# Safe and Adaptive 3-D Locomotion via Constrained Task-Space Imitation Learning

Ding, Jiatao; Lam, Tin Lun; Ge, Ligang; Pang, Jianxin; Huang, Yanlong

**DOI**
10.1109/TMECH.2023.3239099

**Publication date**
2023

**Document Version**
Final published version

**Published in**
IEEE/ASME Transactions on Mechatronics

**Citation (APA)**
Ding, J., Lam, T. L., Ge, L., Pang, J., & Huang, Y. (2023). Safe and Adaptive 3-D Locomotion via Constrained Task-Space Imitation Learning. *IEEE/ASME Transactions on Mechatronics*, *28*(6), 3029-3040. https://doi.org/10.1109/TMECH.2023.3239099

# Safe and Adaptive 3-D Locomotion via Constrained Task-Space Imitation Learning

Jiatao Ding, *Member, IEEE*, Tin Lun Lam, *Senior Member, IEEE*, Ligang Ge, Jianxin Pang, *Member, IEEE*, and Yanlong Huang

***Abstract*—Bipedal locomotion has been widely studied in recent years, where passive safety (i.e., a biped rapidly brakes without falling) is deemed to be a pivotal problem. To realize safe 3-D walking, existing works resort to nonlinear optimization techniques based on simplified dynamics models, requiring hand-tuned reference trajectories. In this article, we propose to integrate safety constraints into constrained task-space imitation learning, endowing a humanoid robot with adaptive walking capability. Specifically, unlike previous work using nonlinear and coupled capturability dynamics, we first linearize the 3-D capture conditions using appropriate extreme values and then seamlessly incorporate them into constrained imitation learning. Furthermore, we propose novel heuristic rules to define control points, enabling adaptive locomotion learning. The resulting framework allows robots to learn locomotion skills from a few demonstrations efficiently and apply the learned skills to unseen 3-D scenarios while satisfying the constraints for passive safety. Unlike deep enforcement learning, our framework avoids the need of a large number of iterations or sim-to-real transfer. By virtue of the task-space adaptability, the proposed imitation learning framework can reuse collected demonstrations in a new robot platform. We validate our method by hardware experiments on Walker2 robot and simulations on COMAN robot.**

***Index Terms*—3-D walking, bipedal locomotion, constrained imitation learning, humanoid robot, passive safety.**

Manuscript received 14 May 2022; revised 20 August 2022 and 19 November 2022; accepted 11 January 2023. Date of publication 20 February 2023; date of current version 15 December 2023. Recommended by Technical Editor H.-W. Park and Senior Editor N. G. Tsagarakis. This work was supported in part by the EU project 101016970 NI, in part by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie under Grant 101018395, in part by the Shenzhen Institute of Artificial Intelligence and Robotics for Society under Grant 2019-ICP002, and in part by the Program of Guangdong Provincial Key Laboratory of Robot Localization and Navigation Technology, under Grant 2020B121202011. *(Corresponding authors: Tin Lun Lam; Jianxin Pang.)*

Jiatao Ding is with the Cognitive Robotics, Delft University of Technology, 2628 CD Delft, The Netherlands (e-mail: jtdingx@163.com).

Tin Lun Lam is with the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen 518172, China, and also with the Shenzhen Institute of Artificial Intelligence and Robotics for Society, Shenzhen 518000, China (e-mail: tllam@cuhk.edu.cn).

Ligang Ge and Jianxin Pang are with the Ubtech Robotics Corporation, Shenzhen 518071, China (e-mail: greg.ge@ubtrobot.com; walton@ubtrobot.com).

Yanlong Huang is with the School of Computing, University of Leeds, Leeds LS29JT, U.K. (e-mail: y.l.huang@leeds.ac.uk).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TMECH.2023.3239099.

Digital Object Identifier 10.1109/TMECH.2023.3239099

## I. Introduction

SAFE walking has been studied from different perspectives, e.g., collision avoidance [1], [2], [3] and balance maintenance [4], [5], [6]. Recently, the concept of *passive safety* which was first exploited in mobile robots navigation [7] has attracted much attention [8], [9], [10]. In addition to preserving balance, passive safety also requires a robot to come to a stop after a finite number of steps or even zero step [9], which can be ensured by obeying $N$- or zero-step capturability constraints [11].

In many previous works, the assumption of a constant height is made to attain the $N$-step or zero-step capturability [11], whereas this requirement will become stringent for 3-D walking tasks (e.g., climbing stairs or walking across noncoplanar terrains) where the height variance is inevitable. There are some extensions of the $N$/zero-step capturability to 3-D cases in terms of divergent component of motion (DCM) analysis, e.g., [9], [12], [13]. Nevertheless, these works often involve nonlinear or implicit constraints, and most of them require extra simplification on the center of mass (CoM) or center of pressure (CoP) motion. For example, sum-of-squares programming was used in [12] to calculate the inner and outer approximations of $N$-step capture region, requiring Taylor approximation of the CoM motion. In [9], 3-D zero-step capturability was attained by linearly constraining the CoM motion, where a specific expression of CoM trajectory is needed.

Usually, the aforementioned techniques, e.g., model predictive control (MPC) [4], [9] and nonlinear programming [12], require a hand-tuned reference trajectory in advance, which may become infeasible or even unstable for a real humanoid robot. To overcome these issues, one can adopt human-inspired locomotion policies, e.g., [14], [15], and [16]. As an emerging topic, *learning from demonstrations* (also known as imitation learning) provides an efficient solution for mimicking expertise motions. By adopting imitation learning schemes, skillful gait representations can be obtained [17], [18], [19]. For instance, dynamic movement primitives (DMP) were employed to generate robust walking patterns against external disturbances in [18] and [19]. However, [17], [18], [19] focus on imitation learning without considering safety constraints.

In order to achieve adaptive 3-D bipedal locomotion with a guarantee of passive safety, this article first derives linear and decoupled conditions for safe walking, without needing extra assumptions on the CoM motion. Then, we integrate these constraints with imitation learning [20] and provide a framework

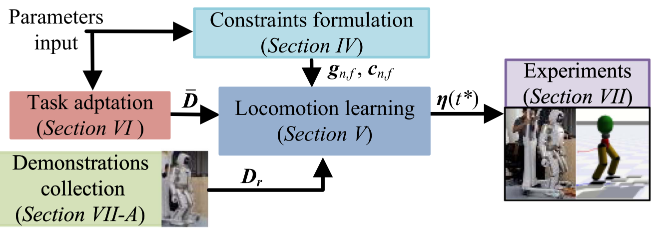


Fig. 1. Flowchart of the proposed approach.

capable of imitating external demonstrations while satisfying the safety constraints. In this framework, the robot learns motions over two-step cycles from demonstrations simultaneously, with the one-step and zero-step capturability applied at the first and second steps, respectively. By virtue of the linear inverted pendulum (LIP) model, desired points in task space are defined to accomplish the adaptive walking.

The contributions are threefold. First, we propose a novel simplification of 3-D capturability (Section IV-A and IV-B), which yields linear and decoupled constraints for safe 3-D locomotion with the brake capability. Second, we integrate linearized capture conditions, together with linear feasibility constraints (Section IV-C), into a constrained imitation learning framework (Section V), achieving safe locomotion by learning from few demonstrates. Third, we design LIP-aided heuristic rules to choose control points, endowing the learning framework with task adaptability (Section VI). In Section VII, we verify our solution by simulations and hardware experiments, showing that our approach needs fewer parameters, requires a smaller time cost and provides a smoother DCM convergence than existing MPC strategies.

An overview of the article's structure is shown in Fig. 1.

## II. Related Work

### A. Capturability in LIP-Based Locomotion

The LIP model [21], has been used widely in locomotion control. In [22], based on the LIP model, stable locomotion was realized by tracking a predefined CoP trajectory [equivalent to zero moment point (ZMP) [23] in 2-D cases]. Differing from [22], where a CoP trajectory was defined in advance, in [24], [25], and [26], MPC was used to find the optimal CoM trajectory by restricting the CoP trajectory within the support region. In essence, the aforementioned works [22], [23], [24], [25], [26] control both the convergent and divergent parts of the LIP dynamics simultaneously. In contrast, only the DCM was controlled in [27]. Note that capture point (CP) [28] is equivalent to DCM in 2-D cases.

Built on the concept of DCM/CP, *capturability* is developed to measure the ability of a robot to stop after a certain number of steps. In this line, *the $N$-step ($N$ could be zero or $\infty$) capture region is defined as an admissible physical region which includes all states from which the robot can come to a stop after taking no more than $N$ steps* [11]. In [11], an analytic solution for computing boundaries of $N$-step capture region was provided, where the LIP model was used under the assumption of a constant height.[1] However, it is nontrivial to extend 2-D capturability to 3-D cases with varying heights.

### B. Capturability in 3-D Walking

To accomplish 3-D walking, Englsberger et al. [29] proposed to manipulate the 3-D DCM by tracking the virtual repellent point. In [30], given ZMP and vertical CoM trajectories, the time-varying natural frequency of the variable-height inverted pendulum (VHIP) was computed and a linear quadratic regulator (LQR) was formulated to compute the DCM trajectory. However, the authors in [29] and [30] fail to derive the capture conditions explicitly.

Some works aim to provide analytic or approximated boundaries of the capture region. In [12], using a quadratic Taylor approximation of height trajectory, the inner and outer approximations of the 3-D capture region were computed via the sum of squares. In [13], assuming that the CoM trajectory only moves along a straight line, the zero-step conditions in multicontact scenarios were derived. A linear approximation was proposed in [9] for walking with passive safety, which, however, requires a specific form (i.e., exponential expressions) of the CoM curve. Caron et al. [31] derived a boundedness condition for the VHIP model to guarantee capturability, yielding a nonconvex optimization problem. Besides, Caron et al. [31] presumed a linear CoP trajectory for achieving the zero-step capturablity and a piecewise constant CoP trajectory for the one-step capturability, restricting the motion space of the bipedal robot. More recently, Liu et al. [32] extended instantaneous capture point (ICP) [11] to instantaneous capture input and provided an analytical solution for the capture region. However, this work took ZMP as the control input, prohibiting its applications in scenarios with noncoplanar contacts.

Unlike the existing work, we propose to linearize the nonlinear part of the 3-D capture region and obtain linear and decoupled boundaries of $N$-step, especially zero-step and one-step, capturability, where extra assumptions on the CoM motion and CoP motion are avoided.

### C. Locomotion by Imitation Learning

For gait control, the MPC schemes, e.g., [24], [25], [26], require a hand-tuned reference trajectory beforehand for each robot and a tedious retuning after switching to a new platform. In contrast, imitation learning provides an efficient tool for obtaining natural and adaptive motion skills through mimicking an expert, such as DMP [33] and kernelized movement primitives (KMP) [34]. DMP was used as a pattern generator to generalize gaits by adjusting the frequency of periodic movements [17]. In [18] and [19], DMP was employed to attain robust gaits against external pushes, with the help of reinforcement learning (RL). Ding et al. [35] integrated KMP with a step adjustment strategy to synthesize adaptive gaits for real-world scenarios. Nevertheless, these works [17], [18], [19], and [35] neglect safety constraints, which are crucial for bipedal locomotion.

[1] Strictly speaking, the LIP allows for vertical motion but requires zero vertical acceleration (i.e., a constant natural frequency).

Recent progress in RL (e.g., [36], [37], and [38]) show potential applications on safe walking, but they require a large number of training iterations and sim-to-real transfer.

To address the above issues, we exploit a constrained imitation learning approach, namely, linearly constrained KMP (LC-KMP) [20], to account for safety constraints in 3-D locomotion. Specifically, we propose to learn task-space actions [38] in order to satisfy high-level task requirements while encouraging demonstrations reuse in a different platform.

## III. Preliminaries

In 2-D cases without height variation, a constant natural frequency ($\omega_0$) of the LIP model is defined as

$$\omega_0 = \sqrt{g/z_c} \tag{1}$$

with $z_c$ being the constant LIP height and $g$ denoting the gravitational acceleration.

Given a constant $\omega_0$, the 2-D DCM (also called CP [28] or ICP [11]) is defined as

$$\xi_\gamma = c_\gamma + \dot{c}_\gamma/\omega_0 \tag{2}$$

where $\xi_\gamma$, $c_\gamma$, and $\dot{c}_\gamma$ ($\gamma \in \{x, y\}$) denote the $\gamma$-component of DCM position, CoM position, and CoM velocity, respectively.

*1) 2-D Zero-Step Capture Conditions:* To achieve zero-step capturability, the DCM $\xi_\gamma$ should be confined within the support region (i.e., zero-step capture region $\mathbb{C}_0$) [11]. Namely,

$$\xi_\gamma \in \mathbb{C}_0. \tag{3}$$

Given a rectangular foot size, we expand (3) as

$$\underline{r}_\gamma \leq \xi_\gamma - d_\gamma \leq \overline{r}_\gamma \tag{4}$$

where $d_\gamma$ denotes the $\gamma$-component of horizontal step location. $\underline{r}_\gamma$ and $\overline{r}_\gamma$ separately denote the minimal and maximal bounds of the support region formed by the support foot.

*2) 2-D One-Step Capture Conditions:* To achieve one-step capturability, the DCM movement should be restricted into the one-step capture region [11]. That is,

$$\xi_\gamma \in \mathbb{C}_1(\omega_0) \tag{5}$$

where $\mathbb{C}_1(\omega_0)$ denotes the one-step capture region.

Assuming a rectangular foot, we can explain (5) as

$$\underline{l}_\gamma e^{-\omega_0 T} + \underline{r}_\gamma \leq \xi_\gamma - d_\gamma \leq \overline{l}_\gamma e^{-\omega_0 T} + \overline{r}_\gamma \tag{6}$$

where $\underline{l}_\gamma$ and $\overline{l}_\gamma$ represent the minimal and maximal bounds of the step size [i.e., step length ($l_x$) and step width ($l_y$)], respectively. $T$ is the step duration.

## IV. Linearized Conditions for Safe 3-D Walking

Herein, we first derive linear and decoupled capture conditions for 3-D walking, based on the VHIP. Then, we provide feasible constraints for robust walking with passive safety.

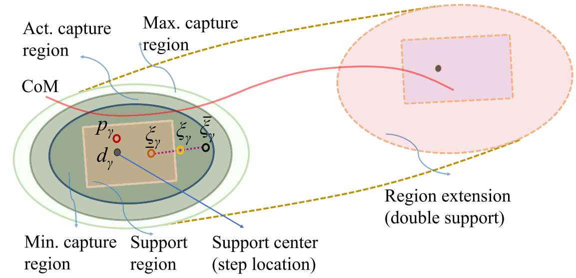


Fig. 2. Admissible physical region for safe walking. "Act. capture region" denotes the actual capture region calculated using the varying natural frequency. The possible maximal (Max.) and minimal (Min.) capture regions are determined by the boundary natural frequency.

### A. 3-D Dynamics With Variable Height

The VHIP model is used to model 3-D walking with varying height. The CoP ($p_\gamma$) of VHIP is

$$p_\gamma = c_\gamma - \ddot{c}_\gamma/\omega^2 \tag{7}$$

where $\omega$, representing the varying frequency, is defined as

$$\omega = \sqrt{(g + \ddot{c}_z)/(c_z - d_z)}. \tag{8}$$

Here, $\ddot{c}_z$ represents the vertical CoM acceleration, $c_z$ denotes the CoM height, and $d_z$ denotes the step height.

Following (8), the 3-D DCM becomes

$$\xi_\gamma = c_\gamma + \dot{c}_\gamma/\omega. \tag{9}$$

### B. 3-D Capturability Conditions

Let us assume that $\omega$ in (7) is bounded by

$$0 < \underline{\omega} \leq \omega \leq \overline{\omega} \tag{10}$$

where $\underline{\omega}$ and $\overline{\omega}$ denote the lower and upper bounds, respectively. As $\xi_\gamma$ in (9) is a monotonic function of $\omega$, we have

$$\underline{\xi}_\gamma \leq \xi_\gamma \leq \overline{\xi}_\gamma \tag{11}$$

where $\underline{\xi}_\gamma = \min\{c_\gamma + \dot{c}_\gamma/\overline{\omega},\ \ c_\gamma + \dot{c}_\gamma/\underline{\omega}\}$, $\overline{\xi}_\gamma = \max\{c_\gamma + \dot{c}_\gamma/\overline{\omega},\ \ c_\gamma + \dot{c}_\gamma/\underline{\omega}\}$ denote the lower and upper bounds of DCM, respectively. An illustration is given in Fig. 2.

*1) Linear Zero-Step Capture Conditions:* Since the support region is merely determined by the foot size, we can achieve 3-D zero-step capture conditions by limiting $\underline{\xi}_\gamma$ and $\overline{\xi}_\gamma$ as

$$\underline{r}_\gamma \leq \underline{\xi}_\gamma - d_\gamma \leq \overline{\xi}_\gamma - d_\gamma \leq \overline{r}_\gamma. \tag{12}$$

*2) Linear One-Step Capture Conditions:* Note that in (6) the one-step capture region only depends on $\omega$, Thus, we can determine the minimal and maximal capture regions (see Min. and Max. capture regions in Fig. 2) under varying height, i.e.,

$$\underline{\mathbb{C}}_1 = \mathbb{C}_1(\underline{\omega}) \cap \mathbb{C}_1(\overline{\omega}), \quad \overline{\mathbb{C}}_1 = \mathbb{C}_1(\underline{\omega}) \cup \mathbb{C}_1(\overline{\omega}) \tag{13}$$

where $\underline{\mathbb{C}}_1$ and $\overline{\mathbb{C}}_1$, respectively, denote the minimal and maximal one-step capture regions when the height is varying.

We define the one-step capturability condition by restricting $\xi_\gamma$ within the minimal capture region[2], i.e.,

$$\xi_\gamma \in \underline{\mathbb{C}}_1 \subset \overline{\mathbb{C}}_1. \tag{14}$$

Since $\xi_\gamma$ is bounded by $\underline{\xi}_\gamma$ and $\overline{\xi}_\gamma$ in (11), the constraint (14) can be ensured by

$$\underline{\xi}_\gamma \in \underline{\mathbb{C}}_1, \quad \overline{\xi}_\gamma \in \underline{\mathbb{C}}_1. \tag{15}$$

Furthermore, with the definition in (6), we have the one-step capturability conditions

$$\underline{s}_\gamma + \underline{r}_\gamma \le \underline{\xi}_\gamma - d_\gamma \le \overline{\xi}_\gamma - d_\gamma \le \overline{s}_\gamma + \overline{r}_\gamma \tag{16}$$

where $\underline{s}_\gamma$ and $\overline{s}_\gamma$ are determined by

$$\underline{s}_\gamma = \max\{\underline{l}_\gamma e^{-\underline{\omega}T}, \underline{l}_\gamma e^{-\overline{\omega}T}\}, \quad \overline{s}_\gamma = \min\{\overline{l}_\gamma e^{-\underline{\omega}T}, \overline{l}_\gamma e^{-\overline{\omega}T}\}. \tag{17}$$

*Remark 1:* Compared with previous works, such as [9], [12], [13], and [31], our linearization does not require extra assumptions/limits on the CoM/CoP movement when deriving 3-D capture conditions. Particularly, unlike [12] and [31], our linear zero/one-step capture conditions facilitate fast deployment. Moreover, the $N$-step capture conditions can be directly obtained by replacing $\underline{s}_\gamma$ and $\overline{s}_\gamma$ with $N$-step capture regions (explicitly derived in [11]), which, however, would be a hard task for [12] and [31].

### C. Feasibility Constraints for Safe Walking

To guarantee the safety without losing robustness, we plan the CoM motion over the current step and the next step (i.e., $0 \le t \le 2T$) at the beginning of each step cycle. In particular, at the current step the one-step capture conditions (16) and (17) are applied to enhance the robustness, while at the next step the zero-step capture constraints (12) are used to obtain the passive safety. Besides, the vertical motion limitation and slip avoidance are considered to guarantee the feasibility.

*1) Capturability Constraints Considering Double Support Phase:* Generally, one step cycle can be divided into a single support (SS) phase and a double support (DS) phase. In contrast to many previous works which ignore the DS phase, we formulate it explicitly for gait synthesis since the DS phase extends the capture region (see Fig. 2). Specifically, we assume that one step cycle starts from the middle of the current DS phase and ends at the middle of the next DS phase, as shown in Fig. 3. Consequently, two steps are divided into

$$\begin{cases} 1^{\text{st}}DS & (0 \le t \le T_d/2) \\ 1^{\text{st}}SS & (T_d/2 < t < T - T_d/2) \\ 2^{\text{nd}}DS & (T - T_d/2 \le t \le T + T_d/2) \\ 2^{\text{nd}}SS & (T + T_d/2 < t \le 2T) \end{cases} \tag{18}$$

where $T_d$ denotes the time period of one DS phase.[3]

[2]The extension of capture region caused by the height variation [12] is ignored here, which provides an inner approximation (subset) of the physically admissible region and yields a conservative but reasonable solution.

[3]Considering that the robot may rest on one leg in an emergency case, we ignore the DS at the end of the next step, leading to a more safe solution.

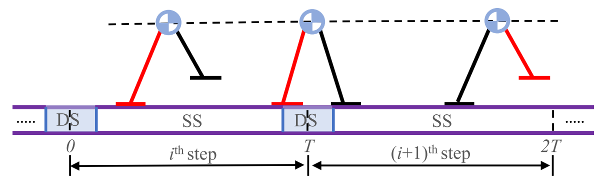


Fig. 3. Bipedal walking with support switch. The DS phase at the ending of the next ($(i+1)$th) step is ignored.

*DCM Constraints During the 1st DS:* During the 1st DS, the capture region extends. To gain the one-step capturablity, the capture region is set to be the sum of the zero-step capture region relative to the last step location and the one-step capture region relative to the current step location. Following (16) and (12), we have

$$\underline{r}_\gamma + \min\{d_\gamma^{i-1}, \underline{s}_\gamma + d_\gamma^i\} \le \underline{\xi}_\gamma \le \overline{\xi}_\gamma \le \overline{r}_\gamma + \max\{d_\gamma^{i-1}, \overline{s}_\gamma + d_\gamma^i\} \tag{19}$$

where $d_\gamma^{i-1}$ and $d_\gamma^i$ denote the previous and current step locations, respectively.

*DCM Constraints During the 1st SS:* The physically admissible region for DCM movement within the 1st SS coincides with the one-step capture region. By substituting the current step location $d_\gamma^i$ into (16), the DCM movement is limited by

$$\underline{s}_\gamma + \underline{r}_\gamma + d_\gamma^i \le \underline{\xi}_\gamma \le \overline{\xi}_\gamma \le \overline{s}_\gamma + \overline{r}_\gamma + d_\gamma^i. \tag{20}$$

*DCM Constraints During the 2nd DS:* At this stage, capture region also extends, which is determined by the one-step capture region relative to the current step location and the zero-step capture region relative to the next step location. Namely,

$$\underline{r}_\gamma + \min\{\underline{s}_\gamma + d_\gamma^i, d_\gamma^{i+1}\} \le \underline{\xi}_\gamma \le \overline{\xi}_\gamma \le \overline{r}_\gamma + \max\{\overline{s}_\gamma + d_\gamma^i, d_\gamma^{i+1}\} \tag{21}$$

where $d_\gamma^{i+1}$ denotes the next step location.

*DCM Constraints During the 2nd SS:* To guarantee the passive safety, the DCM movement relative to the next step location should lie into the zero-step capture region, i.e.,

$$\underline{r}_\gamma + d_\gamma^{i+1} \le \underline{\xi}_\gamma \le \overline{\xi}_\gamma \le \overline{r}_\gamma + d_\gamma^{i+1}. \tag{22}$$

*2) Constraints on the Vertical Motion:* A constraint on $c_z$ and $d_z$ is imposed to address the physical limitations, i.e.,

$$0 < \underline{z} \le c_z - d_z \le \overline{z} \tag{23}$$

where $\underline{z}$ and $\overline{z}$ are the lower and upper bounds of the vertical height. Here, $\underline{z}$ and $\overline{z}$ are chosen to satisfy the joint limits.

Furthermore, as in (8) $\omega$ relies on $g + \ddot{c}_z$, we introduce an additional constraint to ensure the viability of (10), i.e.,

$$(c_z - d_z)(\underline{\omega})^2 \le (g + \ddot{c}_z) \le (c_z - d_z)(\overline{\omega})^2. \tag{24}$$

*3) Friction Cone Constraints:* Using the single-mass VHIP model, the following constraints are added to prevent slippage:

$$-u \le \ddot{c}_\gamma/(g + \ddot{c}_z) \le u \tag{25}$$

where $u$ is the friction coefficient.

So far, we have obtained linear constraints for feasible 3-D walking, complying with passive safety requirements.

## V. Constrained Gait Imitation Learning

Now, we exploit the constrained imitation learning framework LC-KMP [20] to learn gaits, where the acceleration profile is also incorporated in this work. We first explain how the linear constraints for gait learning are formulated (Section V-A), and subsequently, we show how these constraints are integrated into the framework of LC-KMP (Section V-B).

### A. Linear Constraints for Gait Learning

For the sake of brevity, we formulate the linear constraints by taking the DCM constraints during the 2nd SS as an example, while the other constraints in Section IV-C can be tackled in a similar way. Following (9) and (11), the sufficient and necessary conditions for (22) are

$$\begin{aligned} \underline{r}_\gamma + d_\gamma^{i+1} &\leq c_\gamma + \dot{c}_\gamma/\underline{\omega} \leq \overline{r}_\gamma + d_\gamma^{i+1} \\ \underline{r}_\gamma + d_\gamma^{i+1} &\leq c_\gamma + \dot{c}_\gamma/\overline{\omega} \leq \overline{r}_\gamma + d_\gamma^{i+1}. \end{aligned} \tag{26}$$

Let us denote $\vec{\eta} = [c_x, c_y, c_z, \dot{c}_x, \dot{c}_y, \dot{c}_z, \ddot{c}_x, \ddot{c}_y, \ddot{c}_z]^\top$. At the time $t_n$, the constraints in (26) can be rewritten as (taking the first arrow for example)

$$\boldsymbol{g}_{n,1}^\top \eta(t_n) \geq c_{n,1}, \;\; \boldsymbol{g}_{n,2}^\top \eta(t_n) \geq c_{n,2} \tag{27}$$

with

$$\begin{aligned} \boldsymbol{g}_{n,1} &= [1, 0, 0, 1/\underline{\omega}, 0, 0, 0, 0, 0], \quad c_{n,1} = \underline{r}_\gamma + d_\gamma^{i+1} \\ \boldsymbol{g}_{n,2} &= [-1, 0, 0, -1/\underline{\omega}, 0, 0, 0, 0, 0], c_{n,2} = -(\overline{r}_\gamma + d_\gamma^{i+1}) \end{aligned} \tag{28}$$

### B. LC-KMP With Acceleration Learning

Assuming that we have access to $H$ demonstrations[4] $\boldsymbol{D} = \{\{t_{n,h}, \vec{\eta}_{n,h}\}_{n=1}^N\}_{h=1}^H$, where $N$ denotes the length of trajectory and $\vec{\eta}_{n,h} \in \Re^9$ corresponds to the Cartesian trajectory point at the $n$th time step from the $h$th demonstration. We use Gaussian mixture model (GMM) to model the joint probability distribution $\mathcal{P}(t, \boldsymbol{\eta})$ [34], [39], [40], yielding

$$\mathcal{P}(t, \boldsymbol{\eta}) = \sum_{m=1}^M \pi_m \mathcal{N}(\boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m) \tag{29}$$

where $\pi_m$, $\boldsymbol{\mu}_m = \begin{bmatrix} \boldsymbol{\mu}_{t,m} \\ \boldsymbol{\mu}_{\eta,m} \end{bmatrix}$ and $\boldsymbol{\Sigma}_m = \begin{bmatrix} \boldsymbol{\Sigma}_{tt,m} & \boldsymbol{\Sigma}_{t\eta,m} \\ \boldsymbol{\Sigma}_{\eta t,m} & \boldsymbol{\Sigma}_{\eta\eta,m} \end{bmatrix}$, respectively correspond to the prior probability, mean, and covariance of the $m$th Gaussian component in GMM. Subsequently, the Gaussian mixture regression (GMR) [34], [39] is used to retrieve a probabilistic reference trajectory $\boldsymbol{D}_r = \{t_n, \hat{\boldsymbol{\mu}}_n, \hat{\boldsymbol{\Sigma}}_n\}_{n=1}^N$, which encapsulates the distribution of demonstrations.

Let us write $\boldsymbol{\eta}(t)$ in a parametric form (see [20]), i.e.,

$$\boldsymbol{\eta}(t) = \boldsymbol{\Theta}(t)^\mathrm{T} \boldsymbol{w} \tag{30}$$

where $\boldsymbol{\Theta}$ represents a matrix consisting of basis function vectors and $\boldsymbol{w}$ denotes an unknown parameter vector. The constrained imitation learning can be addressed by solving

$$\begin{aligned} &\arg\max_{\boldsymbol{w}} \sum_{n=1}^N \mathcal{P}(\boldsymbol{\eta}(t_n) | \hat{\boldsymbol{\mu}}_n, \hat{\boldsymbol{\Sigma}}_n) \\ &s.t. \;\; \boldsymbol{g}_{n,f}^\top \boldsymbol{\eta}(t_n) \geq c_{n,f}, n \in \{1, 2, \ldots, N\}, f \in \{1, 2, \ldots, F\} \end{aligned} \tag{31}$$

with $\boldsymbol{g}_{n,f} \in \Re^9$ and $c_{n,f} \in \Re$ denoting the $f$th linear constraint acting on $\boldsymbol{\eta}(t_n)$ (see Section V-A).

With the definition of multivariate Gaussian distribution, we can rewrite (31) as

$$\begin{aligned} &\arg\min_{\boldsymbol{w}} \sum_{n=1}^N \frac{1}{2} (\boldsymbol{\Theta}(t_n)^\top \boldsymbol{w} - \hat{\boldsymbol{\mu}}_n)^\top \hat{\boldsymbol{\Sigma}}_n^{-1} (\boldsymbol{\Theta}(t_n)^\top \boldsymbol{w} - \hat{\boldsymbol{\mu}}_n) + \frac{\lambda}{2} \boldsymbol{w}^\top \boldsymbol{w} \\ &s.t. \;\; \boldsymbol{g}_{n,f}^\top \boldsymbol{\eta}(t_n) \geq c_{n,f}, n \in \{1, 2, \ldots, N\}, f \in \{1, 2, \ldots, F\}. \end{aligned} \tag{32}$$

Here, $\frac{\lambda}{2} \boldsymbol{w}^\top \boldsymbol{w}$ with $\lambda > 0$ acts as a regularization term.

By introducing Lagrange multipliers $\vec{\alpha}$ and the kernel trick, the constrained problem (32) can be solved. Specifically, given a query time $t^*$, LC-KMP predicts the corresponding trajectory point as

$$\boldsymbol{\eta}(t^*) = \boldsymbol{k}^* (\boldsymbol{K} + \lambda \boldsymbol{\Sigma})^{-1} (\boldsymbol{\mu} + \boldsymbol{\Sigma} \widetilde{\boldsymbol{G}} \boldsymbol{\alpha}^*) \tag{33}$$

where

$$\begin{aligned} \boldsymbol{k}^* &= \left[ \boldsymbol{k}(t^*, t_1) \; \boldsymbol{k}(t^*, t_2) \; \ldots \; \boldsymbol{k}(t^*, t_N) \right] \\ \boldsymbol{K} &= \begin{bmatrix} \boldsymbol{k}(t_1, t_1) & \boldsymbol{k}(t_1, t_2) & \ldots & \boldsymbol{k}(t_1, t_N) \\ \boldsymbol{k}(t_2, t_1) & \boldsymbol{k}(t_2, t_2) & \ldots & \boldsymbol{k}(t_2, t_N) \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{k}(t_N, t_1) & \boldsymbol{k}(t_N, t_2) & \ldots & \boldsymbol{k}(t_N, t_N) \end{bmatrix} \\ \boldsymbol{\Sigma} &= \mathrm{blockdiag}(\hat{\boldsymbol{\Sigma}}_1, \hat{\boldsymbol{\Sigma}}_2, \ldots, \hat{\boldsymbol{\Sigma}}_N) \\ \boldsymbol{\mu} &= [\hat{\boldsymbol{\mu}}_1, \hat{\boldsymbol{\mu}}_2, \ldots, \hat{\boldsymbol{\mu}}_N] \\ \boldsymbol{G}_n &= [\boldsymbol{g}_{n,1}, \boldsymbol{g}_{n,2}, \ldots, \boldsymbol{g}_{n,N_f}], \forall n \in \{1, 2, \ldots, N\} \\ \widetilde{\boldsymbol{G}} &= \mathrm{blockdiag}(\boldsymbol{G}_1, \boldsymbol{G}_2, \ldots, \boldsymbol{G}_N). \end{aligned} \tag{34}$$

Note that $\boldsymbol{k}(t_i, t_j) \in \Re^{9\times 9}$ in (34) are defined using a kernel function $k(\cdot, \cdot)$. $\alpha^*$ is the optimal Lagrange multiplier. Please see [41] for more details on the kernelization process. Therefore, we can generate safe 3-D locomotion by incorporating all linear constraints defined in Section IV-C into demonstration learning.

## VI. Heuristics for Task-Space Adaptation

In real-world environments, the robot often needs to adapt the learned skills to new scenarios, i.e, meeting task variation requirements. Here, we address the gait adaptation problem by defining proper desired points in terms of the CoM states, enabling the learning framework with locomotion adaptability. Particularly, we focus on the task space movements.

We define four desired points for every two steps (at the time $\{0, T/2, T, 2T\}$), where the first and fourth points are used to determine the start and end states while the second

[4] $3\sim5$ demonstrations under the same mode are sufficient in our evaluations. Data collection is explained in Section VII-A.

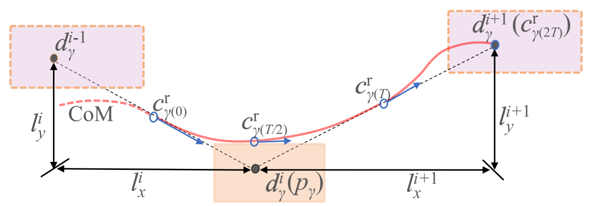

Fig. 4. Horizontal movement assuming zero CoP movement, i.e., $p_\gamma$ coincides with the support center $d_\gamma^i$ at each step.

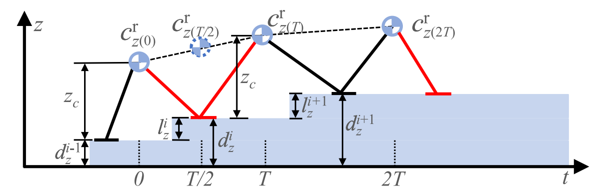

Fig. 5. Heuristic vertical trajectory for walking with varying height.

and third points account for the task variation. To comply with the constraints in Section IV-C, the CoM position, velocity, and acceleration at each desired point are defined. For the horizontal motion, the LIP model is used to choose the desired status. For the height variation, heuristic rules are adopted.

## A. Desired Points for Horizontal Movement Adaptation

*1) Boundary CoM Positions for the Current Step:* The main goal for the current step is to accomplish the desired walking task. We define the first (at $t = 0$) and the third (at $t = T$) desired points (relative to the current support center $d_\gamma^i$) as

$$c^{\mathrm{r}}_{\gamma(0)} = (d_\gamma^{i-1} - d_\gamma^i)/2, \;\; c^{\mathrm{r}}_{\gamma(T)} = (d_\gamma^{i+1} - d_\gamma^i)/2 \tag{35}$$

where $c^{\mathrm{r}}_{\gamma(0)}$ locates at the middle of the last and the current step locations while $c^{\mathrm{r}}_{\gamma(T)}$ locates at the middle of the current and the next step locations. An illustration is given in Fig. 4.

*2) Boundary CoM Positions for the Next Step:* To obtain the passive safety, we expect the robot to stop at the next $((i+1)$th) step. To achieve the highest stability, we set the fourth (at $t = 2T$) desired position (relative to $d_\gamma^{i+1}$) as

$$c^{\mathrm{r}}_{\gamma(2T)} = 0. \tag{36}$$

*3) CoM Position at $t = T/2$:* Assuming zero CoP movement at each step cycle (see Fig. 4), the horizontal CoM trajectory of an LIP is determined by

$$\ddot{c}_\gamma - \omega_0^2 c_\gamma = 0. \tag{37}$$

The CoM trajectory governed by (37) is fully determined by two boundary conditions, e.g., the CoM positions at the time 0 and $T$. An analytical solution for (37) is [42]

$$c_{\gamma(t_e)} = \begin{bmatrix} \cosh(\omega_0 t_e) & \frac{\sinh(\omega_0 t_e)}{\omega_0} \end{bmatrix} \begin{bmatrix} c_{\gamma(0)} \\ \dot{c}_{\gamma(0)} \end{bmatrix} \tag{38}$$

where $\sinh(\cdot)$ and $\cosh(\cdot)$ denote the hyperbolic sine and cosine functions, respectively. $t_e$ is the elapsed time within the current step. $[c_{\gamma(0)} \; \dot{c}_{\gamma(0)}]^{\mathrm{T}}$ comprises the initial CoM position and velocity determined by solving (38) under the constraints

$$c_{\gamma(t_e=0)} = c^{\mathrm{r}}_{\gamma(0)}, \quad c_{\gamma(t_e=T)} = c^{\mathrm{r}}_{\gamma(T)}. \tag{39}$$

As a result, the CoM position at the time $T/2$ is computed by (38) using $t_e = T/2$.

*4) CoM Velocities and Accelerations:* Once the desired CoM positions and $[c_{\gamma(0)} \; \dot{c}_{\gamma(0)}]^{\mathrm{T}}$ are determined, we can resort to the LIP model again to compute the velocity and acceleration for each desired point, i.e.,

$$\begin{bmatrix} \dot{c}_{\gamma(t_e)} \\ \ddot{c}_{\gamma(t_e)} \end{bmatrix} = \begin{bmatrix} \omega_0 \sinh(\omega_0 t_e) & \cosh(\omega_0 t_e) \\ \omega_0^2 \cosh(\omega_0 t_e) & \omega_0 \sinh(\omega_0 t_e) \end{bmatrix} \begin{bmatrix} c_{\gamma(0)} \\ \dot{c}_{\gamma(0)} \end{bmatrix}. \tag{40}$$

Note that the robot is always expected to stop at the end of the next step, so we set

$$\dot{c}^{\mathrm{r}}_{\gamma(2T)} = \ddot{c}^{\mathrm{r}}_{\gamma(2T)} = 0. \tag{41}$$

## B. Desired Points for Vertical Movement Adaptation

To accommodate height variation requirements, Specifically, we assume a piecewise linear height trajectory as the reference and calculate the vertical height using the nominal LIP height $z_c$ and the vertical step location $d_z$, see Fig. 5.

*1) Vertical Movement for the Current Step:* At the current step, the vertical CoM position of each desired point is

$$c^{\mathrm{r}}_{z(0)} = d_z^{i-1} + z_c, \; c^{\mathrm{r}}_{z(T/2)} = (d_z^{i-1} + d_z^i)/2 + z_c, \; c^{\mathrm{r}}_{z(T)} = d_z^i + z_c \tag{42}$$

where $c^{\mathrm{r}}_{z(0)}$, $c^{\mathrm{r}}_{z(T/2)}$, and $c^{\mathrm{r}}_{z(T)}$ separately denote the desired vertical positions at the time $\{0, T/2, T\}$. $d_z^{i-1}$ and $d_z^i$ separately denote the previous and current vertical step locations.

The vertical velocity and acceleration are then computed by

$$\dot{c}^{\mathrm{r}}_{z(0)} = \dot{c}^{\mathrm{r}}_{z(T/2)} = (d_z^i - d_z^{i-1})/T, \; \dot{c}^{\mathrm{r}}_{z(T)} = (d_z^{i+1} - d_z^{i-1})/(2T)$$
$$\ddot{c}^{\mathrm{r}}_{z(0)} = \ddot{c}^{\mathrm{r}}_{z(T/2)} = \ddot{c}^{\mathrm{r}}_{z(T)} = 0 \tag{43}$$

where $\{\dot{c}^{\mathrm{r}}_{z(0)}, \dot{c}^{\mathrm{r}}_{z(T/2)}, \dot{c}^{\mathrm{r}}_{z(T)}\}$ and $\{\ddot{c}^{\mathrm{r}}_{z(0)}, \ddot{c}^{\mathrm{r}}_{z(T/2)}, \ddot{c}^{\mathrm{r}}_{z(T)}\}$ separately comprise the reference vertical velocity and acceleration at the time $\{0, T/2, T\}$.

*2) Vertical Movement for the Next Step:* To stop at the next step, we set the fourth desired point as

$$c^{\mathrm{r}}_{z(2T)} = d_z^{i+1} + z_c, \quad \dot{c}^{\mathrm{r}}_{z(2T)} = \ddot{c}^{\mathrm{r}}_{z(2T)} = 0 \tag{44}$$

where $d_z^{i+1}$ is the next vertical step location. $\{c^{\mathrm{r}}_{z(2T)}, \dot{c}^{\mathrm{r}}_{z(2T)}, \ddot{c}^{\mathrm{r}}_{z(2T)}\}$ is the desired status when $t = 2T$.

*Remark 2:* It should be highlighted that the ending state defined in (36) and (41) ensures that the robot can stop with a single support foot, yielding safe locomotion in the scenario with a limited stepping zone. The designed CoM status at the four discrete points meet the feasible constraints defined in Section IV-C. However, with these desired points the CoM trajectory (see Figs. 4 and 5) does not necessarily comply with the VHIP dynamics in (7) strictly.

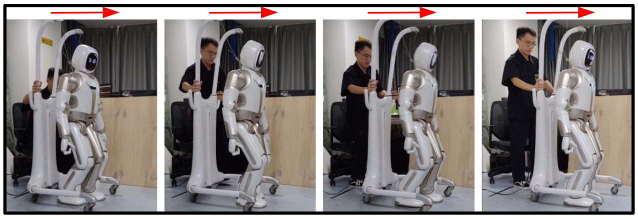

Fig. 6. Walker2 robot walks on a flat ground for data collection.

## C. Task Adaptation Using LC-KMP

Let us denote all desired points defined in Section VI-A and VI-B as $\bar{\boldsymbol{D}} = \{\bar{t}_i, \bar{\boldsymbol{\mu}}_i, \bar{\boldsymbol{\Sigma}}_i\}_{i=1}^4$, where $\bar{\boldsymbol{\mu}}_i \in \Re^9$ comprises the desired CoM position, velocity, and acceleration at the time $\bar{t}_i$ and $\bar{\boldsymbol{\Sigma}}_i \in \Re^{9\times 9}$ is the covariance at each time that is used to control the adaptation precision associated with $\bar{\boldsymbol{\mu}}_i$, i.e., how precisely the adapted CoM trajectory can pass through the desired point $\bar{\boldsymbol{\mu}}_i$ at the time $\bar{t}_i$.

Then we can simply concatenate $\bar{\boldsymbol{D}}$ with $\boldsymbol{D}_r$, resulting in an extended reference trajectory $\tilde{\boldsymbol{D}} = \bar{\boldsymbol{D}} \cup \boldsymbol{D}_r$. By learning $\tilde{\boldsymbol{D}}$ instead of $\boldsymbol{D}_r$, an adapted trajectory that passes through all four desired points while satisfying the linear constraints can be obtained.

*Remark 3:* We can reuse the collected demonstrations in a different robot platform. All we need to do is to scale the demonstrative motions following the kinematic relationship and then define the corresponding desired points.

In summary, by using the LIP-aided heuristic rules, the constrained imitation learning is able to plan safe 3-D locomotion while accounting for the task variation demand.

# VII. Evaluations

This section verifies the advantages of our method. We first introduce the experimental setup (Section VII-A). Then, we conduct simulations on a VHIP (Section VII-B) and compare our method with MPC-based approaches in [9], [43], and [44] (Section VII-C). Afterwards, we conduct hardware experiments on Walker2 [45] (Section VII-D). Finally, we extend our framework to the COMAN robot [46] (Section VII-E).

## A. Setup

We collect four demonstrations for LC-KMP by using the Walker2 robot [45] and apply these motions in the following evaluations. Regarding data collection, the robot repeatedly executes a periodic walking task on a flat ground ($l_z$=0 m, $l_x = 0.1$ m, $|l_y|$=0.2 m, $T$=0.7 s, and $z_c$=0.61 m), where the estimated CoM trajectory (i.e., pelvis center) is recorded as the demonstration. One demonstration is showcased in Fig. 6.

The Gaussian kernel $k(t_i, t_j) = \exp^{(-h(t_i-t_j)^2)}$ with $h$= 14 and the regularized coefficient $\lambda$=20 are used for LC-KMP. The other physical parameters are summarized in Table I.

TABLE I
Parameters Setup for the Walk2 Robot

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\underline{l}_x$[m] | -0.15 | $\bar{l}_x$[m] | 0.35 | $\lvert\underline{l}_y\rvert$[m] | 0.18 | $\lvert\bar{l}_y\rvert$[m] | 0.3 |
| $\underline{r}_x$[m] | -0.11 | $\bar{r}_x$[m] | 0.15 | $\underline{r}_y$[m] | -0.08 | $\bar{r}_y$[m] | 0.08 |
| $\underline{\omega}$[m] | 3.4 | $\bar{\omega}$[m] | 5.1 | $\underline{z}$[m] | 0.51 | $\bar{z}$[m] | 0.71 |

TABLE II
Tracking Errors of CoM Positions

| Error[m] \ Time[s] | 0 | 0.35 | 0.7 | 1.4 |
|---|---|---|---|---|
| $c_x$ | -1.7e-5 | 2.2e-5 | -1.3e-4 | -9.2e-5 |
| $c_y$ | 2.4e-5 | 4.3e-4 | -2.8e-3 | -5.8e-5 |
| $c_z$ | 2.6e-4 | -1.3e-4 | -5.6e-4 | 2.3e-4 |

## B. Safe Gait Learning With Task Adaptation for VHIP

In this test, the desired step lengths for the current and the next steps are separately set to be 0.25 and 0.2 m, and the first and second step heights are separately set to be 0.07 and $-0.07$ m. Note that the step lengths and heights are different from the ones in demonstrations.

The adapted CoM trajectories are illustrated in Fig. 8, showing that the proposed method can generate a 3-D trajectory that passes through the desired points, albeit that 2-D periodic gaits are used as demonstrations. Note that these 2-D demonstrations have negligible height variation, see $c_z^{\text{demo}}$ in Fig. 7. Although a piecewise linear height trajectory is assumed in Section VI-B, continuous vertical motion meeting the task adaptation requirement is obtained, see the smooth $c_z$ and $\dot{c}_z$ in Fig. 8. Numerical analysis in Table II indicates that position tracking errors at these desired points are below $3 \times 10^{-3}$ m, which is acceptable for a locomotion task. Thus, with only four desired points, the adaptive trajectory are generated. The DCM trajectories corresponding to the adapted CoM trajectories are plotted in Fig. 9, where the DS accounts for 20% of the whole period. Note that in Fig. 9 the DCM boundaries are unchanged after 0.77 s, since the DS phase at the end of the next step is ignored. The actual DCM profiles (plotted by yellow curves) computed using the varying $\omega$ fall into the capture region. Therefore, we can conclude that the one-step capturablity constraints are satisfied at the current step and zero-step capturablity constraints are respected at the next step. Also, the passive safety is attained as the CoM state ends with zero velocities (i.e., $\dot{c}_x = \dot{c}_y = \dot{c}_z = 0$ m/s) at $t$ = 1.4 s.

## C. Comparison With MPC-Based Approaches

In order to evidence the advantages of our solution, we compare it with vanilla linear MPC (LMPC) [43], linear MPC (LMPC) [9], and nonlinear MPC (NMPC) [44].

In [43], the 3-D CoM trajectory was generated by a vanilla LMPC approach under the assumption of a bounded natural frequency [see (10)]. The CoP stability was preserved in [43], where, however, the passive safety was overlooked. As an extension of [43], the passive safety was studied in [9], where the CoP movement over the whole prediction horizon was constrained

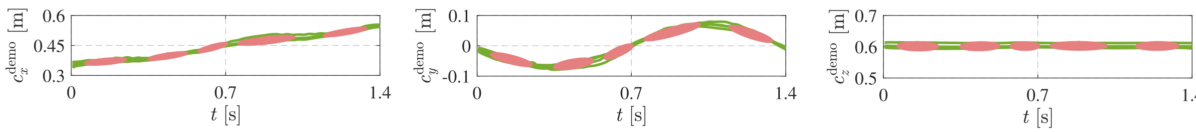


Fig. 7. Demonstrations (green curves) and the modeling of demonstrations using GMM. The red ellipses depict the Gaussian components.

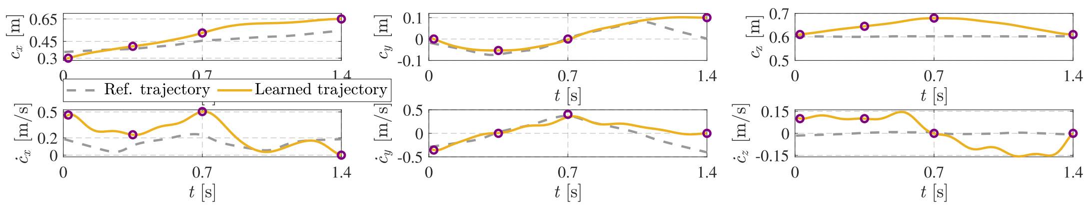


Fig. 8. Adapted CoM trajectories with 3-D capturability. The Ref. trajectories are retrieved from demonstrations via GMR. Circles plot desired points for task adaptation.

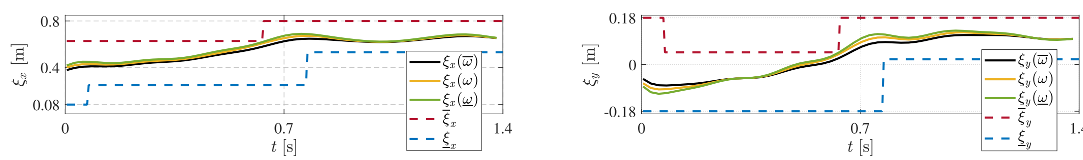


Fig. 9. DCM profiles generated using LC-KMP.

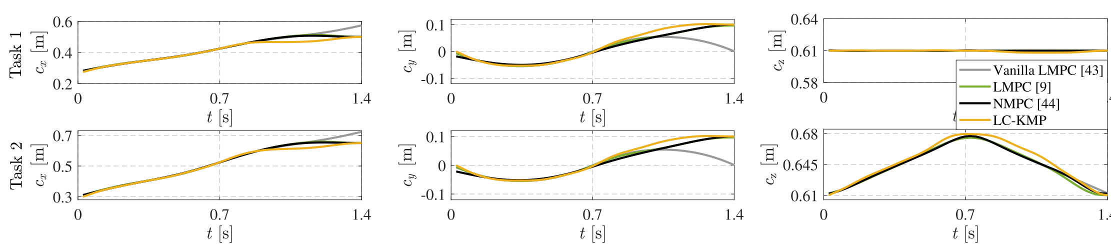


Fig. 10. CoM trajectories generated by different approaches.

and the zero-step capturability constraint was imposed at the end of the prediction horizon.

In [44], an NMPC was designed to generate 3-D gaits, where nonlinear CoP constraints induced by height variation were considered. However, the safety requirements were missing there. To make a fair comparison, we consider an improved version of [44] as a baseline, i.e., the constraints defined in (36), (41), and (44) are imposed at the end of the prediction horizon. Subsequently, NMPC [44] is used to generate safe 3-D gaits via sequential quadratic programming (SQP).

We consider two tasks: one is a 2-D task (Task 1) with a constant CoM height (0.61 m) and a constant step length (0.15 m), and the other (Task 2) is the 3-D task discussed in Section VII-B. For all MPC approaches, the prediction horizon is 1.4 s and time interval is 0.025 s. The CoM trajectories generated by different approaches are plotted in Fig. 10 and the corresponding DCM trajectories at the second step are depicted in Fig. 11.

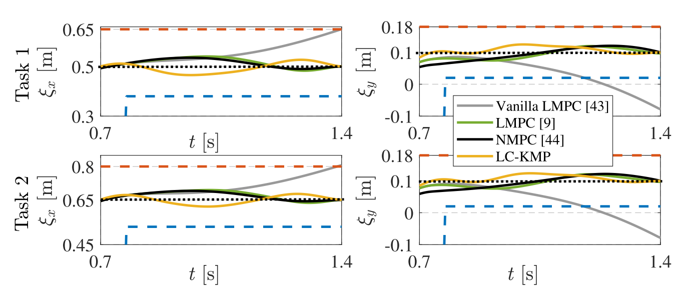


Fig. 11. DCM trajectories generated by different approaches. Red and blue dashed curves separately depict the upper and lower bounds of the support region. The dotted black lines represent the desired DCM values.

*1) Safety Performance:* In Fig. 11, only the vanilla LMPC [43] violates the zero-step captuability constraints considering the generated $\xi_y$ goes beyond the support region in Fig. 11. While LMPC [9], NMPC [44] and our approach all

TABLE III
MAXIMAL DCM (ABSOLUTE) ERRORS AT THE SECOND STEP

| Method \ Error[m] | Task 1 | | Task 2 | |
|---|---|---|---|---|
| | $\xi_x$ | $\xi_y$ | $\xi_x$ | $\xi_y$ |
| LMPC [9] | 0.0405 | 0.0257 | 0.0446 | 0.0242 |
| NMPC [44] | 0.0373 | 0.0412 | 0.0454 | 0.0403 |
| Ours | 0.0180 | 0.0170 | 0.0219 | 0.0167 |

TABLE IV
AVERAGE DCM (ABSOLUTE) ERRORS AT THE SECOND STEP

| Method \ Error[m] | Task 1 | | Task 2 | |
|---|---|---|---|---|
| | $\xi_x$ | $\xi_y$ | $\xi_x$ | $\xi_y$ |
| LMPC [9] | 0.0210 | 0.0153 | 0.0245 | 0.0140 |
| NMPC [44] | 0.0180 | 0.0170 | 0.0219 | 0.0167 |
| Ours | 0.0179 | 0.0099 | 0.0169 | 0.0095 |

TABLE V
TIME COSTS NEEDED BY DIFFERENT GENERATORS

| | vanilla LMPC [43] | LMPC [9] | NMPC [44] | Ours |
|---|---|---|---|---|
| Solver | QP | QP | SQP | QP |
| Safety | ✗ | ✓ | ✓ | ✓ |
| Task 1 time[s] | 0.34±0.04 | 2.40±1.32 | 0.96 ± 0.73 | 0.38±0.03 |
| Task 2 time[s] | 0.36±0.04 | 2.67±1.53 | 1.08 ± 0.76 | 0.43±0.03 |

obey the safety constraints, we compare their DCM trajectories against the desired DCM values (assuming the robot stops at the second step), which are plotted by the dotted black lines in Fig. 11. Tables III and IV summarize the maximal and average errors at the second step, showing that our approach has the smallest DCM errors in both tasks, and thus, a smoother DCM convergence toward the passive safety is achieved.

*2) Computing Efficiency:* The computational costs needed by different approaches[5] are reported in Table V. Compared with the vanilla LMPC [43] which ignored the safety constraints, LMPC with safety concerns [9] needs a larger time cost (around seven times of the vanilla LMPC in both tasks). Specifically, in LMPC, the extreme $\omega$s in linear constraints [see $\underline{\omega}$ and $\overline{\omega}$ in (10) and (11)] impose very strict restrictions to the optimizer and increase the number of constraints (two linear DCM constraints in [9] against one nonlinear DCM constraint in [44]), thus it has a larger computing cost than NMPC [44]. Although NMPC [44] does not tackle the extreme cases, it relies on multiple QP optimization within each iteration loop, leading to an extra time cost. As a result, the computational cost of NMPC is more expensive than our solution which only requires solving a single QP. Therefore, we can conclude that our work is most efficient for gait planning when safety constraints are included.

## D. Hardware Experiments on Walker2 Robot

We here verify our approach on a real Walker2 robot, where rich locomotion tasks are tested, including stone stepping, stair climbing, leg stretching, and brake motion.

*1) Safe 3-D Walking on Uneven Terrains:* In this setting, the robot first climbs a stone and then walks on a noncoplanar surface (the first row in Fig. 12). The corresponding trajectories are plotted in Fig. 13. From Fig. 12(i)(b)−(i)(b) and the forward motion (i.e., $x$ component) during 4.2∼5.6 s in Fig. 13, we can see that the robot makes a larger step (i.e., $l_x = 0.2$ m) than the demonstrated step in order to climb the stone (5 cm in height). Then, the robot walks straight on the noncoplanar surface, see Fig. 12(i)(d)−(i)(b). Finally, the robot comes to a stop on the stepping stone, as evidenced by the lateral ($y$ component) CoM trajectory after 10 s in Fig. 13.

The case of climbing a stair (stair height 10 cm, $T = 0.7$ s) is shown in the second row of Fig. 12. Note that in this climbing task, the tuning of the KMP parameters ($h$ and $\lambda$) is not needed. However, when using the traditional MPC framework, e.g., [9], [44], the prediction horizon needs to be altered and the control parameters (usually more than ten parameters) should also be tuned.

*2) Adaptive Walking With Leg Stretching:* In this case, the robot walks with stretched knees. The sagittal CoM and DCM trajectories are plotted in Fig. 14. Observing the actual CoM profile (magenta curve) in Fig. 14 ($z$ component), the robot increases the vertical height (5 cm) by stretching knee joints after 4 s so that it walks with almost a straight leg, which can be found in the third row of Fig. 12. After 5 s, the robot walks forward with $l_x = 0.1$ m and then walks backward with $l_x = -0.05$ m (see $x$ component during 9.8∼13.3 s in Fig. 14).

Thus, we can conclude that the adaptive and safe locomotion in 3-D challenging scenarios can be accomplished by merely learning 2-D periodic motions.

## E. Transfer Demonstrations to a Different Robot Platform

We now generalize the learned skills to the COMAN robot, with nominal $z_c = 0.457$ m, $l_x = 0.1$ m, and $|l_y| = 0.145$ m. Following the kinematic relationship, we scale the original demonstrations collected on the Walker2 robot through linear transformation. To comply with the actuation capability, the desired step duration is set as 0.8 s.

*1) 3-D Walking With Varying Height and Step Length:* Fig. 15 shows that the COMAN robot first walks across stairs (the stair height is 4 cm, step length $l_z = 0.15$ m) and then crouches across a low passage (reducing the CoM height by 4 cm). Note that the safe gait is generated by learning from Walker2's motion, where the time-consuming iterative search or fine-tuning is avoided, bringing advantages compared with the DRL framework [36], [37], [38] and MPC strategy [9], [44].

*2) Comparison Studies on Safe 3-D Locomotion:* In order to further illustrate the advantages of our solution, we compare our framework with two typical approaches for 3-D locomotion planning, including the DCM-based gait planner [30] and the NMPC approach [47]. In [30], given a height trajectory, the time-varying natural frequency $\omega$ was first calculated. Then, LQR was employed to generate 3-D DCM trajectories. Finally, the CoM profiles were computed by (8). By doing so, the zero-step capturability was obeyed in [30], whereas the ending CoM status was not confined into the support center of one single support foot. In [47], the safety constraints were ignored.

[5] All optimization problems are solved using the "*quadprog*" function provided in Matlab optimization toolbox.

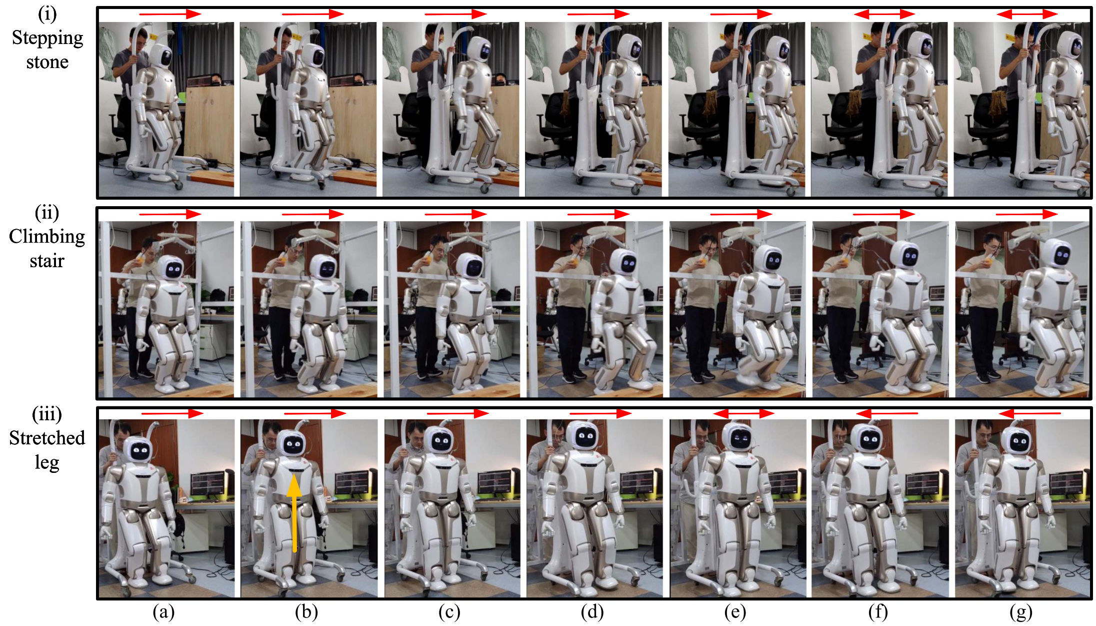


Fig. 12. Walker2 walks safely in real-world scenarios. The *first*, *second*, and *third* rows showcase the applications to stone stepping, stair climbing, and aperiodic walking with stretched legs. The red arrows above each row indicate walking directions and the orange arrow in (iv)(*b*) indicates leg stretching.

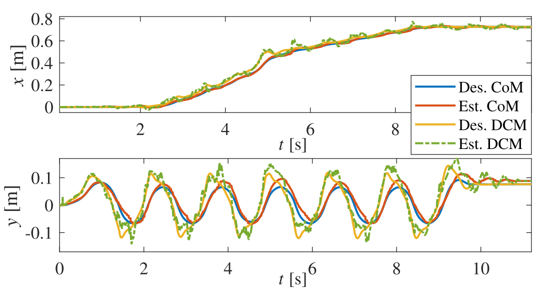


Fig. 13. Walker2's motion in a stone-stepping task. "Des." and "Est." trajectories separately denote the reference and actual trjectories.

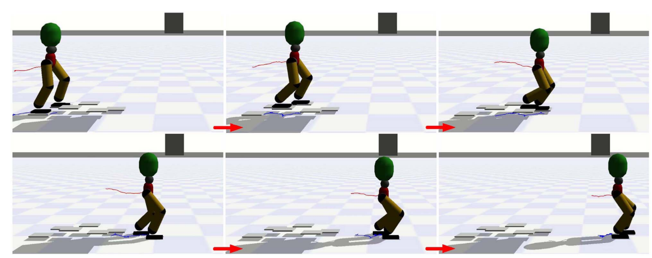

Fig. 15. COMAN robot walks safely in a 3-D scenario by learning from the Walker2's demonstrations. The *first row* shows the motions for stair climbing and the *second row* shows the snapshots when crouching across a passage.

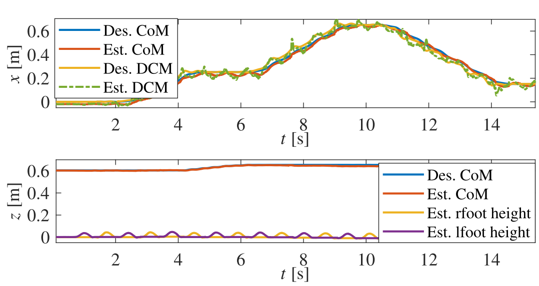


Fig. 14. Walker2's aperiodic walking with stretched knees.

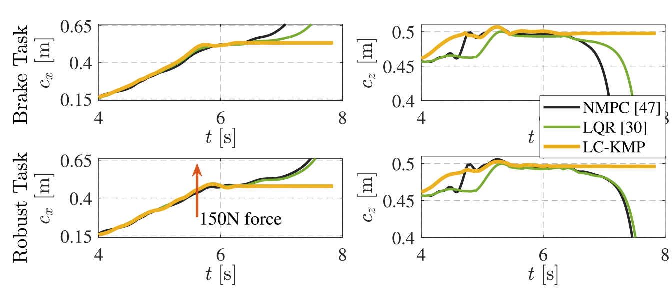


Fig. 16. Real CoM trajectories of the COMAN robot.

We consider two brake motions after climbing stairs: one is an emergency brake (brake task) after making a large step (with $l_x = 0.2$ m), the other is a robust brake against an external push force (robust task). In both tasks, we generated the CoM trajectories by LQR [30], NMPC [47], and our method separately, and mapped the CoM position to the pelvis center.

The actual CoM trajectories are plotted in Fig. 16. From Fig. 16, we can see that all methods accomplish the stair climbing task successfully, see the height ($c_z$) variation during 4–6 s. Nevertheless, when using NMPC [47] and LQR [30], the robot falls down in both tasks when it attempts to stop (see the rapid increase in $c_x$ and the rapid decrease in $c_z$ after 6 s). In contrast, using our solution, the robot stops safely in both tasks (see the constant $c_x$ and $c_z$ in yellow lines after 6 s).

## VIII. Conclusion

In this article, we have developed a task-space imitation learning framework for bipedal locomotion, where linear safety constraints are derived and integrated into a constrained learning approach. Throughout the extensive experiments, we have shown that our method can learn from a few 2-D demonstrations and generalize the learned skills to unseen 3-D scenarios while strictly obeying the safety constraints. Due to the task-space adaptation property, we can reuse the demonstrations in a different platform.

Taking imitation learning as the core, our method holds two key advantages over MPC approaches as follows.

1) The empirical reference trajectory is not needed, reducing the dependence on the simplified models.
2) Our method generates safe and adaptive 3-D gaits efficiently, while providing a smoother convergence.

Our method has a key advantage over the DRL-based framework [36], [37], [38], i.e., the need of a large number of iterations is avoided. Since we learn the task-space policy, high-level locomotion requirements can be met easily. In the future, it would be of interest to extend our framework to highly dynamic motions, e.g., jumping and running.

## References


[1] J. Warnke, A. Shamsah, Y. Li, and Y. Zhao, "Towards safe locomotion navigation in partially observable environments with uneven terrain," in *Proc. IEEE Conf. Decis. Control*., 2020, pp. 958–965.

[2] Z. Zhang, J. Yan, X. Kong, G. Zai, and Y. T. Liu, "Efficient motion planning based on kinodynamic model for quadruped robots following persons in confined spaces," *IEEE/ASME Trans. Mech.*, vol. 26, no. 4, pp. 1997–2006, Aug. 2021.

[3] O. Cebe, C. Tiseo, G. Xin, H.-C. Lin, J. Smith, and M. Mistry, "Online dynamic trajectory optimization and control for a quadruped robot," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2021, pp. 12773–12779.

[4] P.-B. Wieber, "Viability and predictive control for safe locomotion," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst*., 2008, pp. 1103–1108.

[5] J. Ding, L. Han, L. Ge, Y. Liu, and J. Pang, "Robust locomotion exploiting multiple balance strategies: An observer-based cascaded model predictive control approach," *IEEE/ASME Trans. Mech.*, vol. 27, no. 4, pp. 2089–2097, Aug. 2022.

[6] Y.-D. Hong and B. Lee, "Real-time feasible footstep planning for bipedal robots in three-dimensional environments using particle swarm optimization," *IEEE/ASME Trans. Mech.*, vol. 25, no. 1, pp. 429–437, Feb. 2020.

[7] K. Macek, D. A. V. Govea, T. Fraichard, and R. Siegwart, "Towards safe vehicle navigation in dynamic urban scenarios," *Automatika*, vol. 50, no. 3/4, pp. 184–194, 2009.

[8] N. Bohórquez, A. Sherikov, D. Dimitrov, and P.-B. Wieber, "Safe navigation strategies for a biped robot walking in a crowd," in *Proc. IEEE-RAS Int. Conf. Humanoid Robots*, 2016, pp. 379–386.

[9] A. Pajon and P.-B. Wieber, "Safe 3D bipedal walking through linear MPC with 3 D capturability," in *Proc. IEEE Int. Conf. Robot. Autom*., 2019, pp. 1404–1409.

[10] J. Ahn, J. Lee, and L. Sentis, "Data-efficient and safe learning for humanoid locomotion aided by a dynamic balancing model," *IEEE Robot. Autom. Lett.*, vol. 5, no. 3, pp. 4376–4383, Jul. 2020.

[11] T. Koolen, T. De, J. Boer, A. Rebula Goswami, and J. Pratt, "Capturability-based analysis and control of legged locomotion, part 1: Theory and application to three simple gait models," *Int. J. Robot. Res.*, vol. 31, no. 9, pp. 1094–1113, Jul. 2012.

[12] M. A. Posa, T. Koolen, and R. L. Tedrake, "Balancing and step recovery capturability via sums-of-squares optimization," in *Proc. Robot.: Sci. Syst*., Cambridge, MA, USA: MIT Press, 2017.

[13] A. Del Prete, S. Tonneau, and N. Mansard, "Zero step capturability for legged robots in multicontact," *IEEE Trans. Robot.*, vol. 34, no. 4, pp. 1021–1034, Aug. 2018.

[14] D. Tlalolini, C. Chevallereau, and Y. Aoustin, "Human-like walking: Optimal motion of a bipedal robot with toe-rotation motion," *IEEE/ASME Trans. Mech.*, vol. 16, no. 2, pp. 310–320, Apr. 2011.

[15] Q. Huang et al., "Resistant compliance control for biped robot inspired by humanlike behavior," *IEEE/ASME Trans. Mech.*, vol. 27, no. 5, pp. 3463–3473, Oct. 2022.

[16] W. Liu, J. Zhong, R. Wu, B. L. Fylstra, J. Si, and H. H. Huang, "Inferring human-robot performance objectives during locomotion using inverse reinforcement learning and inverse optimal control," *IEEE Robot. Autom. Lett.*, vol. 7, no. 2, pp. 2549–2556, Apr. 2022.

[17] J. Nakanishi, J. Morimoto, G. Endo, G. Cheng, S. Schaal, and M. Kawato, "Learning from demonstration and adaptation of biped locomotion," *Robot. Auton. Syst.*, vol. 47, no. 2/3, pp. 79–91, Jun. 2004.

[18] D. Luo, X. Han, Y. Ding, Y. Ma, Z. Liu, and X. Wu, "Learning push recovery for a bipedal humanoid robot with dynamical movement primitives," in *Proc. IEEE-RAS Int. Conf. Humanoid Robots*, 2015, pp. 1013–1019.

[19] J. Pankert, L. Kaul, and T. Asfour, "Learning efficient omni-directional capture stepping for humanoid robots from human motion and simulation data," in *Proc. IEEE-RAS Int. Conf. Humanoid Robots*, 2018, pp. 431–434.

[20] Y. Huang and D. G. Caldwell, "A linearly constrained nonparametric framework for imitation learning," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2020, pp. 4400–4406.

[21] S. Kajita, F. Kanehiro, K. Kaneko, K. Yokoi, and H. Hirukawa, "The 3D linear inverted pendulum mode: A simple modeling for a biped walking pattern generation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, 2001, pp. 239–246.

[22] S. Kajita et al., "Biped walking pattern generation by using preview control of zero-moment point," in *Proc. IEEE Int. Conf. Robot. Autom*., 2003, pp. 1620–1626.

[23] M. Vukobratović and B. Borovac, "Zero-moment point–thirty five years of its life," *Int J. Humanoid Robot.*, vol. 1, no. 1, pp. 157–173, Jan. 2004.

[24] P.-B. Wieber, "Trajectory free linear model predictive control for stable walking in the presence of strong perturbations," in *Proc. IEEE-RAS Int. Conf. Humanoid Robots*, 2006, pp. 137–142.

[25] H. Diedam, D. Dimitrov, P.-B. Wieber, K. Mombaur, and M. Diehl, "Online walking gait generation with adaptive foot positioning through linear model predictive control," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst*., 2008, pp. 1121–1126.

[26] N. Scianca, D. De Simone, L. Lanari, and G. Oriolo, "MPC for humanoid gait generation: Stability and feasibility," *IEEE Trans. Robot.*, vol. 36, no. 4, pp. 1171–1188, Aug. 2020.

[27] T. Takenaka, T. Matsumoto, and T. Yoshiike, "Real time motion generation and control for biped robot-3rd report: Dynamics error compensation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst*., 2009, pp. 1594–1600.

[28] J. Pratt, J. Carff, S. Drakunov, and A. Goswami, "Capture point: A step toward humanoid push recovery," in *Proc. IEEE-RAS Int. Conf. Humanoid Robots*, 2006, pp. 200–207.

[29] J. Englsberger, C. Ott, and A. Albu-Schäffer, "Three-dimensional bipedal walking control based on divergent component of motion," *IEEE Trans. Robot.*, vol. 31, no. 2, pp. 355–368, Apr. 2015.

[30] M. A. Hopkins, D. W. Hong, and A. Leonessa, "Humanoid locomotion on uneven terrain using the time-varying divergent component of motion," in *Proc. IEEE-RAS Int. Conf. Humanoid Robots*, 2014, pp. 266–272.

[31] S. Caron, A. Escande, L. Lanari, and B. Mallein, "Capturability-based pattern generation for walking with variable height," *IEEE Trans. Robot.*, vol. 36, no. 2, pp. 517–536, Apr., 2020.

[32] J. Liu, H. Chen, P. M. Wensing, and W. Zhang, "Instantaneous capture input for balancing the variable height inverted pendulum," *IEEE Robot. Autom. Lett.*, vol. 6, no. 4, pp. 7421–7428, Oct. 2021.

[33] A. J. Ijspeert, J. Nakanishi, H. Hoffmann, P. Pastor, and S. Schaal, "Dynamical movement primitives: Learning attractor models for motor behaviors," *Neural Comput.*, vol. 25, no. 2, pp. 328–373, Feb. 2013.

[34] Y. Huang, L. Rozo, J. Silvério, and D. G. Caldwell, "Kernelized movement primitives," *Int. J. Robot. Res.*, vol. 38, no. 7, pp. 833–852, May 2019.

[35] J. Ding, X. Xiao, N. Tsagaraki, and Y. Huang, "Robust gait synthesis combining constrained optimization and imitation learning," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst*., 2020, pp. 3473–3480.

[36] Z. Xie, P. Clary, J. Dao, P. Morais, J. Hurst, and M. Panne, "Learning locomotion skills for Cassie: Iterative design and sim-to-real," in *Proc. Conf. Robot Learn.*, 2020, pp. 317–329.

[37] X. B. Peng, E. Coumans, T. Zhang, T.-W. Lee, J. Tan, and S. Levine, "Learning agile robotic locomotion skills by imitating animals," in *Proc. Robot.: Sci. Syst*., 2020, pp. 1–12.

[38] H. Duan, J. Dao, K. Green, T. Apgar, A. Fern, and J. Hurst, "Learning task space actions for bipedal locomotion," in *Proc. IEEE Int. Conf. Robot. Autom*., 2021, pp. 1276–1282.

[39] S. Calinon, "A tutorial on task-parameterized movement learning and retrieval," *Intel. Serv. Robot.*, vol. 9, no. 1, pp. 1–29, 2016.

[40] Y. Huang, J. Silvério, and D. G. Caldwell, "Towards minimal intervention control with competing constraints," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, 2018, pp. 733–738.

[41] Y. Huang, F. J. Abu-Dakka, J. Silvério, and D. G. Caldwell, "Toward orientation learning and adaptation in cartesian space," *IEEE Trans. Robot.*, vol. 37, no. 1, pp. 82–89, Feb. 2020.

[42] J. Ding, C. Zhou, Z. Guo, X. Xiao, and N. Tsagarakis, "Versatile reactive bipedal locomotion planning through hierarchical optimization," in *Proc. IEEE Int. Conf. Robot. Autom*., 2019, pp. 256–262.

[43] C. Brasseur, A. Sherikov, C. Collette, D. Dimitrov, and P.-B. Wieber, "A robust linear MPC approach to online generation of 3D biped walking motion," in *Proc. IEEE-RAS Int. Conf. Humanoid Robots*, 2015, pp. 595–601.

[44] K. Van Heerden, "Real-time variable center of mass height trajectory planning for humanoids robots," *IEEE Robot. Autom. Lett.*, vol. 2, no. 1, pp. 135–142, Jan. 2017.

[45] Ubtech Robotics Corporation. [Online]. Available: https://www.ubtrobot.com/cn/

[46] N. Tsagarakis, S. Morfey, G. M. Cerda, L. Zhibin, and D. G. Caldwell, "Compliant humanoid COMAN: Optimal joint stiffness tuning for modal frequency control," in *Proc. IEEE Int. Conf. Robot. Autom*., 2013, pp. 673–678.

[47] J. Ding, C. Zhou, S. Xin, X. Xiao, and N. Tsagarakis, "Nonlinear model predictive control for robust bipedal locomotion: Exploring angular momentum and com height changes," *Adv. Robot.*, vol. 35, no. 18, pp. 1079–1097, Aug. 2021.

**Jiatao Ding** (Member, IEEE) received the B.Eng. and Ph.D. degrees in engineering from Wuhan University, Wuhan, China, in 2014 and 2020, respectively.

He worked as a Ph.D. Fellow with the Italian Institute of Technology, Genoa, Italy, and an Assistant Research Scientist with the Shenzhen Institute of Artificial Intelligence and Robotics for Society, Shenzhen, China. He is currently a Postdoctoral Researcher with the Department of Cognitive Robotics, 3mE, Delft University of Technology, The Netherlands. His research interests include locomotion control, imitation learning, and model predictive control.

**Tin Lun Lam** (Senior Member, IEEE) received the B.Eng. and Ph.D. degree in robotics and automation from the Chinese University of Hong Kong, Hong Kong, in 2006 and 2010, respectively.

He currently serves as Assistant Professor and Executive Deputy Director with the Robotics & AI Lab, Chinese University of Hong Kong, Shenzhen, China, and the Director of Research Center on Intelligent Robots, Shenzhen Institute of Artificial Intelligence and Robotics for Society, Shenzhen, China. He has authored or coauthored over 40 academic papers in robotics and automation His research interests include field robotics, human–machine interaction, and intelligent control.

**Ligang Ge** received the B.Eng. and M.Sc. degrees in control theory and engineering from the National University of Defense Technology, Changsha, China, in 2014 and 2016, respectively.

He is currently a Senior Engineer with UBTECH Robotics Corporation, Shenzhen, China. His research interests include gait generation for biped robot and balance control.

**Jianxin Pang** (Member, IEEE) received the B.Eng. and Ph.D. degrees in engineering from the University of Science and Technology of China, Hefei, China, in 2002 and 2008, respectively.

He is currently the Director of UBTECH Research Institute, and a Senior Engineer with UBTECH Robotics Corporation, Shenzhen, China. His interests include optimal control, machine learning, motion planning and their applications to humanoid robots, rescue robots, home service robots, and teaching robots.

**Yanlong Huang** received the Ph.D. degree in robotics from the Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2013.

He carried out his research as a Postdoctoral Researcher with the Max-Planck Institute for Intelligent Systems, Stuttgart, Germany, and Italian Institute of Technology, Genoa, Italy. He is currently a University Academic Fellow with the School of Computing, University of Leeds, Leeds, U.K. His interests include imitation learning, optimal control, reinforcement learning, and their applications to robotic systems.