideally, these physical environments share the same benchmark metrics as the simulators to facilitate sim-to-real transfer, while remaining highly customizable to support novel tasks.

### H. Challenge in Data Scarcity

As discussed in Sec. VII-B, the four data sources present a trade-off between quality and availability. The lack of high-quality large-scale robot data becomes a bottleneck for robot skill learning. To solve the bottleneck, much effort has been put into data scaling. There is ongoing debate about whether scaling is the best approach to generalist humanoid robots.

The central question we must answer is which aspect of human motion we want robots to learn. Some humanoid tasks can be achieved simply by mimicking 3D human pose

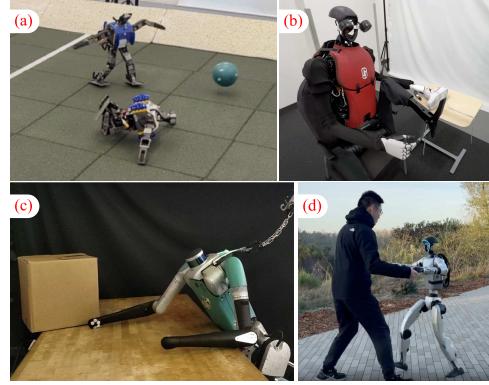


Fig. 16. Learning-based methods for humanoid loco-manipulation skills. (a) Getting up and chasing a ball [282]. (b) Tying shoelace [268]. (c) Multi-contact box manipulation [344]. (d) Dancing with a human [266].

trajectories [267, 268, 308, 311, 324], but a true general-purpose robot emerges from purposive learning: the ability to identify meaningful intentions from human data and adapt past experiences to new tasks or environments [372]. Therefore, human data must teach the robot not only *what* humans are doing, but also *how* and *why* they are doing it. Current data acquisition methods that capture human joint poses only enable learning what humans are doing. In this regard, imitation of human data at the trajectory level is not fundamentally generalizable due to the inevitable gap in morphology and in the surrounding environment.

We argue that generalization in loco-manipulation is achieved by including the motion of the manipulated objects, enhanced with a greater variety of sensing modalities instead of data quantity scaling. To develop truly versatile and adaptive humanoids, human data should also include that of the manipulated objects, cognitive actions (e.g., trust, compete, collaborate) paired with multimodal observations (e.g., whole-body haptic sensing, egocentric images), so that humanoids can learn the ‘how’ and ‘why’. However, instrumenting the environment and the manipulated objects with force and tactile sensing is highly complex and difficult to scale. Potential approaches include leveraging vision-language models to extract information from videos [373, 374, 375], employing video generation models [376] or physics-based simulation [214, 215] to generate synthetic data, and adopting hybrid strategies that effectively combine these synthetic data sources with limited real-world multimodal data [377, 378]. Recent work aims to bridge the gap between human animation and humanoid applications by collecting human multimodal observations [379, 380] and human kinetics [381]. Together with the rapid advancement in humanoid hardware, purposive learning with more informative human data will become the mainstream approach to achieve versatile and general-purpose humanoids.

## VIII. FOUNDATION MODELS FOR HUMANOID ROBOTS

Foundation Models (FMs) are large pre-trained models using Internet-scale data [382]. Recent progress in FMs such as Large Language Models (LLMs) and Vision-Language Models (VLMs) has demonstrated groundbreaking capabilities in

solving a wide range of downstream tasks (through in-context learning or fine-tuning), such as code generation and video understanding [383]. The common sense reasoning capabilities of FMs have inspired many explorations of their applications in robotics [11, 12]. Despite this growing interest, research on FMs specifically for humanoid robots remains sparse. In this section, we first provide an overview of FMs in the context of general robotics (*e.g.*, mobile manipulation, instead of humanoids) and then explore their potential applications to humanoid robots.

Strategies to leverage FMs for robotics can generally be categorized into two, as shown in Fig. 17. The first strategy (Sec. VIII-A) elicits actionable knowledge from pre-trained LLMs/VLMs for robotic tasks, without additional model fine-tuning. The second strategy (Sec. VIII-B) collects abundant robotic data to fine-tune or co-train a *Robot Foundation Model* (RFM) that generalizes to control tasks with common sense reasoning capability [384, 385, 386, 387, 388, 389].

#### A. Applying LLMs/VLMs to Humanoid Robots

Applying LLMs/VLMs to humanoids is still a nascent field. Many studies have demonstrated the successful deployment of LLMs/VLMs across various robot embodiments, such as dexterous hands [390], manipulators [391], mobile manipulators [392], quadrupedal robots [393], and bipedal robots [394].

Among these works, a majority way of using LLMs/VLMs is to leverage pre-trained models without robot data. Although these pre-trained models have semantic understanding capability and context awareness, they often lack embodied knowledge and can prescribe actions that are ambiguous or non-admissible. Therefore, considerable research efforts have focused on task-planning mechanisms to enable the generation of admissible action plans. For example, SayCan [392] ranks available actions based on value functions obtained during the training of corresponding action policies for mobile manipulators. VLM-PC [395] restricts GPT-4o to output plans with skills available only for quadruped navigation.

1) *FMs for selecting low-level skills (task planning):* Such task planning capability of FMs has advanced the complexity of tasks that a humanoid can accomplish. For instance, [396] and [397] use pre-trained LLMs to select skills and task goals for animated humanoid characters. OmniH2O [267] employs GPT-4 [383] to select autonomous skills such as greeting a human. HYPERmotion [398] applies an LLM to construct task graphs that enable a hybrid wheeled-leg robot to execute complex loco-manipulation tasks. However, using FMs with a fixed skill set limits skill versatility. In addition, for complex behaviors such as those in humanoid loco-manipulation tasks, it is essential to allow FMs to author the detailed motions of low-level skills, instead of selecting from the existing, relatively-abstract low-level skills.

2) *FMs for authoring low-level skills:* Thus, many research efforts have focused on identifying the best bridge between FMs and low-level robot skills. For example, researchers have proposed to generate code [391, 399] and reward functions [393, 394, 400] as intermediate representations for bipedal and quadrupedal robots. Compared to

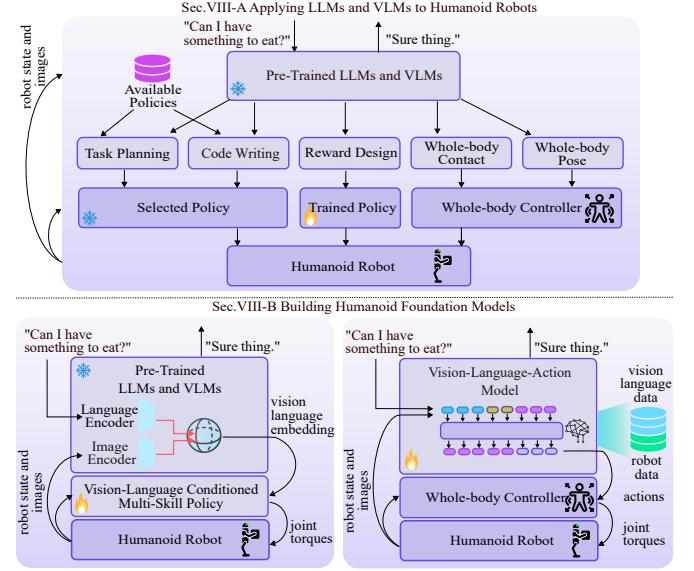


Fig. 17. Sec. VIII-A details different approaches to applying LLMs and VLMs in robotics tasks. These approaches prompt pre-trained LLMs and VLMs to generate task-relevant intermediate representations that can be executed by low-level policies or controllers. Sec. VIII-B presents two approaches to building humanoid foundation models. The first approach uses robot data to train multi-skill policy, and the second approach fine-tunes an existing VLM. Both approaches directly output actions for low-level humanoid control.

selecting the existing skills, these intermediate representations provide additional flexibility in adjusting the generated motion. Furthermore, FMs can generate whole-body poses [401, 402] and whole-body contacts [403, 404] for humanoid robots. These techniques allow users to intuitively direct a robot’s behavior through expressive inputs such as natural language, images, or even gestures.

#### B. Building Humanoid Foundation Models

While most FMs are developed in the vision or language domain, building them in the robotics domain for embodied intelligence is a natural extension. Similar to LLM and VLM, Robot Foundation Models (RFMs) are large models trained from Internet-scale robotic datasets. These RFMs often process multimodal inputs (*e.g.*, egocentric images, and natural language as task description) and enable interaction with the physical world through robot action outputs. Therefore, they are also called Vision-Language-Action (VLA) models. Leveraging the Internet-scale multimodal dataset, RFMs hold the promise of generalization across diverse tasks and provide a natural interface for human-robot interaction, both essential for real-world robot applications.

1) *Hierarchical vision-language-conditioned policy:* One popular approach to building RFMs is to create a multi-task sensorimotor control policy using a high-capacity model that can consume large amounts of robot data. The trained policy can perform a wide variety of low-level skills, even for multiple robot embodiments. To enable high-level reasoning/planning, this approach often employs a hierarchical framework combining a pre-trained LLM or VLM with the low-level control policy, initialized from scratch. During training, this setup aligns the semantic knowledge of the LLM

and VLM with the physical behavior of the control policy, enabling cross-modal capabilities such as language-to-action. Transformer [405] is a common choice for such a low-level policy due to its scalability. To enable interactive user commands, the low-level policy is conditioned on language and/or image inputs, often tokenized with pre-trained text or image encoders from LLMs or VLMs. For example, RT-1 [384] and VIMA [386] both train a language-conditioned visuomotor policy with a large amount of manipulation data and have demonstrated the ability to perform a wide range of skills. However, successful implementations are limited to robots that have stable dynamics and a large amount of high-quality data collected with significant resources [385, 387, 406, 407].

Applying such a hierarchical framework to humanoid robots is an exciting new direction. Recent works have extended RFMs to the humanoid upper body [377, 388, 389]. For example, NVIDIA developed GR00T N1 [377], a general-purpose FM for humanoid robots, and Figure AI demonstrates Helix, a hierarchical VLA for dexterous manipulation and collaboration between two humanoid robots. However, building an RFM for humanoid loco-manipulation remains a challenging endeavor: the inherent instability in locomotion and high-dimensional action space in dexterous hands makes it difficult to collect high-quality data efficiently. Therefore, VLA for humanoid loco-manipulation is shown only in simulation. For example, HumanVLA [312] and SuperPADL [313] train a humanoid action policy aligned with the latent space of pre-trained VLMs, enabling skills based on image and language inputs.

2) *End-to-end Vision-Language-Action model*: To leverage prior knowledge in FMs, another popular approach is to build an end-to-end VLA model, as exemplified by RT-2 [385], OpenVLA [387], and Gato [408]. This approach treats robot data (*i.e.*, observations and actions) as tokens in the language model’s vocabulary, allowing direct fine-tuning or co-training with existing VLMs. Unlike the previous approach, VLA outputs actions as tokens directly without relying on trainable low-level policies. The VLA model not only generates robot actions for diverse skills, but also retains semantic reasoning abilities in the language and vision domain, enhancing its generalization capacity compared to models trained in single domains. For example, RT-2 [385] represents the actions as a string of numbers similar to the tokens from the pre-trained vision and language tokenizer of the base VLMs. However, representing actions with stringified numbers can be token-inefficient with high degrees-of-freedom robots such as humanoids.

A key research question in building RFMs is the design of effective algorithms and model architectures. Although most of today’s RFMs are based on autoregressive transformer models [405], their computational inefficiency over long sequences poses a significant challenge for both training and inference. This has driven the exploration of alternative models that are both efficient and high-capacity, such as state-space models [409].

The key to the success of training an RFM, especially in the case of VLA, heavily depends on the choice of input and output representations. Most RFMs take as input a combination of task descriptions in language, visual observations of

the surrounding environment, and the history of robot states. Outputs typically consist of robot actions, which are derived from either a multi-task policy (the first approach) or an end-to-end VLA model (the second approach). There are variations in the VLA model where its token outputs specify more than just actions. For example, Octo [410] and RDT-1B [411] use the token output for a diffusion denoising process.  $\pi_0$  [406] maps a learned token to robot actions through a diffusion head, enabling high-frequency control (up to 50 Hz for a bimanual manipulator with a wheelbase). GR-2 [412] predicts tokens that represent future images and actions; thus it functions as both a world model and a visuomotor policy. For humanoid robots, input and output are not yet well defined. For tasks involving rich physical interactions, the force feedback is as crucial as egocentric visual input. Given the unstable dynamics of humanoid robots, a more practical approach is to use end-effector and body poses as action outputs from the RFM and adopt additional low-level policies to ensure balance and safety via high-frequency feedback control. Incorporating robotic data as a new modality into state-of-the-art FMs would require significantly more data. Hence, designing effective input/output representation for humanoid loco-manipulation tasks still remains an open research question.

### C. Opportunities and Challenges in Foundation Models

Integrating Foundation Models (FMs) into humanoid robots offers distinct opportunities and challenges. On the opportunity side, since the majority of data used to train FMs is collected by humans, the knowledge embedded in these models is inherently biased towards human-like embodiments. Consequently, humanoid robots could potentially utilize existing knowledge in FMs more effectively due to a smaller embodiment gap. This advantage extends beyond planning and control capabilities to include interactions with humans using natural modalities such as language and gestures.

A major challenge in applying FMs to humanoids arises from the high inference cost. Running large foundation models using only onboard computing is not feasible due to the limited power and computation, which hampers real-time hardware control. To address this challenge, several solutions have been proposed. One effective strategy involves adopting a decentralized hierarchy [388], where FMs operate over the cloud and provide only high-level decisions at a lower frequency, while another controller remains onboard and manages real-time task execution. However, the inference delay and internet latency may impede the control performance. Another approach is to enhance the speed of the computing platform and the efficiency of FMs. For instance, NVIDIA introduced Jetson Thor, an onboard computing platform designed for humanoid robots. Google proposed SARA-RT [413], which accelerates the model speed without compromising its quality.

The training of FM is also resource and time consuming. For example, training the LLaMA model took 34 days on 992 NVIDIA A100-80B GPUs [414], which incurs high cost, high energy consumption, as well as carbon dioxide emission. As FMs continue to scale up, the training cost increases further. A promising approach to maintaining a reasonable cost

for training robotics FMs is to leverage parameter-efficient fine-tuning techniques. For example, OpenVLA leverages the LoRA technique to fine-tune an FM with a robotics dataset [387].

A critical component of the humanoid foundation model that has yet to be developed is a scaling law, similar to the training of large language models [415]. The scaling law provides guidance on how we should scale up model, compute, and data, to meet the desired performance in the most efficient way. A major research effort focuses on scaling the robotics dataset. Open X-Embodiment extends the idea to a much larger robotics dataset across various robot embodiments and tasks [407]. Recent work has already explored a data scaling law for robot manipulation, with a focus on generalization capabilities [416], as well as model scaling behaviors for zero-shot capabilities in action selection [417], which marks important initial steps towards this direction.

Beyond computational limitations, FMs face additional challenges when deployed in humanoid systems. Safety constraints are particularly concerning, as models trained on internet data may generate unsafe actions when controlling physical robots that are inherently unstable. Establishing a thorough evaluation procedure to assess the safety level of humanoid robots will be essential. Moreover, ethical considerations arise from biases present in human-generated training data, which can lead to unfair or discriminatory robot behaviors across different user demographics. These biases may be amplified in embodied systems that interact directly with humans in diverse settings. Federated learning approaches could help address some of these ethical concerns by enabling model training across distributed devices while preserving data privacy and potentially reducing centralized data biases [418].

**Key Takeaways:** With rapid progress in FMs and humanoid robots, a plethora of research studies on humanoid embodied intelligence is anticipated. The first method outlined in Sec. VIII-A provides a more accessible way to leverage VLMs and, thus, has been the popular approach taken by researchers so far, but VLMs as is lack a deep understanding of robot actions. In the near future, we expect increased research efforts devoted to VLA models described in Sec. VIII-B to address this issue. It takes great effort to train RFMs from (i) a large amount of real robot data (e.g., Gemini-Robotics [388, 389]) or (ii) a combination of robot data with synthetic data from simulation or video generation models (e.g., GR00T N1 [377]). With the development of RFMs, we expect to have FMs that can well understand robot action and the physical world, enabling rapid deployment of robotic controllers achieved with little to no additional robot data. For further reading of FMs for robotics, we recommend the survey in [11, 12]. Since most of the FMs are built with the Transformer backbones, please refer to [405] for a comprehensive mathematical foundation.

## IX. ADDITIONAL DISCUSSIONS

### A. Model-based Methods Versus Learning-based Methods

Based on the studies in this survey, we provide a clear comparison between model-based and learning-based methods in Table. VI. In this table, we compare the performance in

TABLE VI  
COMPARISON BETWEEN MODEL-BASED AND LEARNING-BASED METHODS

	Model-based MPC, WBC	Learning-based		
		RL	IL	VLA
Performance measure				
Flexibility to incorporate multi-modal sensor	low	medium	high	very high
Robustness	high	very high	low	very low
Motion accuracy	very high	medium	medium	very low
Real-time feasibility	high	high	high	low
Versatility	high	medium	high	high
Generalizability	medium	low	low	high
Achieved performance of humanoid skills				
Locomotion	very high [6, 30]	very high [37, 40]	medium [283]	NA
Loco-manipulation (hand and foot contact)	high [2, 218]	medium [41, 281]	low [268]	NA
Whole-body multi-contact	medium [140]	high [419]	NA	NA

several aspects: the ability to incorporate multi-modal sensor input, the control robustness, the motion accuracy, the real-time feasibility, the versatility, and the generalizability of each method. Additionally, we evaluate the state-of-the-art capability in several humanoid skills, including locomotion, loco-manipulation, and whole-body multi-contact control. We list representative papers that achieved these skills as supporting evidence for this evaluation. Note that many algorithm performances reported are anecdotal and highly depend on their implementation. Therefore, obtaining a concrete numerical comparison is challenging. This is mentioned in Sec. VII-G challenges in lack of appropriate benchmark.

A key takeaway from this comparison is that learning-based methods exhibit superior robustness via RL and versatility via IL, while model-based approaches offer advantages in motion accuracy. The emerging paradigm of VLA holds promise for enhanced generalizability; however, its effectiveness in mastering humanoid skills has yet to be convincingly demonstrated.

### B. Hardware Limitations for Humanoid Robots

Advancements in control, planning, and learning for loco-manipulation necessitate robot hardware that has high-power-density actuators, strong yet lightweight limbs, and accurate sensing. These requirements are often limitations to hardware performance, which drives hardware designers to push the boundary of humanoid physical capabilities.

The traditional approach to increase torque capacity is through high gear reduction [420] and the inclusion of elastic elements to be impact-resilient [28]. However, the last decade has witnessed advancements in electric motor design such as Quasi-Direct Drive (QDD) actuators (less than 10:1 gear ratio) achieving dynamic and accurate motions [2, 368]. Among the many advantages QDD actuators offer for achieving dynamic motions are their backdrivability and high force control bandwidth, due to reduced actuator backlash and friction [179]. However, achieving high torque requires a high current, which in turn causes strain on power electronics and overheating. Solving this problem involves manufacturing with materials to enhance heat dissipation [420] and continuing to improve controller performance for energy efficiency [43].

Another key element of hardware design is the robot's inertial properties. In building robots capable of gymnastic-style maneuvers, many designs place mass away from end-effectors [2, 46, 179]. The MIT Humanoid, for example, contains around 75% of its mass in the torso [179]. This is often achieved through belt drives or parallel mechanisms to remotely actuate joints such as the ankle or knee [2, 46]. While this significantly reduces rotational inertia and the torque demand for agile maneuvers, there are a number of notable drawbacks. Four-bar transmissions often reduce joint range of motion, while belt drives may require additional modeling and maintenance [420]. Specialized mechanisms for remote actuation that allow a full range of motion can potentially address this problem [421].

Battery technology remains one of the primary hardware limitations in the development of sustained, autonomous humanoid robots. Limited battery capacity restricts both operational time and the feasibility of executing high-power, full-body motions in untethered scenarios. One promising direction is the development of more compact and integrated battery designs. For example, embedding battery cells directly into the structural components of the robot chassis can increase effective energy capacity without compromising weight distribution or form factor. In parallel, solid-state batteries have become commercially available with enhanced energy density and safety profiles [422]. This has increased the operation times of humanoid robots, such as Apptronik's Apollo (4hrs) [423] and Unitree's G1 (2hrs) [424]. Nevertheless, the continued reliance on tethered power in laboratory settings [1, 4, 5, 46] indicates that battery technology remains a limitation.

### C. The Future Applications and Challenges of Humanoid Robots in the Society

Humanoid robots with loco-manipulation capabilities will become an essential part of our lives and a pillar of the future economy, similar to how smartphones and autonomous driving have transformed our lives.

Given a sufficient amount of time, humanoid robots will be increasingly autonomous and capable of performing general human-level tasks across a wide range of domains. The components discussed in this paper—sensing, planning, control, and learning—coordinate in real time to execute complex, goal-directed behaviors. This integration enables humanoid robots to take on tedious, hazardous, and physically demanding tasks.

- In outdoor environments, they could serve as first responders during national crises, providing medical care and emergency support; monitor controlled settings in civil and mechanical infrastructure; deliver and sort building materials at construction sites; explore subterranean or collapsed structures with harsh and unpredictable terrain; and assist in agricultural tasks such as planting crops.
- In factories and warehouses, humanoid robots have the potential to address the growing demand for flexible manufacturing by performing tasks such as delivery, assembly, and inspection, helping to offset the impact of a declining industrial workforce.

- In human-centered settings such as households and hospitals, they have the potential to carry laundry, load dishwashers, and move patients between beds, supporting the needs of an aging population and improving their quality of life.

However, significant challenges remain in integrating humanoid robots into all aspects of society. To be fully accepted, these robots must be physically and mentally reliable, highlighting the core technical challenges of achieving robust and safe mobility and manipulation. Their intentions must also be well predicted and communicated with humans, which calls for advancements in effective human-robot interaction. Beyond the technical hurdles, economic and ethical concerns also pose major obstacles. Humanoid robots are still far from being commercially viable due to high costs, and the absence of legal and regulatory policies further hinders their appearance in our daily lives.

## X. CONCLUSION

Humanoid robotics is advancing at an unprecedented pace, as seen in recent groundbreaking innovations from both industry and academia. In this survey, we go over foundational and state-of-the-art methods for humanoid loco-manipulation, from both model-based and learning-based perspectives. Although humanoid robots still face significant technical challenges, their agility, safety, reliability, and versatility have been improved significantly. More opportunities are emerging through the exploration of new paradigms. Physically, advanced observers (vision and whole-body tactile sensing) and estimators are emerging for contact-rich whole-body loco-manipulation. Cognitively, foundation models grounded on humanoid robots have great potential to unlock the ability of open-world understanding and the development of generalized intelligent agents. More importantly, an integration of foundation models with planning and control policies presents promising capabilities. In the foreseeable future, the cost of humanoid robots will continue to decrease, making them more accessible; their physical capability (hardware intelligence) and cognitive intelligence will significantly advance. We look forward to humanoid robots that are responsive, purposeful, and fully capable of human-like loco-manipulation tasks in the upcoming decade.

## REFERENCES

- [1] S. Samadi, J. Roux, A. Tanguy, S. Caron, and A. Kheddar, "Humanoid control under interchangeable fixed and sliding unilateral contacts," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 4032–4039, 2021.
- [2] J. Li, J. Ma, O. Kolt, M. Shah, and Q. Nguyen, "Dynamic loco-manipulation on hector: Humanoid for enhanced control and open-source research," *arXiv preprint arXiv:2312.11868*, 2023.
- [3] B. Henze, M. A. Roa, and C. Ott, "Passivity-based whole-body balancing for torque-controlled humanoid robots in multi-contact scenarios," *The International Journal of Robotics Research*, vol. 35, no. 12, pp. 1522–1543, 2016.
- [4] N. Figueroa, S. Faraji, M. Koptev, and A. Billard, "A dynamical system approach for adaptive grasping, navigation and co-manipulation with humanoid robots," in *IEEE International Conference on Robotics and Automation*, 2020, pp. 7676–7682.
- [5] D. Calvert, L. Penco, D. Anderson, T. Bialek, A. Chatterjee, B. Mishra, G. Clark, S. Bertrand, and R. Griffin, "A behavior architecture for fast humanoid robot door traversals," *arXiv preprint arXiv:2411.03532*, 2024.

[6] C. Khazoom, S. Hong, M. Chignoli, E. Stanger-Jones, and S. Kim, “Tailoring solution accuracy for fast whole-body model predictive control of legged robots,” *IEEE Robotics and Automation Letters*, vol. 9, no. 12, pp. 11 074–11 081, 2024.

[7] A. Goswami and P. Vadakkepat, *Humanoid Robotics: A Reference*, 1st ed. Springer Publishing Company, Incorporated, 2018.

[8] L. Zhao, Z. Yu, L. Han, X. Chen, X. Qiu, and Q. Huang, “Compliant motion control of wheel-legged humanoid robot on rough terrains,” *IEEE/ASME Transactions on Mechatronics*, vol. 29, no. 3, pp. 1949–1959, 2024.

[9] P. M. Wensing, M. Posa, Y. Hu, A. Escande, N. Mansard, and A. D. Prete, “Optimization-based control for dynamic legged robots,” *IEEE Transactions on Robotics*, vol. 40, pp. 43–63, 2024.

[10] C. Chi, Z. Xu, S. Feng, E. Cousineau, Y. Du, B. Burchfiel, R. Tedrake, and S. Song, “Diffusion policy: Visuomotor policy learning via action diffusion,” *The International Journal of Robotics Research*, vol. 0, no. 0, p. 02783649241273668, 2024.

[11] R. Firooz *et al.*, “Foundation models in robotics: Applications, challenges, and the future,” *The International Journal of Robotics Research*, vol. 0, no. 0, p. 02783649241281508, 2024.

[12] Y. Hu *et al.*, “Toward general-purpose robots via foundation models: A survey and meta-analysis,” *CoRR*, vol. abs/2312.08782, 2023.

[13] R. S. Dahiya, G. Metta, M. Valle, and G. Sandini, “Tactile sensing—from humans to humanoids,” *IEEE Transactions on Robotics*, vol. 26, no. 1, pp. 1–20, 2009.

[14] M. M. Sotaro Katayama and Y. Tazaki, “Model predictive control of legged and humanoid robots: models and algorithms,” *Advanced Robotics*, vol. 37, no. 5, pp. 298–315, 2023.

[15] L. Bao, J. Humphreys, T. Peng, and C. Zhou, “Deep reinforcement learning for bipedal locomotion: A brief survey,” 2024.

[16] S. Ha, J. Lee, M. van de Panne, Z. Xie, W. Yu, and M. Khadiv, “Learning-based legged locomotion: State of the art and future perspectives,” *The International Journal of Robotics Research*, vol. 0, no. 0, p. 02783649241312698, 0.

[17] A. Gams, T. Petrič, B. Nemeć, and A. Ude, “Manipulation learning on humanoid robots,” *Current Robotics Reports*, vol. 3, no. 3, pp. 97–109, 2022.

[18] J. Parra-Moreno, D. Yanguas-Rojas, and E. Mojica-Nava, “Imitation methods for bipedal locomotion of humanoid robots: A survey,” *IEEE Colombian Conference on Automatic Control*, pp. 1–6, 2023.

[19] Y. Tong, H. Liu, and Z. Zhang, “Advancements in humanoid robots: A comprehensive review and future prospects,” *IEEE/CAA Journal of Automatica Sinica*, vol. 11, no. 2, pp. 301–328, 2024.

[20] P.-B. Wieber, R. Tedrake, and S. Kuindersma, “Modeling and control of legged robots,” in *handbook of robotics*. Springer, 2016, pp. 1203–1234.

[21] J. Carpentier and P.-B. Wieber, “Recent progress in legged robots locomotion control,” *Current Robotics Reports*, vol. 2, no. 3, pp. 231–238, 2021.

[22] M. S. Khan and R. K. Mandava, “A review on gait generation of the biped robot on various terrains,” *Robotica*, vol. 41, no. 6, pp. 1888–1930, 2023.

[23] E. R. Westervelt, J. W. Grizzle, C. Chevallereau, J. H. Choi, and B. Morris, *Feedback Control of Dynamic Bipedal Robot Locomotion*. CRC Press, 2018.

[24] S. Kajita, H. Hirukawa, K. Harada, and K. Yokoi, *Introduction to Humanoid Robotics*. Springer Berlin, Heidelberg, 2014.

[25] T. McGeer, “Passive walking with knees,” in *Proceedings., IEEE International Conference on Robotics and Automation*, 1990, pp. 1640–1645 vol.3.

[26] S. Collins, A. Ruina, R. Tedrake, and M. Wisse, “Efficient bipedal robots based on passive-dynamic walkers,” *Science*, vol. 307, no. 5712, pp. 1082–1085, 2005.

[27] 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 *IEEE International Conference on Robotics and Automation*, vol. 2, 2003, pp. 1620–1626.

[28] D. Kim, S. J. Jorgensen, J. Lee, J. Ahn, J. Luo, and L. Sentis, “Dynamic locomotion for passive-ankle biped robots and humanoids using whole-body locomotion control,” *The International Journal of Robotics Research*, vol. 39, no. 8, pp. 936–956, 2020.

[29] J. Li, O. Kolt, and Q. Nguyen, “Continuous dynamic bipedal jumping via adaptive-model optimization,” *arXiv preprint arXiv:2404.11807*, 2024.

[30] P. Wensing, “Optimization and control of dynamic humanoid running and jumping,” Ph.D. dissertation, The Ohio State University, 2014.

[31] S. Pardis, M. Chignoli, and S. Kim, “Probabilistic homotopy optimization for dynamic motion planning,” in *International Conference on Intelligent Robots and Systems*, 2024, pp. 12 856–12 863.

[32] J. Englsberger, A. Dietrich, G.-A. Mesesan, G. Garofalo, C. Ott, and A. O. Albu-Schäffer, “Mptc-modular passive tracking controller for stack of tasks based control frameworks,” *Robotics: Science and Systems*, 2020.

[33] D. J. Agravante, A. Cherubini, A. Sherikov, P.-B. Wieber, and A. Kheddar, “Human-humanoid collaborative carrying,” *IEEE Transactions on Robotics*, vol. 35, no. 4, pp. 833–846, 2019.

[34] J. Morimoto and C. G. Atkeson, “Learning biped locomotion,” *IEEE Robotics and Automation Magazine*, vol. 14, no. 2, pp. 41–51, 2007.

[35] J. Nakanishi, J. Morimoto, G. Endo, G. Cheng, S. Schaal, and M. Kawato, “Learning from demonstration and adaptation of biped locomotion,” *Robotics and Autonomous Systems*, vol. 47, no. 2, pp. 79–91, 2004.

[36] R. Calandra, A. Seyfarth, J. Peters, and M. P. Deisenroth, “Bayesian optimization for learning gaits under uncertainty: An experimental comparison on a dynamic bipedal walker,” *Annals of Mathematics and Artificial Intelligence*, vol. 76, no. 1–2, p. 5–23, 2015.

[37] D. Crowley, J. Dao, H. Duan, K. Green, J. Hurst, and A. Fern, “Optimizing bipedal locomotion for the 100m dash with comparison to human running,” in *IEEE International Conference on Robotics and Automation*, 2023, pp. 12 205–12 211.

[38] Z. Li, X. B. Peng, P. Abbeel, S. Levine, G. Berseth, and K. Sreenath, “Reinforcement learning for versatile, dynamic, and robust bipedal locomotion control,” *The International Journal of Robotics Research*, vol. 0, no. 0, p. 02783649241285161, 2024.

[39] J. Siekmann, K. Green, J. Warila, A. Fern, and J. Hurst, “Blind bipedal stair traversal via sim-to-real reinforcement learning,” in *Robotics: Science and Systems*, 2022.

[40] Z. Zhuang, S. Yao, and H. Zhao, “Humanoid parkour learning,” in *Annual Conference on Robot Learning*, 2024.

[41] J. Dao, K. Green, H. Duan, A. Fern, and J. Hurst, “Sim-to-real learning for bipedal locomotion under unsensed dynamic loads,” in *International Conference on Robotics and Automation*, 2022, pp. 10 449–10 455.

[42] I. Radosavovic, T. Xiao, B. Zhang, T. Darrell, J. Malik, and K. Sreenath, “Real-world humanoid locomotion with reinforcement learning,” *Science Robotics*, vol. 9, no. 89, 2024.

[43] N. Kashiri *et al.*, “An overview on principles for energy efficient robot locomotion,” *Frontiers in Robotics and AI*, vol. 5, 2018.

[44] J. Reher, E. A. Cousineau, A. Hereid, C. M. Hubicki, and A. D. Ames, “Realizing dynamic and efficient bipedal locomotion on the humanoid robot durus,” in *IEEE International Conference on Robotics and Automation*, 2016, pp. 1794–1801.

[45] F. Heremans, S. Vijayakumar, M. Bouri, B. Dehez, and R. Ronse, “Bio-inspired design and validation of the efficient lockable spring ankle (ELSA) prosthesis,” *IEEE Int Conf Rehabil Robot*, vol. 2019, pp. 411–416, Jun. 2019.

[46] B. Xia, B. Li, J. Lee, M. Scutari, and B. Chen, “The duke humanoid: Design and control for energy efficient bipedal locomotion using passive dynamics,” *arXiv preprint arXiv:2409.19795*, 2024.

[47] C. Yang, K. Yuan, S. Heng, T. Komura, and Z. Li, “Learning natural locomotion behaviors for humanoid robots using human bias,” *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2610–2617, 2020.

[48] J.-K. Huang and J. W. Grizzle, “Efficient anytime clif reactive planning system for a bipedal robot on undulating terrain,” *IEEE Transactions on Robotics*, vol. 39, no. 3, pp. 2093–2110, 2023.

[49] K. Muenprasitvej, J. Jiang, A. Shamsah, S. Coogan, and Y. Zhao, “Bipedal safe navigation over uncertain rough terrain: Unifying terrain mapping and locomotion stability,” *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2024.

[50] S. McCrory, B. Mishra, R. Griffin, J. Pratt, and H. E. Sevil, “Bipedal navigation planning over rough terrain using traversability models,” in *SoutheastCon*, 2023, pp. 89–95.

[51] Z. Li, J. Zeng, S. Chen, and K. Sreenath, “Autonomous navigation of underactuated bipedal robots in height-constrained environments,” *The International Journal of Robotics Research*, vol. 42, no. 8, pp. 565–585, 2023.

[52] A.-C. Cheng, Y. Ji, Z. Yang, X. Zou, J. Kautz, E. Biyik, H. Yin, S. Liu, and X. Wang, “Navila: Legged robot vision-language-action model for navigation,” *arXiv preprint arXiv:2412.04453*, 2024.

[53] B. Acosta and M. Posa, “Bipedal walking on constrained footholds with mpc footstep control,” in *IEEE-RAS International Conference on Humanoid Robots*, 2023, pp. 1–8.

[54] H. Duan, B. Pandit, M. S. Gadde, B. Van Marum, J. Dao, C. Kim, and A. Fern, “Learning vision-based bipedal locomotion for challenging terrain,” in *IEEE International Conference on Robotics and Automata*

tion, 2024, pp. 56–62.

[55] R. J. Griffin, G. Wiedebach, S. McCrory, S. Bertrand, I. Lee, and J. Pratt, “Footstep planning for autonomous walking over rough terrain,” in *IEEE-RAS International Conference on Humanoid Robots*, 2019, pp. 9–16.

[56] C. Peng, O. Donca, G. Castillo, and A. Hereid, “Safe bipedal path planning via control barrier functions for polynomial shape obstacles estimated using logistic regression,” in *IEEE International Conference on Robotics and Automation*, 2023, pp. 3649–3655.

[57] Y. Zhao, U. Topcu, and L. Sentis, “High-level planner synthesis for whole-body locomotion in unstructured environments,” in *IEEE 55th Conference on Decision and Control*, 2016, pp. 6557–6564.

[58] A. Shamsah, Z. Gu, J. Warnke, S. Hutchinson, and Y. Zhao, “Integrated task and motion planning for safe legged navigation in partially observable environments,” *IEEE Transactions on Robotics*, vol. 39, no. 6, pp. 4913–4934, 2023.

[59] A. Shamsah, K. Agarwal, N. Katta, A. Raju, S. Kousik, and Y. Zhao, “Socially acceptable bipedal robot navigation via social zonotope network model predictive control,” *IEEE Transactions on Automation Science and Engineering*, 2025.

[60] W. Zhu, A. Raju, A. Shamsah, A. Wu, S. Hutchinson, and Y. Zhao, “Emobipednav: Emotion-aware social navigation for bipedal robots with deep reinforcement learning,” *arXiv preprint arXiv:2503.12538*, 2025.

[61] G. Gibson, O. Dosunmu-Ogunbi, Y. Gong, and J. Grizzle, “Terrain-adaptive, alip-based bipedal locomotion controller via model predictive control and virtual constraints,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2022, pp. 6724–6731.

[62] Z. Xia, J. Xiong, and K. Chen, “Global navigation for humanoid robots using sampling-based footstep planners,” *IEEE/ASME Transactions on Mechatronics*, vol. 16, no. 4, pp. 716–723, 2011.

[63] A. Shamsah, J. Jiang, Z. Yoon, S. Coogan, and Y. Zhao, “Terrain-aware model predictive control of heterogeneous bipedal and aerial robot coordination for search and rescue tasks,” 2024.

[64] J. Jiang, S. Coogan, and Y. Zhao, “Abstraction-based planning for uncertainty-aware legged navigation,” *IEEE Open Journal of Control Systems*, vol. 2, pp. 221–234, 2023.

[65] M. Ahmadi, X. Xiong, and A. D. Ames, “Risk-averse control via cvar barrier functions: Application to bipedal robot locomotion,” *IEEE Control Systems Letters*, vol. 6, pp. 878–883, 2021.

[66] C. Smith, Y. Karayannidis, L. Nalpantidis, X. Gratal, P. Qi, D. V. Dimarogonas, and D. Kragic, “Dual arm manipulation—a survey,” *Robotics and Autonomous systems*, vol. 60, no. 10, pp. 1340–1353, 2012.

[67] N. S. Pollard, “Parallel methods for synthesizing whole-hand grasps from generalized prototypes,” Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, 1994.

[68] J. K. Salisbury, “Whole-arm manipulation,” in *Proceeding of the International Symposium of Robotics Research*. The MIT Press, 1987.

[69] W. T. Townsend and J. K. Salisbury, “Mechanical design for whole-arm manipulation,” in *Robots and Biological Systems: Towards a New Bionics*. Springer, 1993, pp. 153–164.

[70] X. Cheng, S. Patil, Z. Temel, O. Kroemer, and M. T. Mason, “Enhancing dexterity in robotic manipulation via hierarchical contact exploration,” *IEEE Robotics and Automation Letters*, vol. 9, no. 1, pp. 390–397, 2023.

[71] A. Goncalves, N. Kuppuswamy, A. Beaulieu, A. Uttamchandani, K. M. Tsui, and A. Alspach, “Punyo-1: Soft tactile-sensing upper-body robot for large object manipulation and physical human interaction,” in *IEEE International Conference on Soft Robotics*, 2022, pp. 844–851.

[72] O. Stasse and L. Righetti, *Whole-Body Manipulation*. Springer, 2020, pp. 1–9.

[73] G. Cheng, E. Dean-Leon, F. Bergner, J. Rogelio Guadarrama Olvera, Q. Leboulet, and P. Mittendorfer, “A comprehensive realization of robot skin: Sensors, sensing, control, and applications,” *Proceedings of the IEEE*, vol. 107, no. 10, pp. 2034–2051, 2019.

[74] A. Gupta, Y. Mao, A. Bhatia, X. Cheng, J. King, Y. Hou, and M. T. Mason, “Extrinsic dexterous manipulation with a direct-drive hand: A case study,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2022, pp. 4660–4667.

[75] K. Hsiao and T. Lozano-Perez, “Imitation learning of whole-body grasps,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006, pp. 5657–5662.

[76] W. Yuan, K. Hang, H. Song, D. Kragic, M. Y. Wang, and J. A. Stork, “Reinforcement learning in topology-based representation for human body movement with whole arm manipulation,” in *IEEE International Conference on Robotics and Automation*, 2019, pp. 2153–2160.

[77] M. Zhang, J. Barreiros, and A. O. Onol, “Plan-guided reinforcement learning for whole-body manipulation,” in *IROS Workshop on Leveraging Models for Contact-Rich Manipulation*, 2023.

[78] M. Florek-Jasińska, T. Wimböck, and C. Ott, “Humanoid compliant whole arm dexterous manipulation: Control design and experiments,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2014, pp. 1616–1621.

[79] M. Murooka, S. Nozawa, Y. Kakiuchi, K. Okada, and M. Inaba, “Whole-body pushing manipulation with contact posture planning of large and heavy object for humanoid robot,” in *IEEE International Conference on Robotics and Automation*, 2015, pp. 5682–5689.

[80] X. Cheng, A. Kumar, and D. Pathak, “Legs as manipulator: Pushing quadruped agility beyond locomotion,” in *IEEE International Conference on Robotics and Automation*, 2023, pp. 5106–5112.

[81] Z. He, K. Lei, Y. Ze, K. Sreenath, Z. Li, and H. Xu, “Learning visual quadrupedal loco-manipulation from demonstrations,” 2024.

[82] P. Arm, M. Mittal, H. Kolvenbach, and M. Hutter, “Pedipulate: Enabling manipulation skills using a quadruped robot’s leg,” in *IEEE International Conference on Robotics and Automation*, 2024, pp. 5717–5723.

[83] H. Ferrolho, V. Ivan, W. Merkt, I. Havoutis, and S. Vijayakumar, “Roloma: robust loco-manipulation for quadruped robots with arms,” *Autonomous Robots*, vol. 47, no. 8, p. 1463–1481, 2023.

[84] J.-P. Sleiman, F. Farshidian, and M. Hutter, “Versatile multicontact planning and control for legged loco-manipulation,” *Science Robotics*, vol. 8, no. 81, 2023.

[85] T. Portela, G. B. Margolis, Y. Ji, and P. Agrawal, “Learning force control for legged manipulation,” in *International Conference on Robotics and Automation*, 2024, pp. 15 366–15 372.

[86] M. Liu, Z. Chen, X. Cheng, Y. Ji, R. Qiu, R. Yang, and X. Wang, “Visual whole-body control for legged loco-manipulation,” *The 8th Conference on Robot Learning*, 2024.

[87] T. Zhang, Z. McCarthy, O. Jow, D. Lee, X. Chen, K. Goldberg, and P. Abbeel, “Deep imitation learning for complex manipulation tasks from virtual reality teleoperation,” in *IEEE International Conference on Robotics and Automation*, 2018, pp. 5628–5635.

[88] D. Rakita, B. Mutlu, M. Gleicher, and L. M. Hiatt, “Shared control-based bimanual robot manipulation,” *Science Robotics*, vol. 4, no. 30, 2019.

[89] M. Seo, S. Han, K. Sim, S. H. Bang, C. Gonzalez, L. Sentis, and Y. Zhu, “Deep imitation learning for humanoid loco-manipulation through human teleoperation,” *IEEE-RAS International Conference on Humanoid Robots*, pp. 1–8, 2023.

[90] S. J. Jorgensen, M. Vedantam, R. Gupta, H. Cappel, and L. Sentis, “Finding locomanipulation plans quickly in the locomotion constrained manifold,” in *IEEE International Conference on Robotics and Automation*, 2020, pp. 6611–6617.

[91] M. Murooka, K. Chappellet, A. Tanguy, M. Benallegue, I. Kumagai, M. Morisawa, F. Kanehiro, and A. Kheddar, “Humanoid locomanipulations pattern generation and stabilization control,” *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 5597–5604, 2021.

[92] J. C. Vaz, N. Kosanovic, and P. Oh, “Art: Avatar robotics telepresence—the future of humanoid material handling loco-manipulation,” *Intelligent Service Robotics*, vol. 17, no. 2, p. 237–250, 2023.

[93] K. Chappellet, M. Murooka, G. Caron, F. Kanehiro, and A. Kheddar, “Humanoid loco-manipulations using combined fast dense 3d tracking and slam with wide-angle depth-images,” *IEEE Transactions on Automation Science and Engineering*, 2023.

[94] J. Vaillant *et al.*, “Multi-contact vertical ladder climbing with an hrp-2 humanoid,” *Autonomous Robots*, vol. 40, no. 3, pp. 561–580, 2016.

[95] P. Ferrari, L. Rossini, F. Ruscelli, A. Laurenzi, G. Oriolo, N. G. Tsagarakis, and E. M. Hoffman, “Multi-contact planning and control for humanoid robots: Design and validation of a complete framework,” *Robotics and Autonomous Systems*, vol. 166, p. 104448, 2023.

[96] A. Kheddar and A. Billard, “A tactile matrix for whole-body humanoid haptic sensing and safe interaction,” in *IEEE International Conference on Robotics and Biomimetics*, 2011, pp. 1433–1438.

[97] A. Jain, M. D. Killpack, A. Edsinger, and C. C. Kemp, “Reaching in clutter with whole-arm tactile sensing,” *The International Journal of Robotics Research*, vol. 32, no. 4, pp. 458–482, 2013.

[98] S. Gao, Y. Dai, and A. Nathan, “Tactile and vision perception for intelligent humanoids,” *Advanced Intelligent Systems*, vol. 4, no. 2, p. 2100074, 2022.

[99] Z. Kappassov, J.-A. Corrales, and V. Perdereau, “Tactile sensing in dexterous robot hands,” *Robotics and Autonomous Systems*, vol. 74, pp. 195–220, 2015.

[100] Y. Al-Handarish, O. M. Omisore, T. Igbe, S. Han, H. Li, W. Du,

J. Zhang, L. Wang *et al.*, “A survey of tactile-sensing systems and their applications in biomedical engineering,” *Advances in Materials Science and Engineering*, vol. 2020, 2020.

[101] Q. Li, C. Schürmann, R. Haschke, and H. J. Ritter, “A control framework for tactile servoing,” in *Robotics: Science and systems*, 2013.

[102] F. Veiga, B. Edin, and J. Peters, “Grip stabilization through independent finger tactile feedback control,” *Sensors*, vol. 20, no. 6, p. 1748, 2020.

[103] X. Song, H. Liu, K. Althoefer, T. Nanayakkara, and L. D. Seneviratne, “Efficient break-away friction ratio and slip prediction based on haptic surface exploration,” *IEEE Transactions on Robotics*, vol. 30, no. 1, pp. 203–219, 2013.

[104] E. Y. P. Mittendorfer and G. Cheng, “Realizing whole-body tactile interactions with a self-organizing, multi-modal artificial skin on a humanoid robot,” *Advanced Robotics*, vol. 29, no. 1, pp. 51–67, 2015.

[105] N. Sommer, M. Li, and A. Billard, “Bimanual compliant tactile exploration for grasping unknown objects,” in *IEEE International Conference on Robotics and Automation*, 2014, pp. 6400–6407.

[106] F. R. Hogan, J. Ballester, S. Dong, and A. Rodriguez, “Tactile dexterity: Manipulation primitives with tactile feedback,” in *IEEE International Conference on Robotics and Automation*, 2020, pp. 8863–8869.

[107] M. Kaboli and G. Cheng, “Robust tactile descriptors for discriminating objects from textural properties via artificial robotic skin,” *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 985–1003, 2018.

[108] S. Tian, F. Ebert, D. Jayaraman, M. Mudigonda, C. Finn, R. Calandra, and S. Levine, “Manipulation by feel: Touch-based control with deep predictive models,” in *IEEE International Conference on Robotics and Automation*, 2019, pp. 818–824.

[109] M. Lambeta *et al.*, “Digit: A novel design for a low-cost compact high-resolution tactile sensor with application to in-hand manipulation,” *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 3838–3845, 2020.

[110] H. Van Hoof, T. Hermans, G. Neumann, and J. Peters, “Learning robot in-hand manipulation with tactile features,” in *IEEE-RAS International Conference on Humanoid Robots*, 2015, pp. 121–127.

[111] A. Melnik, L. Lach, M. Plappert, T. Korthals, R. Haschke, and H. Ritter, “Using tactile sensing to improve the sample efficiency and performance of deep deterministic policy gradients for simulated in-hand manipulation tasks,” *Frontiers in Robotics and AI*, vol. 8, p. 538773, 2021.

[112] T. Lin, Y. Zhang, Q. Li, H. Qi, B. Yi, S. Levine, and J. Malik, “Learning visuotactile skills with two multifingered hands,” *arXiv preprint arXiv:2404.16823*, 2024.

[113] L. Fu, G. Datta, H. Huang, W. C.-H. Panitch, J. Drake, J. Ortiz, M. Mukadam, M. Lambeta, R. Calandra, and K. Goldberg, “A touch, vision, and language dataset for multimodal alignment,” in *Forty-first International Conference on Machine Learning*, 2024.

[114] S. Yu, L. Kelvin, A. Xiao, J. Duan, and H. Soh, “Octopi: Object Property Reasoning with Large Tactile-Language Models,” in *Proceedings of Robotics: Science and Systems*, 2024.

[115] Y. Chebotar, O. Kroemer, and J. Peters, “Learning robot tactile sensing for object manipulation,” in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2014, pp. 3368–3375.

[116] H. Van Hoof, N. Chen, M. Karl, P. van der Smagt, and J. Peters, “Stable reinforcement learning with autoencoders for tactile and visual data,” in *IEEE/RSJ international conference on intelligent robots and systems*, 2016, pp. 3928–3934.

[117] S. Wang, M. Lambeta, P.-W. Chou, and R. Calandra, “Tacto: A fast, flexible, and open-source simulator for high-resolution vision-based tactile sensors,” *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 3930–3937, 2022.

[118] Y. Lin, J. Lloyd, A. Church, and N. F. Lepora, “Tactile gym 2.0: Sim-to-real deep reinforcement learning for comparing low-cost high-resolution robot touch,” *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 10754–10761, 2022.

[119] I. Akinola, J. Xu, J. Carius, D. Fox, and Y. Narang, “Tacs! A library for visuotactile sensor simulation and learning,” *IEEE Transactions on Robotics*, p. 1–17, 2025.

[120] J. R. Guadarrama-Olvera, F. Bergner, E. Dean, and G. Cheng, “Enhancing biped locomotion on unknown terrain using tactile feedback,” in *IEEE-RAS International Conference on Humanoid Robots*, 2018, pp. 1–9.

[121] X. Guo, B. Blaise, J. Molnar, J. Coholich, S. Padte, Y. Zhao, and F. L. Hammond, “Soft foot sensor design and terrain classification for dynamic legged locomotion,” in *IEEE International Conference on Soft Robotics*, 2020, pp. 550–557.

[122] K. Suwanratchatamanee, M. Matsumoto, and S. Hashimoto, “A simple tactile sensing foot for humanoid robot and active ground slope recognition,” in *IEEE International Conference on Mechatronics*, 2009, pp. 1–6.

[123] F. Veiga, R. Akour, and J. Peters, “Hierarchical tactile-based control decomposition of dexterous in-hand manipulation tasks,” *Frontiers in Robotics and AI*, vol. 7, p. 521448, 2020.

[124] P. Maiolino, M. Maggiali, G. Cannata, G. Metta, and L. Natale, “A flexible and robust large scale capacitive tactile system for robots,” *IEEE Sensors Journal*, vol. 13, no. 10, pp. 3910–3917, 2013.

[125] E. Dean-Leon, J. R. Guadarrama-Olvera, F. Bergner, and G. Cheng, “Whole-body active compliance control for humanoid robots with robot skin,” in *IEEE International Conference on Robotics and Automation*, 2019, pp. 5404–5410.

[126] T. Kobayashi, E. Dean-Leon, J. R. Guadarrama-Olvera, F. Bergner, and G. Cheng, “Whole-body multicontact haptic human–humanoid interaction based on leader–follower switching: A robot dance of the “box step”,” *Advanced Intelligent Systems*, vol. 4, no. 2, p. 2100038, 2022.

[127] W. Yuan, S. Dong, and E. H. Adelson, “Gelsight: High-resolution robot tactile sensors for estimating geometry and force,” *Sensors*, vol. 17, no. 12, p. 2762, 2017.

[128] K. Hirai, M. Hirose, Y. Haikawa, and T. Takenaka, “The development of honda humanoid robot,” in *IEEE International Conference on Robotics and Automation*, vol. 2, 1998, pp. 1321–1326.

[129] S. Dafarra *et al.*, “icub3 avatar system: Enabling remote fully immersive embodiment of humanoid robots,” *Science Robotics*, vol. 9, no. 86, 2024.

[130] S. Kuindersma, R. Deits, M. Fallon, A. Valenzuela, H. Dai, F. Pernerter, T. Koolen, P. Marion, and R. Tedrake, “Optimization-based locomotion planning, estimation, and control design for the atlas humanoid robot,” *Autonomous Robots*, vol. 40, no. 3, pp. 429–455, 2016.

[131] Y. Xu, Z. Wang, W. Hao, W. Zhao, W. Lin, B. Jin, and N. Ding, “A flexible multimodal sole sensor for legged robot sensing complex ground information during locomotion,” *Sensors*, vol. 21, no. 16, p. 5359, 2021.

[132] T. Tyler, V. Malhotra, A. Montague, Z. Zhao, F. L. HammondIII, and Y. Zhao, “Integrating reconfigurable foot design, multi-modal contact sensing, and terrain classification for bipedal locomotion,” *IFAC Modeling, Estimation and Control Conference*, vol. 56, no. 3, pp. 523–528, 2023.

[133] F. Nori, S. Traversaro, J. Eljaik, F. Romano, A. Del Prete, and D. Pucci, “icub whole-body control through force regulation on rigid non-coplanar contacts,” *Frontiers in Robotics and AI*, vol. 2, p. 6, 2015.

[134] A. Kheddar *et al.*, “Humanoid robots in aircraft manufacturing: The airbus use cases,” *IEEE Robotics Automation Magazine*, vol. 26, no. 4, pp. 30–45, 2019.

[135] C. Y. Wong, S. Samadi, W. Suleiman, and A. Kheddar, “Touch semantics for intuitive physical manipulation of humanoids,” *IEEE Transactions on Human-Machine Systems*, vol. 52, no. 6, pp. 1111–1121, 2022.

[136] K. Harada, S. Hattori, H. Hirukawa, M. Morisawa, S. Kajita, and E. Yoshida, “Two-stage time-parametrized gait planning for humanoid robots,” *IEEE/ASME Transactions on Mechatronics*, vol. 15, no. 5, pp. 694–703, 2010.

[137] S. Armleder, E. Dean-Leon, F. Bergner, J. R. Guadarrama Olvera, and G. Cheng, “Tactile-based negotiation of unknown objects during navigation in unstructured environments with movable obstacles,” *Advanced Intelligent Systems*, vol. 6, no. 3, p. 2300621, 2024.

[138] Y. Ohmura and Y. Kuniyoshi, “Humanoid robot which can lift a 30kg box by whole body contact and tactile feedback,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2007, pp. 1136–1141.

[139] L. Natale and G. Cannata, “Tactile sensing,” in *Humanoid Robotics: A Reference*. Springer, 2017, pp. 110–1.

[140] S. Lengagne, J. Vaillant, E. Yoshida, and A. Kheddar, “Generation of whole-body optimal dynamic multi-contact motions,” *The International Journal of Robotics Research*, vol. 32, no. 9–10, pp. 1104–1119, 2013.

[141] B. Ponton, A. Herzog, S. Schaal, and L. Righetti, “A convex model of humanoid momentum dynamics for multi-contact motion generation,” in *IEEE-RAS International Conference on Humanoid Robots*, 2016, pp. 842–849.

[142] A. Bilotnikova, S. Courtois, and A. Kheddar, “Multi-contact planning on humans for physical assistance by humanoid,” *IEEE Robotics and Automation Letters*, vol. 5, no. 1, pp. 135–142, 2020.

[143] K. Bouyarmane, S. Caron, A. Escande, and A. Kheddar, “Multi-contact motion planning and control,” in *Humanoid Robotics: A Reference*.

Springer, 2019, pp. 1763–1804.

[144] M. Posa, C. Cantu, and R. Tedrake, “A direct method for trajectory optimization of rigid bodies through contact,” *The International Journal of Robotics Research*, vol. 33, no. 1, pp. 69–81, 2014.

[145] I. Mordatch, E. Todorov, and Z. Popović, “Discovery of complex behaviors through contact-invariant optimization,” *ACM Transactions on Graphics*, vol. 31, no. 4, pp. 1–8, 2012.

[146] H. Zhu, A. Meduri, and L. Righetti, “Efficient object manipulation planning with monte carlo tree search,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2023, pp. 10 628–10 635.

[147] K. Hauser, T. Bretl, and J.-C. Latombe, “Non-gaited humanoid locomotion planning,” in *IEEE-RAS International Conference on Humanoid Robots*, 2005, pp. 7–12.

[148] P. Ferrari, N. Scianca, L. Lanari, and G. Oriolo, “An integrated motion planner/controller for humanoid robots on uneven ground,” in *European Control Conference*, 2019, pp. 1598–1603.

[149] L. Amatucci, J.-H. Kim, J. Hwangbo, and H.-W. Park, “Monte carlo tree search gait planner for non-gaited legged system control,” in *IEEE International Conference on Robotics and Automation*, 2022, pp. 4701–4707.

[150] I. Taouil, L. Amatucci, M. Khadiv, A. Dai, V. Barasuol, G. Turrisi, and C. Semini, “Non-gaited legged locomotion with monte-carlo tree search and supervised learning,” *IEEE Robotics and Automation Letters*, pp. 1–8, 2024.

[151] A. K. C. Ravi, V. Dhédin, A. Jordana, H. Zhu, A. Meduri, L. Righetti, B. Schölkopf, and M. Khadiv, “Efficient search and learning for agile locomotion on stepping stones,” *arXiv preprint arXiv:2403.03639*, 2024.

[152] M. Murooka, I. Kumagai, M. Morisawa, F. Kanehiro, and A. Kheddar, “Humanoid loco-manipulation planning based on graph search and reachability maps,” *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 1840–1847, 2021.

[153] A. Escande, A. Kheddar, and S. Miossec, “Planning contact points for humanoid robots,” *Robotics and Autonomous Systems*, vol. 61, no. 5, pp. 428–442, 2013.

[154] L. Janson, E. Schmerling, and M. Pavone, “Monte carlo motion planning for robot trajectory optimization under uncertainty,” in *Robotics Research: Volume 2*. Springer, 2017, pp. 343–361.

[155] L. Li, Y. Miao, A. H. Qureshi, and M. C. Yip, “Mpc-mpnet: Model-predictive motion planning networks for fast, near-optimal planning under kinodynamic constraints,” *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 4496–4503, 2021.

[156] S. Brossette, A. Escande, and A. Kheddar, “Multicontact postures computation on manifolds,” *IEEE Transactions on Robotics*, vol. 34, no. 5, pp. 1252–1265, 2018.

[157] F. Abi-Farraj, B. Henze, C. Ott, P. R. Giordano, and M. A. Roa, “Torque-based balancing for a humanoid robot performing high-force interaction tasks,” *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 2023–2030, 2019.

[158] Q. Rouxel, K. Yuan, R. Wen, and Z. Li, “Multicontact motion retargeting using whole-body optimization of full kinematics and sequential force equilibrium,” *IEEE/ASME Transactions on Mechatronics*, vol. 27, no. 5, pp. 4188–4198, 2022.

[159] E. Farnioli, M. Gabiccini, and A. Bicchi, “Optimal contact force distribution for compliant humanoid robots in whole-body loco-manipulation tasks,” in *IEEE International Conference on Robotics and Automation*, 2015, pp. 5675–5681.

[160] J. Li and Q. Nguyen, “Kinodynamic pose optimization for humanoid loco-manipulation,” in *IEEE-RAS International Conference on Humanoid Robots*, 2023, pp. 1–8.

[161] I. Chatzinikolaidis, Y. You, and Z. Li, “Contact-implicit trajectory optimization using an analytically solvable contact model for locomotion on variable ground,” *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 6357–6364, 2020.

[162] Z. Manchester and S. Kuindersma, “Variational contact-implicit trajectory optimization,” in *Robotics Research: The International Symposium*. Springer, 2020, pp. 985–1000.

[163] M. Posa, S. Kuindersma, and R. Tedrake, “Optimization and stabilization of trajectories for constrained dynamical systems,” in *IEEE International Conference on Robotics and Automation*, 2016, pp. 1366–1373.

[164] C. Mastalli, I. Havoutis, M. Focchi, D. G. Caldwell, and C. Semini, “Hierarchical planning of dynamic movements without scheduled contact sequences,” in *IEEE International Conference on Robotics and Automation*, 2016, pp. 4636–4641.

[165] H. Chen, X. Chen, C. Dong, Z. Yu, and Q. Huang, “Online running pattern generation for humanoid robot with direct collocation of reference-tracking dynamics,” *IEEE/ASME Transactions on Mechatronics*, pp. 2091–2102, 2023.

[166] Y. Tassa, T. Erez, and E. Todorov, “Synthesis and stabilization of complex behaviors through online trajectory optimization,” in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012, pp. 4906–4913.

[167] J. Carius, R. Ranftl, V. Koltun, and M. Hutter, “Trajectory optimization with implicit hard contacts,” *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3316–3323, 2018.

[168] S. Le Cleac'h, T. A. Howell, S. Yang, C.-Y. Lee, J. Zhang, A. Bishop, M. Schwager, and Z. Manchester, “Fast contact-implicit model predictive control,” *IEEE Transactions on Robotics*, vol. 40, pp. 1617–1629, 2024.

[169] N. J. Kong, C. Li, G. Council, and A. M. Johnson, “Hybrid iLQR model predictive control for contact implicit stabilization on legged robots,” *IEEE Transactions on Robotics*, vol. 39, no. 6, pp. 4712–4727, 2023.

[170] A. Aydinoglu, A. Wei, W.-C. Huang, and M. Posa, “Consensus complementarity control for multi-contact mpc,” *IEEE Transactions on Robotics*, pp. 1–18, 2024.

[171] V. Kurtz, A. Castro, A. Ö. Önlö, and H. Lin, “Inverse dynamics trajectory optimization for contact-implicit model predictive control,” *arXiv preprint arXiv:2309.01813*, 2023.

[172] S. A. Esteban, V. Kurtz, A. B. Ghansah, and A. D. Ames, “Reduced-order model guided contact-implicit model predictive control for humanoid locomotion,” *arXiv preprint arXiv:2502.15630*, 2025.

[173] G. B. Margolis, T. Chen, K. Paigwar, X. Fu, D. Kim, S. b. Kim, and P. Agrawal, “Learning to jump from pixels,” in *Proceedings of the Conference on Robot Learning*, vol. 164, 2022, pp. 1025–1034.

[174] V. Tsounis, M. Alge, J. Lee, F. Farshidian, and M. Hutter, “Deepgait: Planning and control of quadrupedal gaits using deep reinforcement learning,” *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 3699–3706, 2020.

[175] Y.-C. Lin, B. Ponton, L. Righetti, and D. Berenson, “Efficient humanoid contact planning using learned centroidal dynamics prediction,” in *IEEE International Conference on Robotics and Automation*, 2019, pp. 5280–5286.

[176] S. Xu, Z. Li, Y.-X. Wang, and L.-Y. Gui, “Interdiff: Generating 3d human-object interactions with physics-informed diffusion,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 14 928–14 940.

[177] J. Li, J. Wu, and C. K. Liu, “Object motion guided human motion synthesis,” *ACM Transactions on Graphics*, vol. 42, no. 6, pp. 1–11, 2023.

[178] S. Bahl, R. Mendonca, L. Chen, U. Jain, and D. Pathak, “Affordances from human videos as a versatile representation for robotics,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 13 778–13 790.

[179] M. Chignoli, D. Kim, E. Stanger-Jones, and S. Kim, “The mit humanoid robot: Design, motion planning, and control for acrobatic behaviors,” in *IEEE-RAS International Conference on Humanoid Robots*, 2021, pp. 1–8.

[180] G. Romualdi, S. Dafarra, G. L’Erario, I. Sorrentino, S. Traversaro, and D. Pucci, “Online non-linear centroidal mpc for humanoid robot locomotion with step adjustment,” in *International Conference on Robotics and Automation*, 2022, pp. 10 412–10 419.

[181] M. Elbaid, G. Romualdi, G. Nava, L. Rapetti, H. A. O. Mohamed, and D. Pucci, “Online non-linear centroidal mpc for humanoid robots payload carrying with contact-stable force parametrization,” in *IEEE International Conference on Robotics and Automation*, 2023, pp. 12 233–12 239.

[182] J. Li and Q. Nguyen, “Force-and-moment-based model predictive control for achieving highly dynamic locomotion on bipedal robots,” in *IEEE Conference on Decision and Control*, 2021, pp. 1024–1030.

[183] B. Henze, C. Ott, and M. A. Roa, “Posture and balance control for humanoid robots in multi-contact scenarios based on model predictive control,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2014, pp. 3253–3258.

[184] S. H. Bang, C. Gonzalez, G. Moore, D. H. Kang, M. Seo, R. Gupta, and L. Sentis, “Rpc: A modular framework for robot planning, control, and deployment,” in *IEEE/SICE International Symposium on System Integration*, 2025, pp. 1142–1148.

[185] L. Penco, N. Scianca, V. Modugno, L. Lanari, G. Oriolo, and S. Ivaldi, “A multimode teleoperation framework for humanoid loco-manipulation: An application for the icub robot,” *IEEE Robotics Automation Magazine*, vol. 26, no. 4, pp. 73–82, 2019.

[186] H. Dai, A. Valenzuela, and R. Tedrake, “Whole-body motion planning

with centroidal dynamics and full kinematics,” in *2014 IEEE-RAS International Conference on Humanoid Robots*, 2014, pp. 295–302.

[187] M. P. Polverini, A. Laurenzi, E. M. Hoffman, F. Ruscelli, and N. G. Tsagarakis, “Multi-contact heavy object pushing with a centaur-type humanoid robot: Planning and control for a real demonstrator,” *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 859–866, 2020.

[188] E. Dantec *et al.*, “Whole body model predictive control with a memory of motion: Experiments on a torque-controlled talos,” in *International Conference on Robotics and Automation*, 2021, pp. 8202–8208.

[189] H. Audren, J. Vaillant, A. Kheddar, A. Escande, K. Kaneko, and E. Yoshida, “Model preview control in multi-contact motion-application to a humanoid robot,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2014, pp. 4030–4035.

[190] D. E. Orin, A. Goswami, and S.-H. Lee, “Centroidal dynamics of a humanoid robot,” *Autonomous robots*, vol. 35, pp. 161–176, 2013.

[191] H. Li, R. J. Frei, and P. M. Wensing, “Model hierarchy predictive control of robotic systems,” *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 3373–3380, 2021.

[192] J. Wang, S. Kim, S. Vijayakumar, and S. Tonneau, “Multi-fidelity receding horizon planning for multi-contact locomotion,” in *IEEE-RAS International Conference on Humanoid Robots*, 2021, pp. 53–60.

[193] J. Li, Z. Le, J. Ma, and Q. Nguyen, “Adapting gait frequency for posture-regulating humanoid push-recovery via hierarchical model predictive control,” *arXiv preprint arXiv:2409.14342*, 2024.

[194] R. Budhiraja, J. Carpentier, and N. Mansard, “Dynamics consensus between centroidal and whole-body models for locomotion of legged robots,” in *IEEE International Conference on Robotics and Automation*, 2019, pp. 6727–6733.

[195] A. Herzog, S. Schaal, and L. Righetti, “Structured contact force optimization for kino-dynamic motion generation,” in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2016, pp. 2703–2710.

[196] D. Mayne, “A second-order gradient method for determining optimal trajectories of non-linear discrete-time systems,” *International Journal of Control.*, vol. 3, no. 1, pp. 85–95, 1966.

[197] Y. Tassa, T. Erez, and E. Todorov, “Synthesis and stabilization of complex behaviors through online trajectory optimization,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012, pp. 4906–4913.

[198] L. B. Rall, *Automatic differentiation: Techniques and applications*. Springer, 1981.

[199] C. Rösmann, M. Krämer, A. Makarow, F. Hoffmann, and T. Bertram, “Exploiting sparse structures in nonlinear model predictive control with hypergraphs,” in *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, 2018, pp. 1332–1337.

[200] L. Vanroye, A. Sathy, J. De Schutter, and W. Decré, “Fatrop: A fast constrained optimal control problem solver for robot trajectory optimization and control,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2023, pp. 10 036–10 043.

[201] L. Callens, J. Gillis, W. Decré, and J. Swevers, “Adaptivenlp: a framework for efficient online adaptability in nlp structures for optimal control problems,” in *European Control Conference*, 2024, pp. 1358–1365.

[202] R. B. Wilson, “A simplicial method for convex programming,” *Ph.D. Thesis*, 1963.

[203] A. Zhakatayev, B. Rakhim, O. Adiyatov, A. Baimyshev, and H. A. Varol, “Successive linearization based model predictive control of variable stiffness actuated robots,” in *IEEE International Conference on Advanced Intelligent Mechatronics*, 2017, pp. 1774–1779.

[204] J. L. Jerez, E. C. Kerrigan, and G. A. Constantinides, “A condensed and sparse qp formulation for predictive control,” in *2011 50th IEEE Conference on Decision and Control and European Control Conference*. IEEE, 2011, pp. 5217–5222.

[205] A. Bishop, J. Zhang, S. Gurumurthy, K. Tracy, and Z. Manchester, “Relu-qp: A gpu-accelerated quadratic programming solver for model-predictive control,” in *IEEE International Conference on Robotics and Automation*, 2024, pp. 13 285–13 292.

[206] X. Da and J. Grizzle, “Combining trajectory optimization, supervised machine learning, and model structure for mitigating the curse of dimensionality in the control of bipedal robots,” *The International Journal of Robotics Research*, vol. 38, no. 9, pp. 1063–1097, 2019.

[207] M. Diehl, H. G. Bock, H. Diedam, and P.-B. Wieber, “Fast direct multiple shooting algorithms for optimal robot control,” *Fast motions in biomechanics and robotics: optimization and feedback control*, pp. 65–93, 2006.

[208] J. T. Betts, *Practical Methods for Optimal Control and Estimation Using Nonlinear Programming, Second Edition*, 2010.

[209] E. Dantec, M. Naveau, P. Fernbach, N. Villa, G. Saurel, O. Stasse, M. Taix, and N. Mansard, “Whole-body model predictive control for biped locomotion on a torque-controlled humanoid robot,” in *IEEE-RAS International Conference on Humanoid Robots*, 2022, pp. 638–644.

[210] Y. Ding, A. Pandala, and H.-W. Park, “Real-time model predictive control for versatile dynamic motions in quadrupedal robots,” in *International Conference on Robotics and Automation*, 2019, pp. 8484–8490.

[211] G. Williams, P. Drews, B. Goldfain, J. M. Rehg, and E. A. Theodorou, “Aggressive driving with model predictive path integral control,” in *IEEE International Conference on Robotics and Automation*, 2016, pp. 1433–1440.

[212] J. Carius, R. Ranftl, F. Farshidian, and M. Hutter, “Constrained stochastic optimal control with learned importance sampling: A path integral approach,” *The International Journal of Robotics Research*, vol. 41, no. 2, pp. 189–209, 2022.

[213] T. Howell, N. Gileadi, S. Tunyasuvunakool, K. Zakka, T. Erez, and Y. Tassa, “Predictive sampling: Real-time behaviour synthesis with mujoco,” *arXiv preprint arXiv:2212.00541*, 2022.

[214] E. Todorov, T. Erez, and Y. Tassa, “Mujoco: A physics engine for model-based control,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012, pp. 5026–5033.

[215] M. Mittal *et al.*, “Orbit: A unified simulation framework for interactive robot learning environments,” *IEEE Robotics and Automation Letters*, vol. 8, no. 6, p. 3740–3747, 2023.

[216] G. Turrisi, V. Modugno, L. Amatucci, D. Kanoulas, and C. Semini, “On the benefits of gpu sample-based stochastic predictive controllers for legged locomotion,” *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 13 757–13 764, 2024.

[217] A. Adu-Bredu, G. Gibson, and J. Grizzle, “Exploring kinodynamic fabrics for reactive whole-body control of underactuated humanoid robots,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2023, pp. 10 397–10 404.

[218] J. Carpentier and N. Mansard, “Multicontact locomotion of legged robots,” *IEEE Transactions on Robotics*, vol. 34, no. 6, pp. 1441–1460, 2018.

[219] K. Otani and K. Bouyarmane, “Adaptive whole-body manipulation in human-to-humanoid multi-contact motion retargeting,” in *IEEE-RAS International Conference on Humanoid Robotics*, 2017, pp. 446–453.

[220] K. Harada, S. Kajita, F. Kanehiro, K. Fujiwara, K. Kaneko, K. Yokoi, and H. Hirukawa, “Real-time planning of humanoid robot’s gait for force-controlled manipulation,” *IEEE/ASME Transactions on Mechatronics*, vol. 12, no. 1, pp. 53–62, 2007.

[221] J. Foster, S. McCrary, C. DeBuys, S. Bertrand, and R. J. Griffin, “Physically consistent online inertial adaptation for humanoid locomotion manipulation,” *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 11 278–11 285, 2024.

[222] S. Nozawa, R. Ueda, Y. Kakiuchi, K. Okada, and M. Inaba, “A full-body motion control method for a humanoid robot based on on-line estimation of the operational force of an object with an unknown weight,” in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2010, pp. 2684–2691.

[223] S. H. Bang, J. Lee, C. Gonzalez, and L. Sentis, “Variable inertia model predictive control for fast bipedal maneuvers,” in *IEEE Conference on Decision and Control*, 2024, pp. 4334–4341.

[224] Y. Gao, V. Paredes, Y. Gong, Z. He, A. Hereid, and Y. Gu, “Time-varying foot-placement control for underactuated humanoid walking on swaying rigid surfaces,” *arXiv preprint arXiv:2409.08371*, 2024.

[225] H. Lefèvre, T. Chaki, T. Kawakami, A. Tanguy, T. Yoshiike, and A. Kheddar, “Humanoid-human sit-to-stand-to-sit assistance,” *IEEE Robotics and Automation Letters*, vol. 10, 2025.

[226] J. Zhu, B. Navarro, P. Fraisse, A. Crosnier, and A. Cherubini, “Dual-arm robotic manipulation of flexible cables,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2018, pp. 479–484.

[227] M. Saha and P. Isto, “Manipulation planning for deformable linear objects,” *IEEE Transactions on Robotics*, vol. 23, no. 6, pp. 1141–1150, 2007.

[228] Y. Qin, A. Escande, F. Kanehiro, and E. Yoshida, “Dual-arm mobile manipulation planning of a long deformable object in industrial installation,” *IEEE Robotics and Automation Letters*, vol. 8, no. 5, pp. 3039–3046, 2023.

[229] K. Zhang, B. Li, K. Hauser, and Y. Li, “Adaptigraph: Material-adaptive graph-based neural dynamics for robotic manipulation,” *arXiv preprint arXiv:2407.07889*, 2024.

[230] J. Li and Q. Nguyen, “Multi-contact mpc for dynamic locomotion manipulation on humanoid robots,” in *American Control Conference*, 2023, pp. 1215–1220.

[231] S.-C. Hsu, X. Xu, and A. D. Ames, "Control barrier function based quadratic programs with application to bipedal robotic walking," in *American Control Conference*, 2015, pp. 4542–4548.

[232] O. Khatib, "A unified approach for motion and force control of robot manipulators: The operational space formulation," *IEEE Journal on Robotics and Automation*, vol. 3, no. 1, pp. 43–53, 1987.

[233] M. A. Hopkins, D. W. Hong, and A. Leonessa, "Compliant locomotion using whole-body control and divergent component of motion tracking," in *IEEE International Conference on Robotics and Automation*, 2015, pp. 5726–5733.

[234] T. Koolen, S. Bertrand, G. Thomas, T. De Boer, T. Wu, J. Smith, J. Englsberger, and J. Pratt, "Design of a momentum-based control framework and application to the humanoid robot atlas," *International Journal of Humanoid Robotics*, vol. 13, no. 01, p. 1650007, 2016.

[235] D. Marew, M. Lvovsky, S. Yu, S. Sessions, and D. Kim, "Integration of riemannian motion policy with whole-body control for collision-free legged locomotion," in *IEEE-RAS International Conference on Humanoid Robots*, 2023, pp. 1–8.

[236] S. Yu, N. Perera, D. Marew, and D. Kim, "Learning generic and dynamic locomotion of humanoids across discrete terrains," in *International Conference on Humanoid Robots*, 2024, pp. 1048–1055.

[237] M. Mistry, J. Nakanishi, G. Cheng, and S. Schaal, "Inverse kinematics with floating base and constraints for full body humanoid robot control," in *IEEE-RAS International Conference on Humanoid Robots*, 2008, pp. 22–27.

[238] K. Darvish, L. Penco, J. Ramos, R. Cisneros, J. Pratt, E. Yoshida, S. Ivaldi, and D. Pucci, "Teleoperation of humanoid robots: A survey," *IEEE Transactions on Robotics*, vol. 39, no. 3, pp. 1706–1727, 2023.

[239] P. Marion *et al.*, "Director: A user interface designed for robot operation with shared autonomy," *The DARPA Robotics Challenge Finals: Humanoid Robots To The Rescue*, pp. 237–270, 2018.

[240] L. Penco, B. Clement, V. Modugno, E. Mingo Hoffman, G. Nava, D. Pucci, N. G. Tsagarakis, J. B. Mouret, and S. Ivaldi, "Robust real-time whole-body motion retargeting from human to humanoid," in *IEEE-RAS International Conference on Humanoid Robots*, 2018, pp. 425–432.

[241] J. Oh, O. Sim, B. Cho, K. Lee, and J.-H. Oh, "Online delayed reference generation for a humanoid imitating human walking motion," *IEEE/ASME Transactions on Mechatronics*, vol. 26, no. 1, pp. 102–112, 2021.

[242] S. J. Jorgensen, M. Wonsick, M. Paterson, A. Watson, I. Chase, and J. S. Mehling, "Cockpit interface for locomotion and manipulation control of the nasa valkyrie humanoid in virtual reality (vr)," White Paper, 2022.

[243] D. Wallace, Y. H. He, J. C. Vaz, L. Georgescu, and P. Y. Oh, "Multimodal teleoperation of heterogeneous robots within a construction environment," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2020, pp. 2698–2705.

[244] F. Abi-Farraj, B. Henze, A. Werner, M. Panzirsch, C. Ott, and M. A. Roa, "Humanoid teleoperation using task-relevant haptic feedback," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2018, pp. 5010–5017.

[245] J. Ramos and S. Kim, "Humanoid dynamic synchronization through whole-body bilateral feedback teleoperation," *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 953–965, 2018.

[246] G. Colin, Y. Sim, and J. Ramos, "Bipedal robot walking control using human whole-body dynamic teleoperation," in *IEEE International Conference on Robotics and Automation*, 2023, pp. 12 191–12 197.

[247] L. Sentis, "Synthesis and control of whole-body behaviors in humanoid systems," Ph.D. dissertation, Stanford University, 2007.

[248] M. Mistry, J. Buchli, and S. Schaal, "Inverse dynamics control of floating base systems using orthogonal decomposition," in *International Conference on Robotics and Automation*, 2010, pp. 3406–3412.

[249] L. Righetti, J. Buchli, M. Mistry, and S. Schaal, "Inverse dynamics control of floating-base robots with external constraints: A unified view," in *2011 IEEE International Conference on Robotics and Automation*, 2011, pp. 1085–1090.

[250] J. Nakanishi, R. Cory, M. Mistry, J. Peters, and S. Schaal, "Operational space control: A theoretical and empirical comparison," *The International Journal of Robotics Research*, vol. 27, no. 6, pp. 737–757, 2008.

[251] N. Mansard, O. Khatib, and A. Kheddar, "A unified approach to integrate unilateral constraints in the stack of tasks," *IEEE Transactions on Robotics*, vol. 25, no. 3, pp. 670–685, 2009.

[252] F. Keith, P.-B. Wieber, N. Mansard, and A. Kheddar, "Analysis of the discontinuities in prioritized tasks-space control under discreet task scheduling operations," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2011, pp. 3887–3892.

[253] A. Rocchi, E. M. Hoffman, D. G. Caldwell, and N. G. Tsagarakis, "OpenSoT: A whole-body control library for the compliant humanoid robot COMAN," in *IEEE International Conference on Robotics and Automation*, 2015, pp. 6248–6253.

[254] A. Escande, N. Mansard, and P.-B. Wieber, "Hierarchical quadratic programming: Fast online humanoid-robot motion generation," *The International Journal of Robotics Research*, vol. 33, no. 7, pp. 1006–1028, 2014.

[255] K. Bouyarmane and A. Kheddar, "Using a multi-objective controller to synthesize simulated humanoid robot motion with changing contact configurations," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2011, pp. 4414–4419.

[256] M. de Las, I. Mordatch, and A. Hertzmann, "Feature-based locomotion controllers," *ACM Transactions On Graphics*, vol. 29, no. 4, pp. 131:1–131:10, 2010.

[257] K. Pfeiffer, A. Escande, and A. Kheddar, "Singularity Resolution in Equality and Inequality Constrained Hierarchical Task-Space Control by Adaptive Nonlinear Least Squares," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3630–3637, 2018.

[258] K. Bouyarmane and A. Kheddar, "On weight-prioritized multitask control of humanoid robots," *IEEE Transactions on Automatic Control*, vol. 63, no. 6, pp. 1632–1647, 2018.

[259] M. Djeha, P. Gergondet, and A. Kheddar, "Robust task-space quadratic programming for kinematic-controlled robots," *IEEE Transactions on Robotics*, vol. 39, no. 5, pp. 3857–3874, 2023.

[260] S. Feng, E. Whitman, X. Jinjilefu, and C. G. Atkeson, "Optimization-based full body control for the darpa robotics challenge," *Journal of Field Robotics*, vol. 32, no. 2, pp. 293–312, 2015.

[261] H. Hirukawa, S. Hattori, K. Harada, S. Kajita, K. Kaneko, F. Kanehiro, K. Fujiwara, and M. Morisawa, "A universal stability criterion of the foot contact of legged robots - adios zmp," in *IEEE International Conference on Robotics and Automation*, 2006, pp. 1976–1983.

[262] L. Righetti, J. Buchli, M. Mistry, M. Kalakrishnan, and S. Schaal, "Optimal distribution of contact forces with inverse-dynamics control," *The International Journal of Robotics Research*, vol. 32, no. 3, pp. 280–298, 2013.

[263] P.-B. Wieber, *Holonomy and Nonholonomy in the Dynamics of Articulated Motion*. Springer Berlin Heidelberg, p. 411–425.

[264] A. Laurenzi, E. Mingo Hoffman, M. Parigi Polverini, and N. G. Tsagarakis, "Balancing control through post-optimization of contact forces," in *IEEE-RAS International Conference on Humanoid Robots*, 2018, pp. 320–326.

[265] K. Bouyarmane, K. Chappellet, J. Vaillant, and A. Kheddar, "Quadratic programming for multirobot and task-space force control," *IEEE Transactions on Robotics*, vol. 35, no. 1, pp. 64–77, 2019.

[266] M. Ji, X. Peng, F. Liu, J. Li, G. Yang, X. Cheng, and X. Wang, "Exbody2: Advanced expressive humanoid whole-body control," *arXiv preprint arXiv:2412.13196*, 2024.

[267] T. He, Z. Luo, X. He, W. Xiao, C. Zhang, W. Zhang, K. M. Kitani, C. Liu, and G. Shi, "Omnih2o: Universal and dexterous human-to-humanoid whole-body teleoperation and learning," in *Annual Conference on Robot Learning*, 2024.

[268] Z. Fu, Q. Zhao, Q. Wu, G. Wetzstein, and C. Finn, "Humanplus: Humanoid shadowing and imitation from humans," in *Annual Conference on Robot Learning*, 2024.

[269] Y. Wang, N. Dehio, A. Tanguy, and A. Kheddar, "Impact-aware task-space quadratic-programming control," *The International Journal of Robotics Research*, vol. 42, no. 14, pp. 1265–1282, 2023.

[270] A. Patel, S. L. Shield, S. Kazi, A. M. Johnson, and L. T. Biegler, "Contact-implicit trajectory optimization using orthogonal collocation," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 2242–2249, 2019.

[271] A. Aydinoglu and M. Posa, "Real-time multi-contact model predictive control via admmm," in *International Conference on Robotics and Automation*, 2022, pp. 3414–3421.

[272] Sharanya Puthige Venkatesh, "Approximating global mpc for contact rich manipulation using local feedback," Master's thesis, University of Pennsylvania, 2024.

[273] B. Plancher and S. Kuindersma, "A performance analysis of parallel differential dynamic programming on a gpu," in *the 13th Workshop on the Algorithmic Foundations of Robotics*. Springer, 2020, pp. 656–672.

[274] E. Adabag, M. Atal, W. Gerard, and B. Plancher, "Mpcgpu: Real-time nonlinear model predictive control through preconditioned conjugate gradient on the gpu," in *IEEE International Conference on Robotics and Automation*, 2024, pp. 9787–9794.

[275] N. J. Kong, J. J. Payne, J. Zhu, and A. M. Johnson, "Saltation matrices: The essential tool for linearizing hybrid dynamical systems,"

*Proceedings of the IEEE*, 2024.

[276] H. J. T. Suh, T. Pang, and R. Tedrake, “Bundled gradients through contact via randomized smoothing,” *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 4000–4007, 2022.

[277] B. Hammoud, A. Jordana, and L. Righetti, “irisc: Iterative risk sensitive control for nonlinear systems with imperfect observations,” in *2022 American Control Conference (ACC)*. IEEE, 2022, pp. 3550–3557.

[278] Q. Chen, J. Li, S. Cheng, N. Hovakimyan, and Q. Nguyen, “Auto-tuning bipedal locomotion mpc with grfm-net for efficient sim-to-real transfer,” *arXiv preprint arXiv:2409.15710*, 2024.

[279] C. Sartore, M. Rando, G. Romualdi, C. Molinari, L. Rosasco, and D. Pucci, “Automatic gain tuning for humanoid robots walking architectures using gradient-free optimization techniques,” in *IEEE-RAS International Conference on Humanoid Robots*, 2024, pp. 996–1003.

[280] R. Jiang, B. He, Z. Wang, X. Cheng, H. Sang, and Y. Zhou, “Robot skill learning and the data dilemma it faces: a systematic review,” *Robotic Intelligence and Automation*, vol. 44, no. 2, pp. 270–286, 2024.

[281] C. Zhang, W. Xiao, T. He, and G. Shi, “Wococo: Learning whole-body humanoid control with sequential contacts,” in *Annual Conference on Robot Learning*, 2024.

[282] T. Haarnoja *et al.*, “Learning agile soccer skills for a bipedal robot with deep reinforcement learning,” *Science Robotics*, vol. 9, no. 89, 2024.

[283] X. Huang, Y. Chi, R. Wang, Z. Li, X. B. Peng, S. Shao, B. Nikolic, and K. Sreenath, “Diffuseloco: Real-time legged locomotion control with diffusion from offline datasets,” in *Annual Conference on Robot Learning*, 2024.

[284] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, “Proximal policy optimization algorithms,” *arXiv preprint:1707.06347*, 2017.

[285] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, “Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor,” in *International Conference on Machine Learning*, vol. 80, 2018, pp. 1856–1865.

[286] C. Sferrazza, D.-M. Huang, X. Lin, Y. Lee, and P. Abbeel, “Humanoid-Bench: Simulated Humanoid Benchmark for Whole-Body Locomotion and Manipulation,” in *Robotics: Science and Systems*, 2024.

[287] V. Mnih, A. P. Badia, M. Mirza, A. Graves, T. P. Lillicrap, T. Harley, D. Silver, and K. Kavukcuoglu, “Asynchronous methods for deep reinforcement learning,” in *International Conference on Machine Learning*, 2016.

[288] S. Fujimoto, H. Hoof, and D. Meger, “Addressing function approximation error in actor-critic methods,” in *International Conference on Machine Learning*, 2018, pp. 1582–1591.

[289] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. M. O. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, “Continuous control with deep reinforcement learning,” *CoRR*, vol. abs/1509.02971, 2015.

[290] Z. Xie, P. Clary, J. Dao, P. Morais, J. Hurst, and M. van de Panne, “Learning locomotion skills for cassie: Iterative design and sim-to-real,” in *the Conference on Robot Learning*, vol. 100, 2020, pp. 317–329.

[291] C. Schwarke, V. Klemm, M. v. d. Boon, M. Bjelonic, and M. Hutter, “Curiosity-driven learning of joint locomotion and manipulation tasks,” in *Conference on Robot Learning*, vol. 229, 2023, pp. 2594–2610.

[292] Y. Kim, H. Oh, J. Lee, J. Choi, G. Ji, M. Jung, D. Youm, and J. Hwangbo, “Not only rewards but also constraints: Applications on legged robot locomotion,” *IEEE Transactions on Robotics*, vol. 40, pp. 2984–3003, 2024.

[293] J. Lee, J. Hwangbo, L. Wellhausen, V. Koltun, and M. Hutter, “Learning quadrupedal locomotion over challenging terrain,” *Science Robotics*, vol. 5, no. 47, 2020.

[294] J. Tan, Z. Xie, B. Boots, and C. K. Liu, “Simulation-based design of dynamic controllers for humanoid balancing,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2016, pp. 2729–2736.

[295] Y. Li, J. Li, W. Fu, and Y. Wu, “Learning agile bipedal motions on a quadrupedal robot,” in *IEEE International Conference on Robotics and Automation*, 2024, pp. 9735–9742.

[296] T. He *et al.*, “Asap: Aligning simulation and real-world physics for learning agile humanoid whole-body skills,” *arXiv preprint arXiv:2502.01143*, 2025.

[297] K.-C. Hsu, A. Z. Ren, D. P. Nguyen, A. Majumdar, and J. F. Fisac, “Sim-to-lab-to-real: Safe reinforcement learning with shielding and generalization guarantees,” *Artificial Intelligence*, p. 103811, 2022.

[298] L. Brunke, M. Greeff, A. W. Hall, Z. Yuan, S. Zhou, J. Panerati, and A. P. Schoellig, “Safe learning in robotics: From learning-based control to safe reinforcement learning,” *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 5, no. Volume 5, 2022, pp. 411–444, 2022.

[299] J. Ibarz, J. Tan, C. Finn, M. Kalakrishnan, P. Pastor, and S. Levine, “How to train your robot with deep reinforcement learning: lessons we have learned,” *The International Journal of Robotics Research*, vol. 40, no. 4-5, pp. 698–721, 2021.

[300] F. Wu, Z. Gu, H. Wu, A. Wu, and Y. Zhao, “Infer and adapt: Bipedal locomotion reward learning from demonstrations via inverse reinforcement learning,” in *IEEE International Conference on Robotics and Automation*, 2024, pp. 16 243–16 250.

[301] X. Cheng, J. Li, S. Yang, G. Yang, and X. Wang, “Open-television: Teleoperation with immersive active visual feedback,” in *Annual Conference on Robot Learning*, 2024.

[302] Y. Wang, J. Lin, A. Zeng, Z. Luo, J. Zhang, and L. Zhang, “Physhoi: Physics-based imitation of dynamic human-object interaction,” *arXiv preprint arXiv:2312.04393*, 2023.

[303] J. Merel, Y. Tassa, D. TB, S. Srinivasan, J. Lemmon, Z. Wang, G. Wayne, and N. Heess, “Learning human behaviors from motion capture by adversarial imitation,” *arXiv preprint arXiv:1707.02201*, 2017.

[304] A. Tang, T. Hiraoka, N. Hiraoka, F. Shi, K. Kawaharazuka, K. Kojima, K. Okada, and M. Inaba, “Humanmimic: Learning natural locomotion and transitions for humanoid robot via wasserstein adversarial imitation,” in *IEEE International Conference on Robotics and Automation*, 2024, pp. 13 107–13 114.

[305] Q. Zhang *et al.*, “Whole-body humanoid robot locomotion with human reference,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2024, pp. 11 225–11 231.

[306] X. B. Peng, A. Kanazawa, J. Malik, P. Abbeel, and S. Levine, “Sfv: Reinforcement learning of physical skills from videos,” *ACM Transactions On Graphics*, vol. 37, no. 6, pp. 1–14, 2018.

[307] H. Zhang, Y. Liu, and W. Zhou, “Deep adversarial imitation learning of locomotion skills from one-shot video demonstration,” in *IEEE Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems*, 2019, pp. 1257–1261.

[308] J. Li, Y. Zhu, Y. Xie, Z. Jiang, M. Seo, G. Pavlakos, and Y. Zhu, “Okami: Teaching humanoid robots manipulation skills through single video imitation,” in *Annual Conference on Robot Learning*, 2024.

[309] R. Cisneros-Limón *et al.*, “A cybernetic avatar system to embody human telepresence for connectivity, exploration, and skill transfer,” *International Journal of Social Robotics*, 2024.

[310] C. Lu, X. Cheng, J. Li, S. Yang, M. Ji, C. Yuan, G. Yang, S. Yi, and X. Wang, “Mobile-television: Predictive motion priors for humanoid whole-body control,” *arXiv preprint arXiv:2412.07773*, 2024.

[311] A. Dallard, M. Benallegue, F. Kanehiro, and A. Kheddar, “Synchronized human-humanoid motion imitation,” *IEEE Robotics and Automation Letters*, vol. 8, no. 7, pp. 4155–4162, 2023.

[312] X. Xu, Y. Zhang, Y.-L. Li, L. Han, and C. Lu, “HumanVLA: Towards vision-language directed object rearrangement by physical humanoid,” in *Annual Conference on Neural Information Processing Systems*, 2024.

[313] J. Juravsky, Y. Guo, S. Fidler, and X. B. Peng, “Superpadl: Scaling language-directed physics-based control with progressive supervised distillation,” in *Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers '24*. ACM, Jul. 2024, p. 1–11.

[314] M. A. Goodrich, J. W. Crandall, and E. Barakova, “Teleoperation and beyond for assistive humanoid robots,” *Reviews of Human Factors and Ergonomics*, vol. 9, no. 1, pp. 175–226, 2013.

[315] B. L. Bhatnagar, X. Xie, I. A. Petrov, C. Sminchisescu, C. Theobalt, and G. Pons-Moll, “Behave: Dataset and method for tracking human object interactions,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 15 935–15 946.

[316] Carnegie Mellon University, “Cmu graphics lab motion capture database,” 2002.

[317] Simon Fraser University and National University of Singapore, “Sfu motion capture database.”

[318] F. G. Harvey, M. Yurick, D. Nowrouzezahrai, and C. Pal, “Robust motion in-betweening,” vol. 39, no. 4, 2020.

[319] N. Mahmood, N. Ghorbani, N. F. Troje, G. Pons-Moll, and M. J. Black, “AMASS: Archive of motion capture as surface shapes,” in *International Conference on Computer Vision*, 2019, pp. 5442–5451.

[320] P. Varin, “Estimation and planning for dynamic robot behaviors,” Ph.D. dissertation, Harvard University, 2021.

[321] OpenAI, “Video generation models as world simulators,” 2024.

[322] A. Blattmann *et al.*, “Stable video diffusion: Scaling latent video diffusion models to large datasets,” *arXiv preprint:2311.15127*, 2023.

[323] W. Zhu, X. Ma, D. Ro, H. Ci, J. Zhang, J. Shi, F. Gao, Q. Tian, and Y. Wang, “Human motion generation: A survey,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, no. 4, pp. 2430–

2449, 2024.

[324] X. Cheng, Y. Ji, J. Chen, R. Yang, G. Yang, and X. Wang, “Expressive whole-body control for humanoid robots,” 2024.

[325] A. Fabisch, M. Uliano, D. Marschner, M. Laux, J. Brust, and M. Controzzi, “A Modular Approach to the Embodiment of Hand Motions from Human Demonstrations,” in *IEEE-RAS International Conference on Humanoid Robots*, 2022, pp. 801–808, iSSN: 2164-0580.

[326] X. B. Peng, Z. Ma, P. Abbeel, S. Levine, and A. Kanazawa, “Amp: adversarial motion priors for stylized physics-based character control,” *ACM Transactions On Graphics*, vol. 40, no. 4, 2021.

[327] P. Dugar, A. Shrestha, F. Yu, B. van Marum, and A. Fern, “Learning multi-modal whole-body control for real-world humanoid robots,” *arXiv preprint arXiv:2408.07295*, 2024.

[328] X. B. Peng, P. Abbeel, S. Levine, and M. van de Panne, “Deepmimic: example-guided deep reinforcement learning of physics-based character skills,” *ACM Transactions On Graphics*, vol. 37, no. 4, 2018.

[329] R. Yu, H. Park, and J. Lee, “Human dynamics from monocular video with dynamic camera movements,” *ACM Transactions on Graphics*, vol. 40, no. 6, pp. 1–14, 2021.

[330] Z. Xie, J. Tseng, S. Starke, M. van de Panne, and C. K. Liu, “Hierarchical planning and control for box loco-manipulation,” *Proc. ACM Comput. Graph. Interact. Tech.*, vol. 6, no. 3, 2023.

[331] M. Hassan, Y. Guo, T. Wang, M. Black, S. Fidler, and X. B. Peng, “Synthesizing physical character-scene interactions,” in *ACM SIGGRAPH 2023 Conference Proceedings*, 2023, pp. 1–9.

[332] F. Wu, Z. Gu, Y. Zhao, and A. Wu, “Learn to teach: Improve sample efficiency in teacher-student learning for sim-to-real transfer,” *arXiv preprint arXiv:2402.06783*, 2024.

[333] D. Youm, H. Jung, H. Kim, J. Hwangbo, H.-W. Park, and S. Ha, “Imitating and finetuning model predictive control for robust and symmetric quadrupedal locomotion,” *IEEE Robotics and Automation Letters*, vol. 8, no. 11, pp. 7799–7806, 2023.

[334] L. M. Smith, J. C. Kew, T. Li, L. Luu, X. B. Peng, S. Ha, J. Tan, and S. Levine, “Learning and Adapting Agile Locomotion Skills by Transferring Experience,” in *Robotics: Science and Systems*, 2023.

[335] F. Jenelten, J. He, F. Farshidian, and M. Hutter, “Dtc: Deep tracking control,” *Science Robotics*, vol. 9, no. 86, 2024.

[336] D. Kang, J. Cheng, M. Zamora, F. Zargarbashi, and S. Coros, “Rl + model-based control: Using on-demand optimal control to learn versatile legged locomotion,” *IEEE Robotics and Automation Letters*, vol. 8, no. 10, pp. 6619–6626, 2023.

[337] S. Levine and V. Koltun, “Guided policy search,” in *Proceedings of the International Conference on Machine Learning*, vol. 28, no. 3, 2013, pp. 1–9.

[338] D. Marew, N. Perera, S. Yu, S. Roelker, and D. Kim, “A biomechanics-inspired approach to soccer kicking for humanoid robots,” in *IEEE-RAS International Conference on Humanoid Robots*, 2024, pp. 722–729.

[339] T. Kwon, Y. Lee, and M. Van De Panne, “Fast and flexible multilegged locomotion using learned centroidal dynamics,” *ACM Transactions On Graphics*, vol. 39, no. 4, 2020.

[340] A. J. Ijspeert, J. Nakanishi, and S. Schaal, “Learning attractor landscapes for learning motor primitives,” in *Proceedings of the International Conference on Neural Information Processing Systems*, 2002, p. 1547–1554.

[341] Z. Xie, G. Berseth, P. Clary, J. Hurst, and M. van de Panne, “Feedback control for cassie with deep reinforcement learning,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2018, pp. 1241–1246.

[342] R. P. Singh, Z. Xie, P. Gergondet, and F. Kanehiro, “Learning bipedal walking for humanoids with current feedback,” *IEEE Access*, vol. 11, p. 82013–82023, 2023.

[343] Y. Fuchioka, Z. Xie, and M. Van de Panne, “Opt-mimic: Imitation of optimized trajectories for dynamic quadruped behaviors,” in *International Conference on Robotics and Automation*, 2023, pp. 5092–5098.

[344] F. Liu, Z. Gu, Y. Cai, Z. Zhou, S. Zhao, H. Jung, S. Ha, Y. Chen, D. Xu, and Y. Zhao, “Opt2skill: Imitating dynamically-feasible whole-body trajectories for versatile humanoid loco-manipulation,” *arXiv preprint arXiv:2409.20514*, 2024.

[345] V. C. Paredes and A. Hereid, “Resolved motion control for 3d under-actuated bipedal walking using linear inverted pendulum dynamics and neural adaptation,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2022, pp. 6761–6767.

[346] S. Coros, P. Beaudoin, and M. van de Panne, “Robust task-based control policies for physics-based characters,” *ACM Transactions on Graphics*, vol. 28, no. 5, p. Article 170, 2009.

[347] C. Li, E. Stanger-Jones, S. Heim, and S. bae Kim, “FLD: Fourier latent dynamics for structured motion representation and learning,” in *The International Conference on Learning Representations*, 2024.

[348] D. P. Kingma and M. Welling, “Auto-encoding variational bayes,” in *International Conference on Learning Representations*, 2014.

[349] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial networks,” *Communications of the ACM*, vol. 63, no. 11, p. 139–144, 2020.

[350] S. Bohez *et al.*, “Imitate and repurpose: Learning reusable robot movement skills from human and animal behaviors,” *arXiv preprint arXiv:2203.17138*, 2022.

[351] C. Tessler, Y. Kasten, Y. Guo, S. Mannor, G. Chechik, and X. B. Peng, “Calm: Conditional adversarial latent models for directable virtual characters,” in *ACM SIGGRAPH 2023 Conference Proceedings*, 2023, pp. 1–9.

[352] Z. Luo, J. Cao, J. Merel, A. Winkler, J. Huang, K. M. Kitani, and W. Xu, “Universal humanoid motion representations for physics-based control,” in *The International Conference on Learning Representations*, 2024.

[353] M. Liu, M. Zhu, and W. Zhang, “Goal-conditioned reinforcement learning: Problems and solutions,” in *Proceedings of the International Joint Conference on Artificial Intelligence*, 2022, pp. 5502–5511.

[354] A. Serifi, R. Grandia, E. Knoop, M. Gross, and M. Bächer, “Vmp: Versatile motion priors for robustly tracking motion on physical characters,” in *Computer Graphics Forum*. Wiley Online Library, 2024, p. e15175.

[355] D. Ha and J. Schmidhuber, “World models,” *Zenodo*, 2018.

[356] T. Hester, M. Quinlan, and P. Stone, “Generalized model learning for reinforcement learning on a humanoid robot,” in *IEEE International Conference on Robotics and Automation*, 2010, pp. 2369–2374.

[357] X. Gu, Y.-J. Wang, X. Zhu, C. Shi, Y. Guo, Y. Liu, and J. Chen, “Advancing Humanoid Locomotion: Mastering Challenging Terrains with Denoising World Model Learning,” in *Robotics: Science and Systems*, 2024.

[358] N. Hansen, J. S. V. V. Sobal, Y. LeCun, X. Wang, and H. Su, “Hierarchical world models as visual whole-body humanoid controllers,” in *International Conference on Learning Representations*, 2025.

[359] H. Shi, T. Li, Q. Zhu, J. Sheng, L. Han, and M. Q.-H. Meng, “An efficient model-based approach on learning agile motor skills without reinforcement,” in *IEEE International Conference on Robotics and Automation*, 2024, pp. 5724–5730.

[360] J. Gao, Z. Wang, Z. Xiao, J. Wang, T. Wang, J. Cao, X. Hu, S. Liu, J. Dai, and J. Pang, “CooHOI: Learning cooperative human-object interaction with manipulated object dynamics,” in *The Annual Conference on Neural Information Processing Systems*, 2024.

[361] J. Dao, H. Duan, and A. Fern, “Sim-to-real learning for humanoid box loco-manipulation,” in *IEEE International Conference on Robotics and Automation*, 2024, pp. 16930–16936.

[362] H. Munn, B. Tidd, D. Howard, and M. Gallagher, “Whole-body dynamic throwing with legged manipulators,” *arXiv preprint arXiv:2410.05681*, 2024.

[363] C. Yang, C. Pu, G. Xin, J. Zhang, and Z. Li, “Learning complex motor skills for legged robot fall recovery,” *IEEE Robotics and Automation Letters*, vol. 8, no. 7, pp. 4307–4314, 2023.

[364] F. Al-Hafez, G. Zhao, J. Peters, and D. Tateo, “Locomujoco: A comprehensive imitation learning benchmark for locomotion,” in *Robot Learning Workshop, NeurIPS*, 2023.

[365] B. Liu, Y. Zhu, C. Gao, Y. Feng, Q. Liu, Y. Zhu, and P. Stone, “Libero: Benchmarking knowledge transfer for lifelong robot learning,” *NeurIPS Datasets and Benchmarks Track*, vol. 36, 2023.

[366] Y. Liu, B. Yang, L. Zhong, H. Wang, and L. Yi, “Mimicking-bench: A benchmark for generalizable humanoid-scene interaction learning via human mimicking,” *arXiv preprint arXiv:2412.17730*, 2024.

[367] Z. Luo *et al.*, “Smplolympics: Sports environments for physically simulated humanoids,” *arXiv preprint arXiv:2407.00187*, 2024.

[368] Q. Liao, B. Zhang, X. Huang, X. Huang, Z. Li, and K. Sreenath, “Berkeley humanoid: A research platform for learning-based control,” *arXiv preprint arXiv:2407.21781*, 2024.

[369] H. Qiao, “The q-series humanoid robot system developed by the institute of automation, chinese academy of sciences,” 2024.

[370] S. Wilson, P. Glotfelter, L. Wang, S. Mayya, G. Notomista, M. Mote, and M. Egerstedt, “The robotarium: Globally impactful opportunities, challenges, and lessons learned in remote-access, distributed control of multirobot systems,” *IEEE Control Systems Magazine*, vol. 40, no. 1, pp. 26–44, 2020.

[371] M. Wuthrich *et al.*, “Trifinger: An open-source robot for learning dexterity,” in *Conference on Robot Learning*, vol. 155, 16–18 Nov 2021, pp. 1871–1882.

[372] G. Cheng, K. Ramirez-Amaro, M. Beetz, and Y. Kuniyoshi, “Purposive learning: Robot reasoning about the meanings of human activities,” *Science Robotics*, vol. 4, no. 26, p. eaav1530, 2019.

[373] A. Radford *et al.*, “Learning transferable visual models from natural language supervision,” *Proceedings of the 38th International Conference on Machine Learning*, vol. 139, pp. 8748–8763, 2021.

[374] J.-B. Alayrac *et al.*, “Flamingo: a visual language model for few-shot learning,” *Advances in Neural Information Processing Systems*, vol. 35, pp. 23 716–23 736, 2022.

[375] J. Mao, S. Zhao, S. Song, T. Shi, J. Ye, M. Zhang, H. Geng, J. Malik, V. Guizilini, and Y. Wang, “Learning from massive human videos for universal humanoid pose control,” *arXiv preprint:2412.14172*, 2024.

[376] J. Ho *et al.*, “Imagen video: High definition video generation with diffusion models,” *arXiv preprint arXiv:2210.02303*, 2022.

[377] NVIDIA *et al.*, “Gr00t n1: An open foundation model for generalist humanoid robots,” *arXiv preprint arXiv:2503.14734*, 2025.

[378] R.-Z. Qiu *et al.*, “Humanoid policy human policy,” *arXiv preprint arXiv:2503.13441*, 2025.

[379] J. DelPreto, C. Liu, Y. Luo, M. Foshey, Y. Li, A. Torralba, W. Matusik, and D. Rus, “Actionsense: A multimodal dataset and recording framework for human activities using wearable sensors in a kitchen environment,” in *Advances in Neural Information Processing Systems*, vol. 35, 2022, pp. 13 800–13 813.

[380] K. Grauman *et al.*, “Ego-exo4d: Understanding skilled human activity from first- and third-person perspectives,” in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 19 383–19 400.

[381] K. Werling, N. A. Bianco, M. Raitor, J. Stigel, J. L. Hicks, S. H. Collins, S. L. Delp, and C. K. Liu, “Add biomechanics: Automating model scaling, inverse kinematics, and inverse dynamics from human motion data through sequential optimization,” *PLOS ONE*, vol. 18, no. 11, pp. 1–25, 2023.

[382] R. Bonmassari *et al.*, “On the opportunities and risks of foundation models,” *arXiv preprint arXiv:2108.07258*, 2022.

[383] OpenAI, “Gpt-4 technical report,” 2023.

[384] A. Brohan *et al.*, “RT-1: Robotics Transformer for Real-World Control at Scale,” in *Proceedings of Robotics: Science and Systems*, 2023.

[385] B. Zitkovich *et al.*, “Rt-2: Vision-language-action models transfer web knowledge to robotic control,” in *Conference on Robot Learning*, vol. 229, 2023, pp. 2165–2183.

[386] Y. Jiang, A. Gupta, Z. Zhang, G. Wang, Y. Dou, Y. Chen, L. Fei-Fei, A. Anandkumar, Y. Zhu, and L. Fan, “Vima: General robot manipulation with multimodal prompts,” in *International Conference on Machine Learning*, 2023.

[387] M. J. Kim *et al.*, “Openvla: An open-source vision-language-action model,” in *Conference on Robot Learning*, vol. 270, 2025, pp. 2679–2713.

[388] G. R. Team *et al.*, “Gemini robotics: Bringing ai into the physical world,” 2025.

[389] Q. Bu *et al.*, “Agibot world colosseo: A large-scale manipulation platform for scalable and intelligent embodied systems,” *arXiv preprint arXiv:2503.06669*, 2025.

[390] Y. J. Ma, W. Liang, G. Wang, D.-A. Huang, O. Bastani, D. Jayaraman, Y. Zhu, L. Fan, and A. Anandkumar, “Eureka: Human-level reward design via coding large language models,” in *International Conference on Learning Representations*, 2024.

[391] J. Liang, W. Huang, F. Xia, P. Xu, K. Hausman, B. Ichter, P. Florence, and A. Zeng, “Code as policies: Language model programs for embodied control,” in *IEEE International Conference on Robotics and Automation*, 2023, pp. 9493–9500.

[392] brian ichter *et al.*, “Do as i can, not as i say: Grounding language in robotic affordances,” in *Annual Conference on Robot Learning*, 2022.

[393] Y. J. Ma, W. Liang, H.-J. Wang, Y. Zhu, L. Fan, O. Bastani, and D. Jayaraman, “DrEureka: Language Model Guided Sim-To-Real Transfer,” in *Robotics: Science and Systems*, 2024.

[394] Y. Yao, W. He, C. Gu, J. Du, F. Tan, Z. Zhu, and J. Lu, “Anybipe: An end-to-end framework for training and deploying bipedal robots guided by large language models,” *arXiv preprint arXiv:2409.08904*, 2024.

[395] A. S. Chen, A. M. Lessing, A. Tang, G. Chada, L. Smith, S. Levine, and C. Finn, “Commonsense reasoning for legged robot adaptation with vision-language models,” *arXiv preprint arXiv:2407.02666*, 2024.

[396] J. Juravsky, Y. Guo, S. Fidler, and X. B. Peng, “Padl: Language-directed physics-based character control,” in *SIGGRAPH Asia 2022 Conference Papers*, 2022, pp. 1–9.

[397] J. Sun, Q. Zhang, Y. Duan, X. Jiang, C. Cheng, and R. Xu, “Prompt, plan, perform: Llm-based humanoid control via quantized imitation learning,” in *IEEE International Conference on Robotics and Automation*, 2024, pp. 16 236–16 242.

[398] J. Wang, R. Dai, W. Wang, L. Rossini, F. Ruscelli, and N. Tsagarakis, “HYPERmotion: Learning hybrid behavior planning for autonomous loco-manipulation,” in *Annual Conference on Robot Learning*, 2024.

[399] I. Singh, V. Blukis, A. Mousavian, A. Goyal, D. Xu, J. Tremblay, D. Fox, J. Thomason, and A. Garg, “Progprompt: Generating situated robot task plans using large language models,” in *IEEE International Conference on Robotics and Automation*, 2023, pp. 11 523–11 530.

[400] W. Yu *et al.*, “Language to rewards for robotic skill synthesis,” in *Annual Conference on Robot Learning*, 2023.

[401] Z. Jiang, Y. Xie, J. Li, Y. Yuan, Y. Zhu, and Y. Zhu, “Harmon: Whole-body motion generation of humanoid robots from language descriptions,” in *Annual Conference on Robot Learning*, 2024.

[402] K. N. Kumar, I. Essa, and S. Ha, “Words into action: Learning diverse humanoid robot behaviors using language guided iterative motion refinement,” *CoRR*, vol. abs/2310.06226, 2023.

[403] Z. Xiao, T. Wang, J. Wang, J. Cao, W. Zhang, B. Dai, D. Lin, and J. Pang, “Unified human-scene interaction via prompted chain-of-contacts,” in *The International Conference on Learning Representations*, 2024.

[404] D. Totsila, Q. Rouxel, J.-B. Mouret, and S. Ivaldi, “Words2contact: Identifying support contacts from verbal instructions using foundation models,” in *IEEE-RAS International Conference on Humanoid Robots*, 2024, pp. 9–16.

[405] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. u. Kaiser, and I. Polosukhin, “Attention is all you need,” in *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[406] K. Black *et al.*, “π0: A vision-language-action flow model for general robot control,” *arXiv preprint arXiv:2410.24164*, 2024.

[407] A. O’Neill *et al.*, “Open x-embodiment: Robotic learning datasets and rt-x models : Open x-embodiment collaboration0,” in *IEEE International Conference on Robotics and Automation*, 2024, pp. 6892–6903.

[408] S. Reed *et al.*, “A generalist agent,” *Transactions on Machine Learning Research*, issn: 2835-8856, 2022.

[409] A. Gu and T. Dao, “Mamba: Linear-time sequence modeling with selective state spaces,” *arXiv preprint arXiv:2312.00752*, 2023.

[410] D. Ghosh *et al.*, “Octo: An Open-Source Generalist Robot Policy,” in *Robotics: Science and Systems*, 2024.

[411] S. Liu, L. Wu, B. Li, H. Tan, H. Chen, Z. Wang, K. Xu, H. Su, and J. Zhu, “RDT-1b: a diffusion foundation model for bimanual manipulation,” in *International Conference on Learning Representations*, 2025.

[412] C.-L. Cheang *et al.*, “Gr-2: A generative video-language-action model with web-scale knowledge for robot manipulation,” *arXiv preprint arXiv:2410.06158*, 2024.

[413] I. Leal *et al.*, “Sara-rt: Scaling up robotics transformers with self-adaptive robust attention,” in *International Conference on Robotics and Automation*, 2024, pp. 6920–6927.

[414] H. Touvron *et al.*, “Llama: Open and efficient foundation language models,” *arXiv preprint arXiv:2302.13971*, 2023.

[415] J. Kaplan, S. McCandlish, T. Henighan, T. B. Brown, B. Chess, R. Child, S. Gray, A. Radford, J. Wu, and D. Amodei, “Scaling laws for neural language models,” *arXiv preprint arXiv:2001.08361*, 2020.

[416] F. Lin, Y. Hu, P. Sheng, C. Wen, J. You, and Y. Gao, “Data scaling laws in imitation learning for robotic manipulation,” *arXiv preprint arXiv:2410.18647*, 2024.

[417] S. Nasiriany *et al.*, “Pivot: Iterative visual prompting elicits actionable knowledge for vlm,” *arXiv preprint arXiv:2402.07872*, 2024.

[418] D. Kim, H. Woo, and Y. Lee, “Addressing bias and fairness using fair federated learning: A synthetic review,” *Electronics*, vol. 13, no. 23, p. 4664, 2024.

[419] Z. Zhuang and H. Zhao, “Embrace collisions: Humanoid shadowing for deployable contact-agnostic motions,” 2025.

[420] G. Ficht and S. Behnke, “Bipedal humanoid hardware design: A technology review,” *Current Robotics Reports*, vol. 2, pp. 201–210, 2021.

[421] M.-H. Park and B.-K. Cho, “Design of new drive mechanism for dynamic movement of humanoid robot legs,” in *2023 IEEE-RAS 22nd International Conference on Humanoid Robots (Humanoids)*, 2023, pp. 1–8.

[422] A. Machín, C. Morant, and F. Márquez, “Advancements and challenges in solid-state battery technology: An in-depth review of solid electrolytes and anode innovations,” *Batteries*, vol. 10, no. 1, p. 29, 2024.

[423] Apptronik, “Apptronik apollo,” 2024.

[424] U. robotics, “Unitree h1 robot,” 2024.