Machine Intelligence Research
www.mi-research.net

22(4), August 2025, 627-640
DOI: 10.1007/s11633-024-1538-9

Reactive Whole-body Locomotion-integrated Manipulation Based on Combined Learning and Optimization

Jianzhuang Zhao[1,2] Tao Teng[3,4] Elena De Momi[2] Arash Ajoudani[1]

[1] Human-Robot Interfaces and Interaction Lab, Italian Institute of Technology, Genova 16163, Italy

[2] Department of Electronics, Information, and Bioengineering, Polytechnic University of Milan, Milano 20133, Italy

[3] Department of Mechanical and Automation Engineering, T-Stone Robotics Institute, The Chinese University of Hong Kong, Hong Kong 999077, China

[4] Hong Kong Center for Logistics Robotics, Hong Kong 999077, China

Abstract: Reactive planning and control capacity for collaborative robots is essential when the tasks change online in an unstructured environment. This is more difficult for collaborative mobile manipulators (CMM) due to high redundancies. To this end, this paper proposed a reactive whole-body locomotion-integrated manipulation approach based on combined learning and optimization. First, human demonstrations are collected, where the wrist and pelvis movements are treated as whole-body trajectories, mapping to the end-effector (EE) and the mobile base (MB) of CMM, respectively. A time-input kernelized movement primitive (T-KMP) learns the whole-body trajectory, and a multi-dimensional kernelized movement primitive (M-KMP) learns the spatial relationship between the MB and EE pose. According to task changes, the T-KMP adapts the learned trajectories online by inserting the new desired point predicted by M-KMP. Then, the updated reference trajectories are sent to a hierarchical quadratic programming (HQP) controller, where the EE and the MB trajectories tracking are set as the first and second priority tasks, generating the feasible and optimal joint level commands. An ablation simulation experiment with CMM of the HQP is conducted to show the necessity of MB trajectory tracking in mimicking human whole-body motion behavior. Finally, the tasks of the reactive pick-and-place and reactive reaching were undertaken, where the target object was randomly moved, even out of the region of demonstrations. The results showed that the proposed approach can successfully transfer and adapt the human whole-body loco-manipulation skills to CMM online with task changes.

Keywords: Embodied intelligence, robot learning, mobile manipulation, whole-body motion planning and control, learning from human demonstrations.

Citation: J. Zhao, T. Teng, E. De Momi, A. Ajoudani. Reactive whole-body locomotion-integrated manipulation based on combined learning and optimization. *Machine Intelligence Research*, vol.22, no.4, pp.627–640, 2025. http://doi.org/10.1007/s11633-024-1538-9

1 Introduction

Collaborative robots, including fixed-base and mobile base, have become increasingly vital in contemporary settings due to their transformative impact on industrial and service sectors[1–4]. In today's rapidly evolving landscape, characterized by dynamic market demands and technological advancements, collaborative robots must offer a flexible and adaptive solution to various industries and work closely with humans in unstructured environments[1, 5]. Therefore, with the rapid development of artificial intelligence, reactive planning and control play a pivotal role in collaborative robots, enabling them to respond promptly to changes in their surroundings, such as unexpected obstacles or alterations in the task requirements[6, 7]. Furthermore, integrating reactive control mechanisms allows collaborative robots to seamlessly interact with their environment, human counterparts, and other robots, fostering a collaborative ecosystem characterized by fluid adaptability and enhanced performance[3, 8, 9].

Fixed-base collaborative robotic manipulators only need to handle the changes inside their limited workspace. A lot of computer version techniques have been applied to detect task changes in the workspace. For example, a reactive motion planning and control approach based on a cascade of dependent quadratic programmings (QP) in [10] is proposed to ensure the end-effector (EE) motion of a 7 degree of freedom (DoF) collaborative arm is always inside the view of the camera (namely, obey this constraint) and is compliant to the randomly applied human's external force at running time. Based on hierarchical quadratic programming (HQP), an adaptative motion controller for robotic arm in [11] can respond to human disturbance online and generate compliant behavior, subjecting to the joint limits. Besides, model predictive control (MPC) in [12] is also applied for reactive

Research Article
Special Issue on Embodied Intelligence
Manuscript received on September 10, 2024; accepted on December 23, 2024; published online on June 27, 2025
Recommended by Associate Editor Wei He
Colored figures are available in the online version at https://link.springer.com/journal/11633
© The Author(s) 2025

motion planning for manipulators. In particular, this method can generate real-time collision-free trajectories following a moving target from a joystick in dynamic environments.

The HQP and MPC methods mentioned above work well for fixed-based robotic arms. However, reactive planning and control for collaborative mobile manipulators (CMM) is more complicated due to the high redundancies. CMM includes a mobile platform and a collaborative robotic arm mounted on top of the former, combining mobility and dexterity[3, 13, 14]. With enhanced mobility, CMM can work in larger workspaces and play a more and more important role in various applications, such as smart manufacturing factories, warehouses, domestic, and health care. Therefore, the learning/planning and control for CMM are the hot topics in the current robotics field. In general, the planning and control techniques for CMMs can be categorized into two groups based on how they handle the two components: Either as two distinct subsystems or as a unified whole-body system[15–17]. For the former category, usually, the mobile base moves close to the target object (locomotion mode) first, and the robotic arm starts to move after the base stopping (manipulation mode)[18, 19]. To ensure the feasibility of the manipulation phase of the robotic arm, the inverse reachability map (IRM)[18] is applied to determine the final mobile base placement during the locomotion period. Once given a desired EE pose, the IRM can find a feasible base placement for the mobile base. Although the IRM can be constructed offline and used online, it varies from one CMM to another. Meanwhile, it is non-trivial to construct the IRM.

For CMM, it is straightforward that the tasks′ efficiency and performance can be improved if the arm and mobile base can move simultaneously without stopping, namely, in loco-manipulation mode[16, 17]. To this end, many unified whole-body CMM methods have been proposed recently. For example, a whole-body kinematics perceptive MPC framework for CMM is proposed in [20], where the visual information and haptic sensing are used to avoid obstacles and control interaction forces separately. A similar MPC method is used to achieve nonprehensile object transportation and simultaneously avoid the static and dynamic obstacles for CMM in [21]. Besides, a holistic approach based on QP for CMM in [22] maximizes the manipulability of the arm and minimizes the angle between the mobile base and end-effector. The MPC/QP-based whole-body approaches can generate feasible motion for CMM as the arm, and base joint limits are set as constraints inside the optimization scheme. Besides, the MPC/QP approaches for collaborative robots (both fixed-base and mobile base) can respond to task changes by adapting trajectories online for the short (predicted) horizon. However, these methods can only be used as local planners since the long-horizon reference trajectory still needs to be provided. To generate the reference trajectory, a sampling-based method is proposed in [17] as a high-level planner, which is tracked by an optimization-based low-level controller. However, it is challenging for these methods to pass by a via-point with specific velocity requirements, like the contact velocity with the target object in the pick-and-place task in [16] (i.e., the trajectory modulation capability).

Besides, the learning-from-humans demonstration approach is applied[23–25] to generate the reference trajectory for collaborative robots. For example, based on an impedance controller, the reference trajectory and the variable impedance profile are learned for fixed-base robots in [23, 24]. For CMM, the EE trajectory and the desired interaction force in a table-cleaning task are learned from human demonstrations in [8]. The learned EE reference trajectory is sent to a whole-body Cartesian impedance controller to generate the overall joint-level command for CMM. Differently, the EE and mobile base motion in a door-opening task are learned from human demonstrations for CMM in [26]. However, the IRM is still necessary as a constraint to ensure the feasibility of the EE movements, making it challenging to apply this approach to other CMMs. Recently, a combined learning and optimization approach has been proposed in [16], transferring human whole-body loco-manipulation skills to a CMM with different geometry. Specifically, the kernelized movement primitives (KMP)[27] encode the human demonstrated whole-body trajectory, where the waist and pelvis motion are mapped to the EE and mobile base, respectively. Then, the learned trajectories are sent to the HQP controller, which separately sets the EE and mobile base tracking as primary and secondary tasks. This approach successfully learns human-demonstrated fluent and single whole-body movement (i.e., unified loco-manipulation mode) for CMM and adapts the learned trajectory to a new initial point. However, the target object position is the same (fixed) for all the experiments in [16]. It is still an open question of how to generate the whole-body reactive locomotion-integrated manipulation skills based on the learned reference trajectories for CMM when the target object is moved on the fly.

To address these issues, we propose an online reactive whole-body locomotion-integrated manipulation framework based on combined learning and optimization (see Fig. 1). In addition, to address the novelty of the proposed method, Table 1 compares our method with the most relevant works in the literature. According to the task changes, we aim to adapt the learned whole-body locomotion-integrated manipulation skills online reactively. Specifically, once the desired EE pose is changed, the corresponding feasible base placement will be found first, and the learned whole-body reference trajectory will be updated online later. The generated whole-body reference trajectory is sent to an HQP controller to generate the joint-level commands for CMM. The main contributions of this paper are as follows:

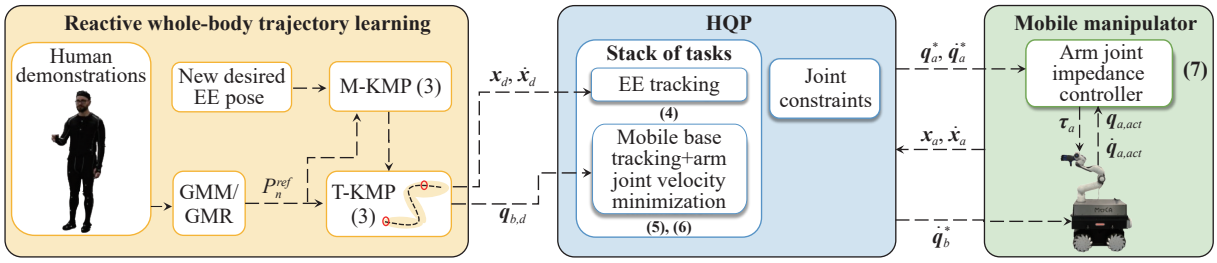

Fig. 1 Overall proposed reactive whole-body mobile manipulation framework (Colored figures are available in the online version at https://link.springer.com/journal/11633)

Table 1 Comparison with relevant prior works (here, both reference trajectory and velocity modulation refer to task space)

Paper	Feature			
	Reference trajectory	Trajectory modulation	Reactive adaptation	Optimal
[16]	√	√	×	√
[17]	√	×	√	√
[20]	×	×	√	√
[22]	×	×	√	√
Ours	√	√	√	√

1) An overall online reactive whole-body locomotion-integrated manipulation framework that adapts to the task's changes online;

2) A reactive task space whole-body reference trajectory learning and adaptation formulation with trajectory modulation capability based on human demonstrations;

3) A tailored HQP controller with deeper ablation investigation that transfers the updated task space whole-body trajectory to optimal joint-level commands for CMM.

The following sections of the paper are organized as follows. The overall approach and the details of the proposed methodology are illustrated in Section 2. The experiments and results are presented in Section 3, and Section 4 concludes the overall paper.

2 Methodology

2.1 Overall framework description

The proposed framework mainly includes three parts (see Fig. 1).

1) **Step 1. Reactive whole-body trajectory learning:** Reference trajectories are extracted from human demonstrations offline by Gaussian mixture model (GMM)/Gaussian mixture regression (GMR), where the human wrist and pelvis motion are mapped to the EE and mobile base of CMM, respectively. The following parts are implemented online to gain reactive planning and control for CMM. Specifically, the human-demonstrated spatial relationship between the EE pose of CMM and the mobile base pose is learned by multi-dimensional input KMP (i.e., M-KMP), where the EE pose and mobile base pose are set as input and output variables, respectively. Besides, the learned whole-body reference trajectories, including EE and mobile base movements, are encoded by the time-input KMP, namely T-KMP, where the time and trajectories are treated as input and output variables separately. Once the new desired EE pose is defined, for example, the target object pose is detected by a vision system in a pick-and-place task, M-KMP will calculate the corresponding mobile base pose. Then, the new desired point is inserted into the reference trajectory, and T-KMP updates and generates the whole-body reference trajectories.

2) **Step 2. HQP formulation:** The real-time generated whole-body trajectories are sent to the HQP controller, subjecting to joint constraints, where the EE tracking and mobile base tracking with arm joint velocity minimization are treated as primary and secondary tasks, respectively.

3) **Step 3. Joint-level controller:** The optimal and feasible joint velocities from HQP are sent to the arm low-level joint impedance controller and mobile base velocity controller.

2.2 Reactive whole-body trajectory learning and adaptation

2.2.1 Whole-body trajectory mapping and learning from human demonstrations

To acquire comprehensive whole-body motion patterns from human demonstrations, we establish a mapping between the human wrist/pelvis and the EE/mobile base of the CMM, respectively. Due to disparities in the geometric characteristics of the human arm and the mounting configuration of the robotic arm, it is unnecessary to individually map each joint of the human arm to the corresponding joint in the robotic arm. Consequently, we consider the human wrist pose and velocity $(\boldsymbol{x}_{wr}, \dot{\boldsymbol{x}}_{wr}) \in \mathbf{R}^6$, and pelvis pose $\boldsymbol{x}_{pe} \in \mathbf{R}^6$, in the world frame as the whole-body trajectories for CMMs. It is worth mentioning that this mapping method allows the learned trajectories to be applied to different CMMs.

In this study, we employ KMP[27] to encode the whole-body trajectory through human demonstrations. For the gathered M human demonstrations $\{\{\boldsymbol{s}_{n,m}, \boldsymbol{\xi}_{n,m}\}_{n=1}^{N_m}\}_{m=1}^{M}$, where each time step is selected as the in-

put variable $\boldsymbol{s}$ comprising human wrist pose and velocity, and pelvis pose as the output variables $\boldsymbol{\xi} = [\boldsymbol{x}_{wr}^{\mathrm{T}}\ \dot{\boldsymbol{x}}_{wr}^{\mathrm{T}}\ \boldsymbol{x}_{pe}^{\mathrm{T}}]^{\mathrm{T}}$ for each data point. In this paper, we call this time-driven KMP as T-KMP. Utilizing these demonstrations, a distribution $\hat{\boldsymbol{\xi}}_n|\boldsymbol{s}_n \sim \mathcal{N}(\hat{\boldsymbol{\mu}}_n, \hat{\boldsymbol{\Sigma}}_n)$ and the reference trajectory $\{\boldsymbol{s}_n, \hat{\boldsymbol{\mu}}_n, \hat{\boldsymbol{\Sigma}}_n\}_{n=1}^{N}$ can be derived through GMMs/GMR, where $\hat{\boldsymbol{\mu}}_n$ and $\hat{\boldsymbol{\Sigma}}_n$ denote the mean and covariance, respectively. Subsequently, the parametric model of KMP is expressed as

$$\boldsymbol{\xi}(\boldsymbol{s}) = \boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{s})\boldsymbol{w} \tag{1}$$

where $\boldsymbol{\Phi}(\boldsymbol{s}) = \boldsymbol{I}_{\mathcal{O}} \otimes \boldsymbol{\phi}(\boldsymbol{s}) \in \mathbf{R}^{B\mathcal{O}\times\mathcal{O}}$. $\boldsymbol{\phi}(\boldsymbol{s})$ represents the basis function with dimension of B, and $\boldsymbol{I}_{\mathcal{O}}$ is the identity matrix with dimension of $\mathcal{O}$. The weight vector is denoted as $\boldsymbol{w} \sim \mathcal{N}(\boldsymbol{\mu}_w, \boldsymbol{\Sigma}_w)$, where $\boldsymbol{\mu}_w$ and $\boldsymbol{\Sigma}_w$ are unknown. To determine these variables, KMP utilizes (2) to minimize the KL-divergence between the probabilistic trajectory generated by (1) and the reference trajectory $\mathcal{P}_n^{ref}$:

$$\sum_{n=1}^{N} KL(\mathcal{P}_n^{para}||\mathcal{P}_n^{ref}) \tag{2}$$

where $\mathcal{P}_n^{para} = \mathcal{N}(\boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{s}_n)\boldsymbol{\mu}_w, \boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{s}_n)\boldsymbol{\Sigma}_w\boldsymbol{\Phi}(\boldsymbol{s}_n))$ models the distribution of demonstration and $\mathcal{P}_n^{ref} = \mathcal{N}(\hat{\boldsymbol{\mu}}_n, \hat{\boldsymbol{\Sigma}}_n)$. Through the decomposition of the aforementioned objective function, the mean output corresponding to any input s^* is as follows:

$$\mathbb{E}(\boldsymbol{\xi}(\boldsymbol{s}^*)) = \boldsymbol{k}^*(\boldsymbol{K} + \lambda_1\boldsymbol{\Sigma})^{-1}\boldsymbol{\mu} \tag{3}$$

$$\mathbb{D}(\boldsymbol{\xi}(\boldsymbol{s}^*)) = \frac{N}{\lambda_2}(\boldsymbol{k}(\boldsymbol{s}^*, \boldsymbol{s}^*) - \boldsymbol{k}^*(\boldsymbol{K} + \lambda_2\boldsymbol{\Sigma})^{-1}\boldsymbol{k}^{*\mathrm{T}}) \tag{4}$$

where $\lambda_1 > 0$ and $\lambda_2 > 0$ represent the regularization factor. $\boldsymbol{k}^* \in \mathbf{R}^{\mathcal{O}\times N\mathcal{O}}$ is a $1 \times N$ block matrix, where the i-th column element is $\boldsymbol{k}(\boldsymbol{s}^*, \boldsymbol{s}_i)\boldsymbol{I}_{\mathcal{O}}$. $\boldsymbol{K} \in \mathbf{R}^{N\mathcal{O}\times N\mathcal{O}}$ denotes an $N \times N$ block matrix, where the i-th row and the j-th column item is $\boldsymbol{k}(\boldsymbol{s}_i, \boldsymbol{s}_j)\boldsymbol{I}_{\mathcal{O}}$. Besides, $\boldsymbol{\mu} = [\hat{\boldsymbol{\mu}}_1^{\mathrm{T}}\ \hat{\boldsymbol{\mu}}_2^{\mathrm{T}}\ \cdots\ \hat{\boldsymbol{\mu}}_N^{\mathrm{T}}]^{\mathrm{T}}$ and $\boldsymbol{\Sigma} = \mathrm{blockdiag}\{\hat{\boldsymbol{\Sigma}}_1, \hat{\boldsymbol{\Sigma}}_2, \cdots, \hat{\boldsymbol{\Sigma}}_N\}$.

2.2.2 Reactive trajectory adaptation

To repeat learned skills, the whole-body reference trajectories can be generated directly by (3) in Section 2.2.1. Furthermore, the new CMM whole-body trajectory should be adapted online in an unstructured environment according to the task requirements. For example, the CMM EE pose should be adapted to a new initial pose or the new position of the target object in the pick-and-place task (i.e., new desired points set $\{\overline{\boldsymbol{s}}_m, \overline{\boldsymbol{\mu}}_m, \overline{\boldsymbol{\Sigma}}_m\}_{m=1}^{M}$). Thanks to the trajectory modulation capacity of KMP, we can insert the new desired point set into the reference trajectory $\mathcal{P}_n^{ref}$, resulting in an $N + M$ new trajectory. Then, the T-KMP utilizes (3) to undertake the learning and generation of the new trajectory, encompassing all the specified desired points. The EE pose and velocity are easy to define based on the task requirements when we insert a new point (including EE pose/velocity and base pose), but finding the corresponding base pose is non-trivial. Nevertheless, the mobile base placement is essential to ensure the feasibility of the EE motion of CMM, as we discussed in Section 1.

To generate the corresponding mobile base pose with respect to a new desired EE pose of a new point, we propose a reactive whole-body motion learning and adaption approach for CMM, which can generate feasible whole-body reference trajectories based on human demonstrations, even out of the demonstration region. Since the sequence of the human pelvis can ensure a feasible wrist pose to achieve the task during the demonstration, the EE pose is set as the input variable ($\boldsymbol{s}' = \boldsymbol{x}_{wr}$), and the corresponding base pose is treated as the output variable ($\boldsymbol{\xi}' = \boldsymbol{x}_{pe}$), which is formulated as a multi-dimensional KMP (M-KMP in this paper). Then, this spatial relationship between EE pose and mobile base pose can be learned by (1) and (2). Once a new EE pose is given (i.e., $\boldsymbol{s}'^*$), the corresponding mobile base pose can be defined by (3).

Eventually, the overall reactive whole-body trajectory learning and adaptation approach (Algorithm 1, left module in Fig. 1) is as follows: The EE pose and velocity are defined by the updated task requirements. Then, the corresponding base pose is generated by M-KMP using the new EE pose ($\boldsymbol{s}'^*$) as input. After that, the overall new desired point set is inserted into the reference trajectory. Finally, the new whole-body reference trajectory is generated by T-KMP, which uses the current time ($\boldsymbol{s}^*$) as the input variable. It is worth noting that this algorithm is implemented online, and the whole-body trajectory is adapted whenever the task changes during the execution process.

Algorithm 1. Reactive whole-body trajectory generation
Input: $\mathcal{P}_n^{ref}, s'^*, s^*$
Output: Updated whole-body reference trajectory (T-KMP)
1) **if** $\boldsymbol{s}'^*$ (desired EE pose) changed **then**
2) Calculate $\boldsymbol{\xi}' = \boldsymbol{x}_{pe}$ of M-KMP by (3) using $\boldsymbol{s}'^*$
3) Define covariance $\overline{\boldsymbol{\Sigma}}$
4) Insert new desired point $(\overline{\boldsymbol{s}}_m, \overline{\boldsymbol{\mu}}_m, \overline{\boldsymbol{\Sigma}})$ into $\mathcal{P}_n^{ref}$
5) Update T-KMP
6) Generate trajectory by (3) with T-KMP from $\boldsymbol{s}^*$
7) **end if**

2.3 Hierarchical quadratic programming controller

We use a two-priority level HQP controller to transfer the learned whole-body reference trajectory to joint-level command for CMM (middle module in Fig. 1). This design ensures that the solution obtained at the first level is rigorously imposed on the subsequent second-level pri-

ority. Since the EE trajectories $\boldsymbol{x}_d, \dot{\boldsymbol{x}}_d \in \mathbf{R}^6$ (i.e., the $\boldsymbol{x}_{wr}$, $\dot{\boldsymbol{x}}_{wr}$ in Section 2.2) are the most important to achieve the tasks in general, the whole-body closed-loop inverse kinematics (CLIK) is set as the primary (i.e., the first-level priority) task, which can be expressed as

$$\min_{\dot{\boldsymbol{q}}} \|\boldsymbol{J}\dot{\boldsymbol{q}} - (\boldsymbol{K_v}(\dot{\boldsymbol{x}}_d - \dot{\boldsymbol{x}}_a) + \boldsymbol{K_p}(\boldsymbol{x}_d - \boldsymbol{x}_a))\|^2 \quad (5)$$

where $\boldsymbol{x}_a, \dot{\boldsymbol{x}}_a \in \mathbf{R}^6$ represent the actual EE pose and velocity, respectively. Additionally, $\boldsymbol{K_p}, \boldsymbol{K_v} \in \mathbf{R}^{6\times 6}$ refer to the positive-definite diagonal gain matrices responsible for ensuring the accuracy of position and velocity trajectory tracking. Moreover, $\dot{\boldsymbol{q}} \in \mathbf{R}^n$ represents the joint velocities, and $\boldsymbol{J} \in \mathbf{R}^{6\times n}$ denotes the whole-body Jacobian matrix of CMM. Specifically, $\dot{\boldsymbol{q}} = [\dot{\boldsymbol{q}}_b^{\mathrm{T}}\ \dot{\boldsymbol{q}}_a^{\mathrm{T}}]^{\mathrm{T}}$, with $n = n_b + n_a$, where $\dot{\boldsymbol{q}}_b = \dot{\boldsymbol{x}}_b \in \mathbf{R}^{n_b}$ and $\dot{\boldsymbol{q}}_a \in \mathbf{R}^{n_a}$ represent the joint velocities of the mobile base and robotic arm, respectively.

The second-level task is set to track the learned mobile base pose and minimize the joint velocity of the robotic arm, which is defined as

$$\min_{\boldsymbol{q}} \|\boldsymbol{H}_b\boldsymbol{q} - \boldsymbol{q}_{b,d}\|^2 + \|\boldsymbol{H}_a\boldsymbol{q} - \boldsymbol{q}_{a,init}\|^2 \quad (6)$$

which can be expressed as the function of the optimization variable:

$$\min_{\dot{\boldsymbol{q}}} \|\boldsymbol{H}_b\dot{\boldsymbol{q}}\Delta t - (\boldsymbol{q}_{b,d} - \boldsymbol{q}_b^*)\|^2 + \|\boldsymbol{H}_a\dot{\boldsymbol{q}}\Delta t - (\boldsymbol{q}_{a,init} - \boldsymbol{q}_a^*)\|^2 \quad (7)$$

where $\boldsymbol{H}_a, \boldsymbol{H}_b \in \mathbf{R}^n$ denote selection matrices. $\Delta t \in \mathbf{R}^+$ represents the control period, while $\boldsymbol{q}_b^* \in \mathbf{R}^{n_b}, \boldsymbol{q}_a^* \in \mathbf{R}^{n_a}$ correspond to the integrated optimal solutions derived by the HQP at the preceding time step, where $\boldsymbol{q}^* = [\boldsymbol{q}_b^{*\mathrm{T}}\ \boldsymbol{q}_a^{*\mathrm{T}}]^{\mathrm{T}}$. Moreover, $\boldsymbol{q}_{a,init} \in \mathbf{R}^{n_a}$ represents the initial configuration of the robotic arm. The first term in (7) is responsible for tracking the learned mobile base pose, while the second term minimizes the joint velocity of the robotic arm.

Besides, the constraints governing the HQP problem are established at position, velocity, and acceleration levels, adhering to the physical limitations of the CMM actuators. Specifically, by keeping the mobile base Cartesian coordinates unconstrained, the constraints are solely applied to the arm at the position level: $\boldsymbol{q}_{a,min} \le \boldsymbol{q}_a^* + \dot{\boldsymbol{q}}_a\ \Delta t \le \boldsymbol{q}_{a,max}$, where $\boldsymbol{q}_{a,min}, \boldsymbol{q}_{a,max} \in \mathbf{R}^{n_a}$ are the minimum and maximum mechanical limits of the robotic arm joints. Furthermore, the constraints of the velocity and acceleration level for the overall CMM are represented as $\dot{\boldsymbol{q}}_{min} \le \dot{\boldsymbol{q}} \le \dot{\boldsymbol{q}}_{max}$ and $\ddot{\boldsymbol{q}}_{min} \le \dfrac{\dot{\boldsymbol{q}} - \dot{\boldsymbol{q}}^*}{\Delta t} \le \ddot{\boldsymbol{q}}_{max}$, respectively, where $\dot{\boldsymbol{q}}^* \in \mathbf{R}^n$ denotes the optimal joint velocity identified at the previous time step. $\dot{\boldsymbol{q}}_{min}, \ddot{\boldsymbol{q}}_{min} \in \mathbf{R}^n$ and $\dot{\boldsymbol{q}}_{max}, \ddot{\boldsymbol{q}}_{max} \in \mathbf{R}^n$ represent the velocity and acceleration limits, respectively.

The optimal joint velocity $\dot{\boldsymbol{q}}_a^*$, obtained through the HQP controller, is compatible with direct integration with a conventional joint velocity controller designed for robotic arms that lack joint torque control functionality. Nevertheless, incorporating an additional layer comprising a joint impedance controller enhances safety in instances of unexpected contacts or impacts occurring anywhere along the kinematic chain (right module in Fig. 1). This joint impedance controller can be formulated as follows:

$$\boldsymbol{\tau}_a = \boldsymbol{K}_{\boldsymbol{q_{a,d}}}(\dot{\boldsymbol{q}}_a^* - \dot{\boldsymbol{q}}_{a,act}) + \boldsymbol{K}_{\boldsymbol{q_{a,p}}}(\boldsymbol{q}_a^* - \boldsymbol{q}_{a,act}) + \boldsymbol{g}_a \quad (8)$$

where $\boldsymbol{q}_a^*$ and $\dot{\boldsymbol{q}}_a^* \in \mathbf{R}^{n_a}$ denote the desired positions and velocities of the robotic arm′s optimal joints, respectively, derived from the HQP. Simultaneously, $\boldsymbol{q}_{a,act}$ and $\dot{\boldsymbol{q}}_{a,act} \in \mathbf{R}^{n_a}$ represent the actual joint positions and velocities. The matrices $\boldsymbol{K}_{\boldsymbol{q_{a,p}}} \in \mathbf{R}^{n_a\times n_a}$ and $\boldsymbol{K}_{\boldsymbol{q_{a,d}}} \in \mathbf{R}^{n_a\times n_a}$ denote the positive definite joint stiffness and damping matrices, respectively, while $\boldsymbol{g_a} \in \mathbf{R}^{n_a}$ corresponds to the gravity compensation vector. Regarding the mobile base, the optimal joint velocity $\dot{\boldsymbol{q}}_b^* \in \mathbf{R}^{n_b}$ is directly conveyed to the default controller, given that the majority of mobile platforms primarily support velocity control interfaces.

2.4 Technical details

As discussed in Section 2.2, the covariance $\overline{\boldsymbol{\Sigma}}$ of a new desired point $(\overline{\boldsymbol{s}}, \overline{\boldsymbol{\mu}}, \overline{\boldsymbol{\Sigma}})$ should be given when we insert them into the reference trajectory of T-KMP. Generally, a small value (close to zero) ensures that the new trajectory passes all the new desired points. However, this can be adjusted according to the task requirements. For example, a close zero value of covariance can be used for the EE pose and base pose to guarantee the feasibility of the EE motion of CMM. A high value, for example, ten thousand, can be set for the EE velocity covariance simultaneously if there is no specific requirement for the velocity profile at that point. In this way, there is no need to find a particular value for the EE velocity when the requirements for the new points are just related to the EE pose and mobile base pose, and the learned whole-body trajectory can be flexibly adapted to the task changes.

3 Experiments results and discussions

The proposed approach was evaluated using the mobile collaborative robotic assistant (MOCA) platform, as detailed in [8], in locomotion-integrated reactive pick-and-place and reactive reaching tasks. MOCA is a mobile manipulator featuring a 7-degree-of-freedom (DoFs) Franka Emika Panda robotic arm with joint torque control capabilities mounted atop a velocity-controlled 3-DoFs Robotnik SUMMIT-XL STEEL mobile base. Besides, the Franka default gripper and Pisa/IIT SoftHand were used

for the two tasks. Initially, human demonstrations were conducted, as depicted in Fig. 2, where the human demonstrations were collected and used to learn the T-KMP and M-KMP. After that, an ablation experiment in simulation for the proposed HQP controller was conducted to show why the mobile base reference trajectory is necessary to keep the human-demonstrated whole-body movement behavior. Finally, the reactive pick-and-place and reactive reaching experiments with the real MOCA were conducted. Specifically, the target object was randomly moved out of the demonstration region. MOCA was required to pick, replace, or reach it based on the online reactively adapted trajectories using the proposed framework. We focus on the whole-body coordination movement generation between the arm and mobile base of CMM instead of the grasping in all experiments.

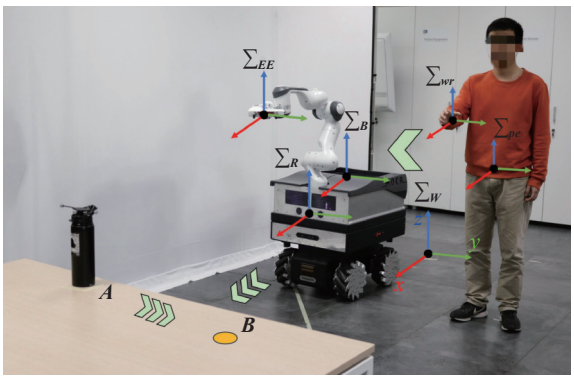

Fig. 2 Experimental settings. The subject was required to naturally pick the bottle at point A without stopping and place it at point B later. All the movements were described in the world frame $\boldsymbol{\Sigma_W}$ if without specification. The corresponding frames between the subject and MOCA are also depicted. (Colored figures are available in the online version at https://link.springer.com/journal/11633)

3.1 Human demonstrations collection

A healthy human subject was asked to perform the pick-and-place task to collect demonstrations. The overall setup is shown in Fig. 2. Demonstrations were collected using the following protocol[1]: The subject was instructed to don a Lycra suit adorned with markers, enabling real-time tracking of whole-body motion through the OptiTrack system for data collection purposes. The sensor data acquisition and synchronization processes were coordinated through the robot operating system (ROS) environment, operating at a frequency of 100 Hz. As described in Fig. 2, the bottle was placed at $x = 3.687\,\text{m}$, $y = 0.423\,\text{m}$, point A on a table with a height of 0.723 m. If not specified, all distances were described in the world frame ($\boldsymbol{\Sigma_W}$). Using the right hand, the subject was required to naturally pick up this bottle without stopping and place it at ($x = 3.73\,\text{m}$, $y = 0.97\,\text{m}$, point B). Eventually, 5 demonstrations were applied to learn the whole-body pick-and-place task.

[1] The research protocol employed in this study received formal approval from the ethics committee of Azienda Sanitaria Locale (ASL) Genovese N.3, as documented under protocol IIT_HRII_ERGOLEAN 156/2 020.

3.2 Experimental setup

3.2.1 Parameters setting

In this study, MOCA was employed to execute all the experiments, leveraging learned human skills. Specifically, the pose elements $\boldsymbol{q}_{b,d} = [x_{pe}\ y_{pe}\ \varphi_{pe}]^{\mathrm{T}} = [x_{b,d}\ y_{b,d}\ \varphi_{b,d}]^{\mathrm{T}}$ corresponding to the pelvis′ pose ($\boldsymbol{x}_{pe}$) were acquired, considering that MOCA′s mobile base possesses 3 DoFs. Moreover, only the position and translational speed of the wrist′s pose and velocity ($(\boldsymbol{x}_{wr}\ \dot{\boldsymbol{x}}_{wr}) = (\boldsymbol{x}_d\ \dot{\boldsymbol{x}}_d)$) were utilized, and the EE′s desired orientation was fixed across all experimental trials. Additionally, a mapping between the pelvis frame ($\boldsymbol{\Sigma_{pe}}$) and the robotic arm base frame ($\boldsymbol{\Sigma_R}$) was established. Given that the mobile base commands are referenced in $\boldsymbol{\Sigma_B}$, compensation for the relative transformation between $\boldsymbol{\Sigma_R}$ and $\boldsymbol{\Sigma_B}$ was necessary, with an offset of 0.387 5 m observed in the x direction for MOCA.

The regularization factors were set for the parameters as $\lambda_1 = \lambda_2 = 1$ for both T-KMP and M-KMP. Please note that T-KMP and M-KMP are the same from the algorithm point of view, which have different input and output variables in our application. Besides, we set the corresponding value as 1×10^{-6} in the covariance matrix $\overline{\boldsymbol{\Sigma}}$ of the new desired point when we want the new trajectory to pass by the desired position. While the related value is set as 1×10^{6} when we do not have requirements for the velocity profile, as presented in Section 2.4. Besides, to get accurate tracking performance of EE position and velocity, the control gains of CLIK (5) were chosen as follows: $\boldsymbol{K_p} = \text{diag}\{4, 4, 4, 4, 4, 4\}$ and $\boldsymbol{K_v} = \text{diag}\{1, 1, 1, 1, 1, 1\}$. The selection matrixes in the secondary task were set as $\boldsymbol{H}_b = \text{diag}\{1, 1, 1, 0, 0, 0, 0, 0, 0, 0\}$ and $\boldsymbol{H}_a = \text{diag}\{0, 0, 0, 1, 1, 1, 1, 1, 1, 1\}$, resulting in the learned mobile base pose tracking (i.e., the first term in (7)) only worked on the optimal base velocity $\boldsymbol{q}_b^*$. The optimal robotic arm velocity $\boldsymbol{q}_a^*$ was determined by the second term in (7) to minimize the joint velocity. Furthermore, to generate a suitable trade-off between tracking accuracy and compliance, we use $\boldsymbol{K}_{\boldsymbol{q_{a,p}}} = \text{diag}\{20, 20, 20, 20, 20, 20, 20\}$ and $\boldsymbol{K}_{\boldsymbol{q_{a,v}}} = \text{diag}\{8, 8, 8, 8, 8, 8, 8\}$ for arm joint impedance controller (8). The control period for the following simulation and actual robot experiment was $\Delta t = 0.001\,\text{s}$.

3.2.2 Implementation details

We use the ALGLIB QP-BLEIC solver (http://www.alglib.net/optimization/quadraticprogramming.php) for the HQP formulation by C++ in ROS on Ubuntu 20.04.

All experiments were run on a computer with an Intel Core i7-4790S 3.2 GHz × 8-cores CPU and 16 GB RAM. Based on this setup, we successfully ran the HQP formulation in real-time with a frequency of 1 KHz. Furthermore, a mark was stuck on the top of the bottle, which was detected by the OptiTrack system to get the real-time 3D position. The overall ROS node structure of the two reactive whole-body mobile manipulation experiments was described in Fig. 3, which follows the process of Algorithm 1. Four nodes were working parallelly after starting the experiment. Once the bottle was moved to a new position, node 1 sent the new position (i.e., new EE position) to node 2, M-KMP, which predicted the corresponding mobile base position. Node 3, T-KMP, updated the reference trajectory immediately after receiving the new desired point. Node 4, HQP controller generated the joint-level trajectory for MOCA by commanding the latest reference trajectory. It is worth mentioning that there were two independent timers of T-KMP and HQP. Node 4 (timer 2) was still executing with the previous time instant reference trajectory when node 3 was updating with timer 1 stopping, resulting in timer 2 being possibly larger than timer 1. It is essential to stop timer 1 when node 1 is updating to avoid time jumps since T-KMP is time-driven, and a time jump may generate trajectory jumps.

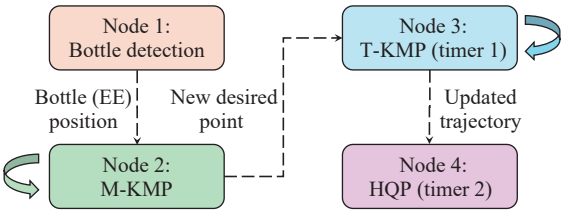

Fig. 3 ROS nodes structure for the two reactive whole-body mobile manipulation tasks

3.3 Experimental results and analysis

3.3.1 Ablation experiment of HQP

In the proposed HQP formulation, we have two items in the secondary task (7), where the first tracks the learned base trajectory and the second tries to minimize the robotic arm joint velocity. Since the proposed HQP without the base tracking item could generate joint-level commands, one may argue that the learned base pose tracking is unnecessary to imitate the human-demonstrated whole-body tasks. Here, we conducted a simulated ablation experiment to interstage the effect of this item. Specifically, MOCA was required to repeat the whole-body trajectory generated by T-KMP of the pick-and-place task without/with the learned mobile base pose tracking item in (7). For simplicity, we used with and without to express these two settings in the following part. All the simulations were implemented in Gazebo. Besides, since we focus on the effect of the HQP on distributing the EE motion of CMM to the arm and mobile base, bottle grasping was not considered here, and the reference trajectories were planned offline.

The snapshots of the without ($Wo1 - Wo5$) and with ($W1 - W5$) experiments were expressed in the first and second rows of Fig. 4, respectively. The overall movements of these two experiments can be found in Fig. 5. Specifically, the actual and desired EE position and velocity in the world frame ($\boldsymbol{\Sigma_W}$) were separately in the first and fourth rows. The second row showed the EE position in the robotic arm base frame ($\boldsymbol{\Sigma_R}$), and the learned mobile base trajectory with the actual ones could be found in the third row. According to the results, the human demonstrated trajectory was in a unified loco-manipulation mode, with a duration of 26 s. To measure the tracking performances of the two-level priority tasks of HQP, the root mean squared error (RMSE) was used as the metric, which can be written as

$$RMSE_{\boldsymbol{x}_{des}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{x}_{des,i} - \boldsymbol{x}_{act,i})^2} \tag{9}$$

where N is the total number of samples. Besides, the $\boldsymbol{x}_{des}$ and $\boldsymbol{x}_{act}$ represent the general desired and actual variables, respectively.

For the EE position and velocity tracking in the world frame, both with and without settings showed similar good performances since the CLIK was set as the primary task of HQP. However, the EE motion was distributed differently for the robotic arm and the mobile base of the two cases. In general, for the with case, MOCA imitated the whole-body motion of the human demonstrations. Specifically, the mobile base took the domination role, moving toward the object along the x direction. Then, the robotic arm started to move in all three directions because it was close to the object. Meanwhile, the mobile base rotated and still moved toward the final point. Obviously, according to Table 2, the mobile base tracking was not as good as the EE position tracking, which was expected for the secondary task in HQP. However, the mobile base was never rotated for the without case. Furthermore, the robotic arm basically did not move along x and y directions due to the minimizing joint velocity item in (7). Since the mobile base of MOCA can not move along the z direction, the robotic arm movements in this direction were similar for both cases. Although the without case could generate joint-level commands for MOCA, the human-demonstrated EE motion distribution relationship (i.e., the coordinated motion) between the robotic arm and mobile base was clearly changed. Therefore, the learned trajectory tracking item of the mobile base in the proposed HQP is essential for imitating the human-demonstrated whole-body behavior.

Fig. 4 Snapshots of the ablation experiments in Gazebo. The first row: the proposed approach is without the learned mobile base pose tracking item in the secondary task of HQP. The second row: the overall proposed approach (i.e., with the mobile base tracking item in HQP).

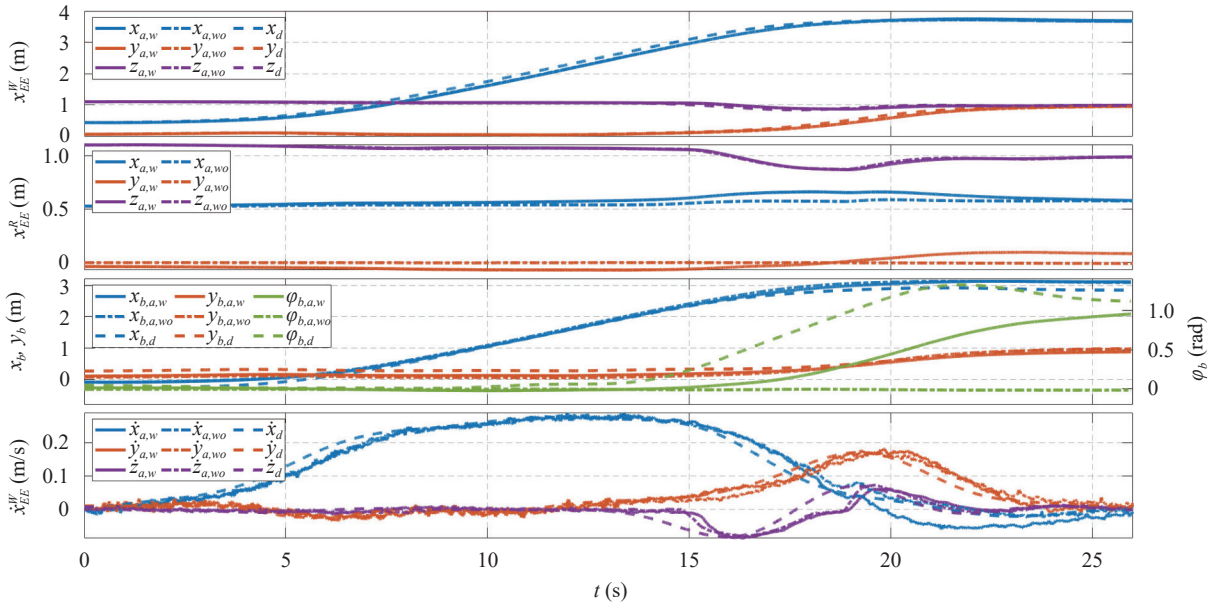

Fig. 5 Results of simulated ablation experiments without/with the base pose tracking item in the secondary task of HQP. From top to bottom, the first row: EE desired position $\boldsymbol{x}_d \in \mathbf{R}^3$ and actual position $\boldsymbol{x}_{a,wo} \in \mathbf{R}^3$ for without case ($\boldsymbol{x}_{a,w} \in \mathbf{R}^3$ for with case) in the world frame $\boldsymbol{\Sigma_W}$; the second row: EE actual position in the robotic arm base frame $\boldsymbol{\Sigma_R}$. $\boldsymbol{x}_{a,wo} \in \mathbf{R}^3$ and $\boldsymbol{x}_{a,w} \in \mathbf{R}^3$ represent for the without and with case, respectively; the third row: desired mobile base learned pose $\boldsymbol{q}_{b,d} \in \mathbf{R}^3$, and actual pose $\boldsymbol{q}_{b,a,wo} \in \mathbf{R}^3$ for the without case ($\boldsymbol{q}_{b,a,w} \in \mathbf{R}^3$ for with case); the fourth row: EE desired velocity $\dot{\boldsymbol{x}}_d \in \mathbf{R}^3$ and actual velocity $\dot{\boldsymbol{x}}_{a,wo} \in \mathbf{R}^3$ for the without case ($\dot{\boldsymbol{x}}_{a,w} \in \mathbf{R}^3$ for with case). (Colored figures are available in the online version at https://link.springer.com/journal/11633)

Table 2 RMSE results of the ablation experiment

Experiment	With	Without
$RMSE_{EEPos}$ (m)	(0.085, 0.027, 0.019)	(0.085, 0.033, 0.018)
$RMSE_{EEVel}$ (m/s)	(0.023, 0.018, 0.015)	(0.019, 0.019, 0.015)
$RMSE_{base}$ (m, m, rad)	(0.142, 0.145, 0.260)	–

3.3.2 Reactive pick-and-place experiment

We conducted a reactive pick-and-place task to evaluate the proposed combined learning and optimization framework with actual MOCA. Specifically, the bottle (target object) was randomly moved twice when MOCA executed the task. We hope the proposed method can reactively generate the whole-body trajectory to adapt to online changes. Besides, the bottle, with a marker on top of it, was tracked by Optitrack in real-time. The Franka gripper was preconfigured to close/open at defined positions. The results were in Fig. 6, with the snapshots (from $A-F$) on top and movement plots at the bottom. At the start, the bottle was placed in the same position as the human demonstration process (snapshot A). Then, a human randomly moved the bottle from the demonstration region to another position. Once the new bottle position (i.e., the latest desired EE position) was updated, the M-KMP gave the corresponding base pose. For the first update, the desired EE position and related base pose were $\boldsymbol{x}_{d,fir} = [1.799\ 0.699\ 1.017]^{\mathrm{T}}$ m and $\boldsymbol{q}_{b,d,fir} = [1.075\,\mathrm{m}\ 0.689\,\mathrm{m}\ -0.032\,\mathrm{rad}]^{\mathrm{T}}$ (magenta circles marked in Fig. 6), respectively. It is clear that the new EE position and the generated base position were outside the demonstrations region, especially in the y direction. Here, we did not have requirements for the EE velocity with the corresponding covariance value of 1×10^6 since the

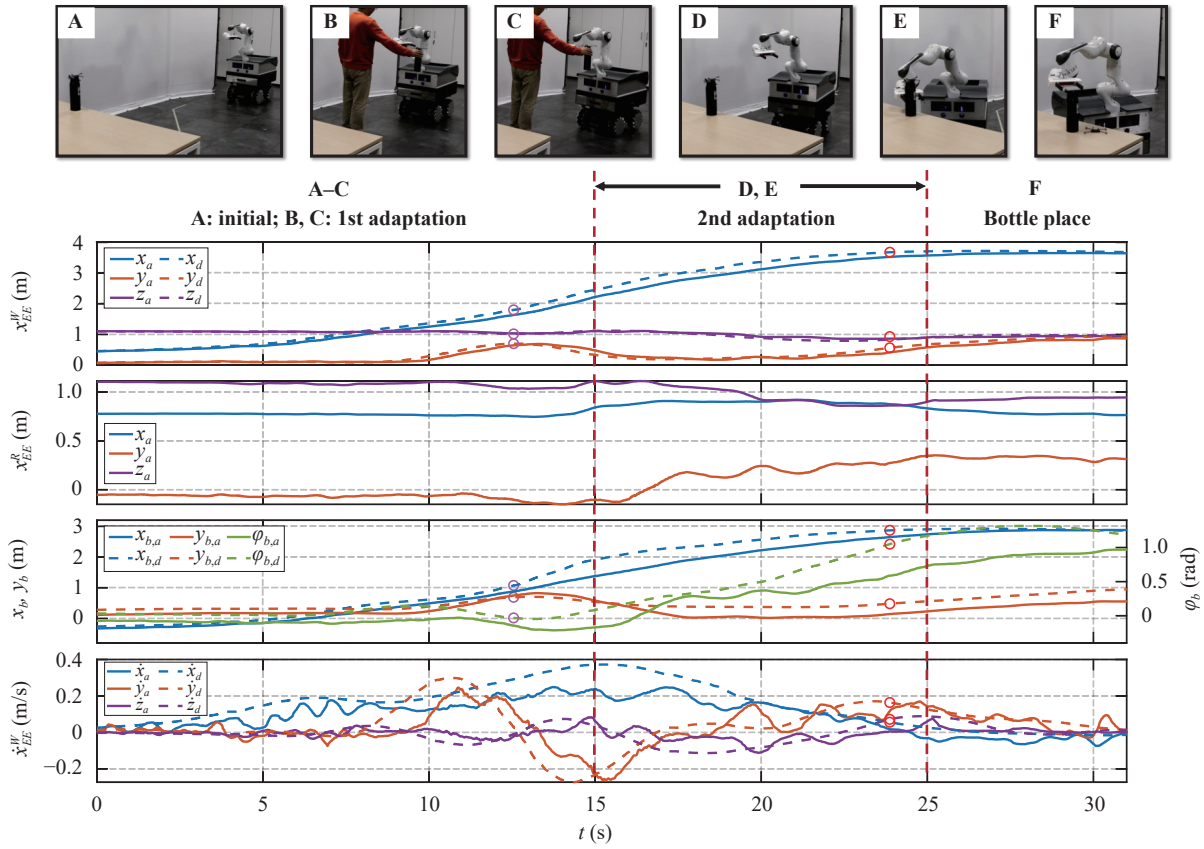

Fig. 6 Results of the reactive pick-and-place experiment. Top: snapshots during the task process. The content of the current figure is similar to Fig. 5. Around $t = 12.5$, MOCA reached the first desired bottle position marked by magenta circles (snapshot C), where the desired EE position (i.e., the position of bottle tracked by Optitrack) and the corresponding base pose generated by M-KMP were $\boldsymbol{x}_{d,fir} = [1.799\ 0.699\ 1.017]^{\mathrm{T}}$ m and $\boldsymbol{q}_{b,d,fir} = [1.075\,\mathrm{m}\ 0.689\,\mathrm{m}\ -0.032\,\mathrm{rad}]^{\mathrm{T}}$, respectively; around $t = 23.9$ s, MOCA grasped the bottle at the second desired position highlighted in red circles (E): $\boldsymbol{x}_{d,sec} = [3.678\ 0.563\ 0.930]^{\mathrm{T}}$ m and $\boldsymbol{q}_{b,d,sec} = [2.872\,\mathrm{m}\ 0.480\,\mathrm{m}\ 1.049\,\mathrm{rad}]^{\mathrm{T}}$. (Colored figures are available in the online version at https://link.springer.com/journal/11633)

bottle would not be grasped here. On the contrary, the EE position and base pose-related covariance were set as 1×10^{-6} to ensure the new trajectory will pass by the desired point, as we discussed in Section 2.4. Then, the new via point was inserted into the reference trajectory, and the T-KMP was updated online. Furthermore, the mobile base contributed significantly and moved to the bottle position (B). Then, the arm started to move when it was close to the bottle, and the desired position was achieved (C). However, we did not want MOCA to grasp the bottle here and randomly move it on the table. For the second update, the desired EE position and related base pose were $\boldsymbol{x}_{d,sec} = [3.678\ 0.563\ 0.930]^{\mathrm{T}}$ m and $\boldsymbol{q}_{b,d,sec} = [2.872\,\mathrm{m}\ 0.480\,\mathrm{m}\ 1.049\,\mathrm{rad}]^{\mathrm{T}}$ (red circles marked in Fig. 6), respectively. This time we set the desired EE velocity as human demonstrated contact velocity ($\dot{\boldsymbol{x}}_d = [0.056\ 0.164\ 0.072]\,\mathrm{m/s}^{\mathrm{T}}$, red circle in the fourth row in Fig. 6) because MOCA was required to grasp the bottle at this point. The whole-body trajectory was updated again as before, and the mobile base mainly moved toward the new bottle position (D). Then, the base and arm moved together and grasped the bottle (E). Finally, the bottle was placed at the desired point (F). The RMSE of this experiment is listed in Table 3.

Table 3 RMSE results of two reactive experiments

Experiment	Reactive pick-and-place	Reactive reaching
$RMSE_{EEPos}$ (m)	(0.137, 0.067, 0.032)	(0.139, 0.086, 0.044)
$RMSE_{EEVel}$ (m/s)	(0.073, 0.064, 0.037)	–
$RMSE_{base}$ (m, m, rad)	(0.224, 0.220, 0.227)	(0.192, 0.202, 0.197)

Although the bottle was randomly moved twice, even out of the demonstration region, the learned whole-body trajectory reactively adapted online based on the proposed method, achieving a successful pick-and-place task. Compared with Fig. 5, the reactively generated whole-body trajectory differed from the demonstrations, especially around the two inserted new points. While the overall shape of the reference trajectories (i.e., the desired EE and base movements) was still similar, indicating that the EE motion distribution relationship between arm and base was reserved because of the proposed learn-

ing algorithm. Besides, the overall motion duration was 31 s of this reactive pick-and-place task, which was longer than the ablation experiment in Section 3.3.1 since the reference trajectory was generalized online for both reactive experiments according to real-time task changes. We thus generated smooth reference trajectories by avoiding time discontinuities, as detailed in Section 3.2.2.

3.3.3 Reactive reaching experiment

To further evaluate the proposed framework's whole-body trajectory online adaptation capacity, we conducted a reactive reaching experiment based on previous demonstrations. Specifically, a human moved the bottle randomly four times during the task execution, and MOCA was required to reach the position autonomously. We did not have the reaching velocity requirements here, and the Pisa/IIT SoftHand was used.

The results of the reactive reaching experiment can be found in Fig. 7, where the snapshots $A-I$ and corresponding data plot are presented from top to bottom, respectively. Starting from the initial setting (A), following the online adapted whole-body trajectory, MOCA approached (B) and reached (C) the first new desired position (magenta circles in Fig. 7) around $t = 6.8$ s. Then, according to the updated bottle position around $t = 15.5$ s (red circles in Fig. 7), MOCA achieved there gradually (D, E). After that, MOCA reached the third (F, G, and cyan circles in Fig. 7) and the fourth (H, I, and brown circles in Fig. 7) desired points, subsequently around $t = 23.5$ s and $t = 31.7$ s. Therefore, thanks to the proposed framework, MOCA successfully finished the reactive reaching task. As for the tracking performance, the RMSE results were presented in the fifth column of

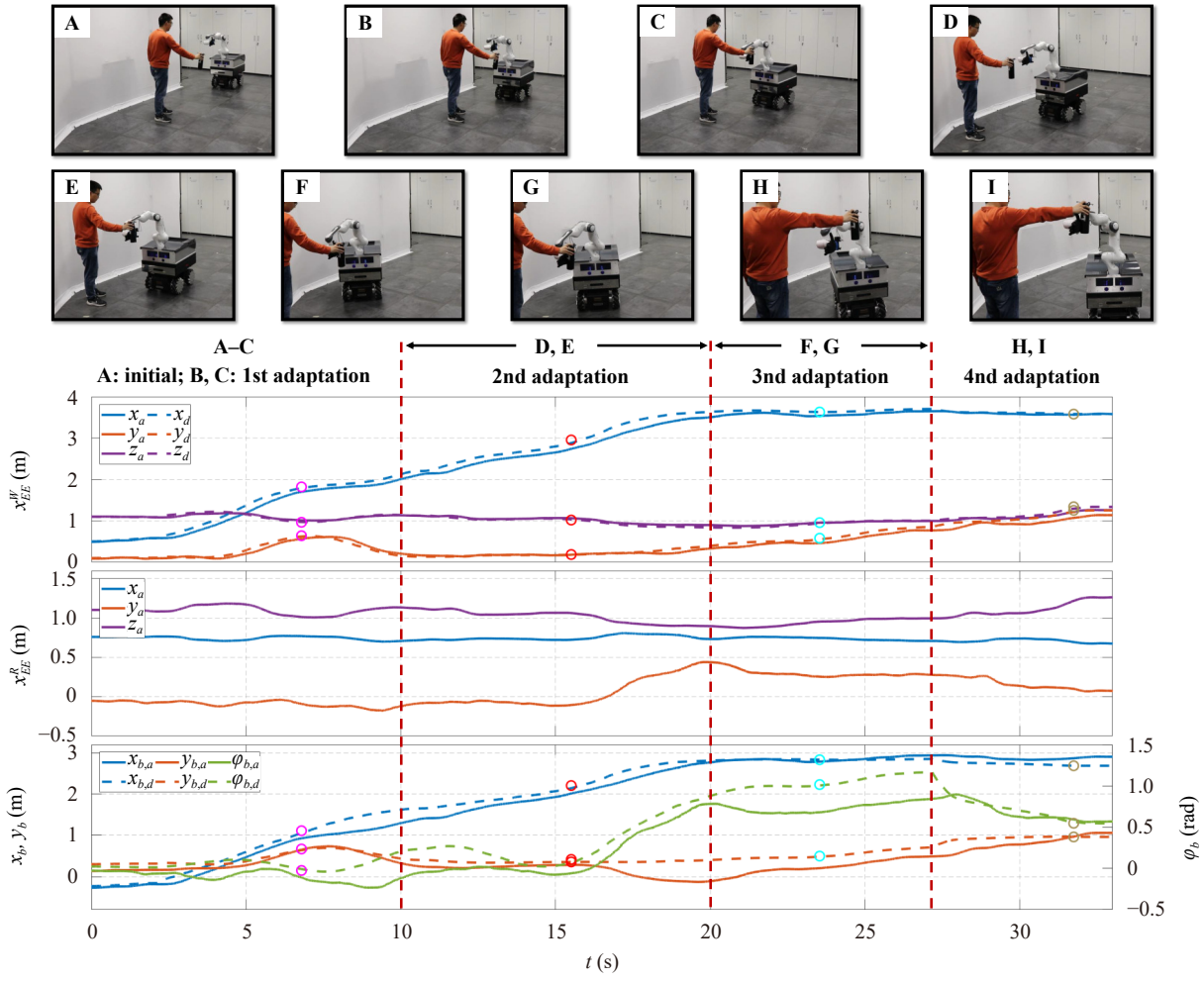

Fig. 7 Results of the reactive reaching experiment. Around $t = 6.8$, the first desired EE position and the corresponding base pose (snapshot C, marked by magenta circles) were $\boldsymbol{x}'_{d,fir} = [1.819\ 0.639\ 0.964]^{\mathrm{T}}$ m and $\boldsymbol{q}'_{b,d,fir} = [1.114\,\mathrm{m}\ 0.671\,\mathrm{m}\ -0.025\,\mathrm{rad}]^{\mathrm{T}}$, respectively; around $t = 15.5$ s, the second desired EE position and the corresponding base pose (E, marked by red circles) were $\boldsymbol{x}'_{d,sec} = [2.954\ 0.185\ 1.018]^{\mathrm{T}}$ m and $\boldsymbol{q}'_{b,d,sec} = [2.209\,\mathrm{m}\ 0.357\,\mathrm{m}\ 0.109\,\mathrm{rad}]^{\mathrm{T}}$, respectively; around $t = 23.5$ s, the third desired EE position and the corresponding base pose (G, marked by cyan circles) were $\boldsymbol{x}'_{d,thi} = [3.629\ 0.579\ 0.955]^{\mathrm{T}}$ m and $\boldsymbol{q}'_{b,d,thi} = [2.835\,\mathrm{m}\ 0.499\,\mathrm{m}\ 1.016\,\mathrm{rad}]^{\mathrm{T}}$, respectively; around $t = 31.7$ s, the fourth desired EE position and the corresponding base pose (I, marked by bronze circles) were $\boldsymbol{x}'_{d,fou} = [3.579\ 1.252\ 1.339]^{\mathrm{T}}$ m and $\boldsymbol{q}'_{b,d,fou} = [2.689\,\mathrm{m}\ 0.963\,\mathrm{m}\ 0.547\,\mathrm{rad}]^{\mathrm{T}}$, respectively. (Colored figures are available in the online version at https://link.springer.com/journal/11633)

Table 3, which were comparable to the reactive pick-and-place ones. As expected, the EE tracking outperformed the base tracking due to the hierarchical design of the stack of tasks (SoT). Besides, the overall period of this experiment was 33 s, which is longer than the reactive pick-and-place because of two more updates. Thanks to the proposed learning algorithm, the EE motion distribution relationship between the arm and base was kept, resulting in a similar trajectory shape compared with Figs. 5 and 6. However, MOCA's final state (I in 7) was clearly different from the demonstrated one ($W5$ in 5 and $F5$ in 6). This is because the fourth desired bottle position was far from the demonstration region, generating a corresponding desired mobile base pose different from the original one.

3.4 Discussions

The without case of the ablation experiments was a baseline compared with our overall proposed approach (i.e., with case) with quantitative results presented in Table 2. According to the ablation experiment results, only the with case successfully reserved the human-demonstrated EE motion distribution relationship (i.e., coordinated whole-body motion) between the mobile base and the robotic arm. Thanks to the M-KMP, the suitable base placement can be found based on human demonstrations when given a new EE pose. Different from the QP and MPC approaches for CMM in [20, 22] without long-term planning capacity, the proposed T-KMP can generate the long-term whole-body reference trajectory based on human demonstrations. Furthermore, thanks to the trajectory modulation capacity of the learning algorithm, the T-KMP adapts the learned whole-body reference trajectories online by inserting the new desired via-point with velocity requirements from M-KMP, which is challenging for previous research[17, 20, 22]. Combined with the tailored HQP formulation, the proposed learning and optimization framework reactively adapts the learned locomotion-integrated manipulation skills to the online task changes, keeping the human-demonstrated EE motion distribution relationship as much as possible. Compared with our previous work[16], the proposed approach in this paper can not only transfer the human mobile manipulation skills to CMM with different geometry as before but also reactively adapts the learned whole-body reference trajectory to task changes, even outside the demonstration region. Meanwhile, a deeper ablation investigation was conducted for the HQP formulation. In general, our proposed reactive whole-body locomotion-integrated manipulation framework distinguishes itself from existing approaches through the seamless integration of learning and optimization processes. This integration allows for enhanced reactivity, enabling real-time responses to dynamic task changes, and improved adaptability. Unlike traditional methods that treat learning and optimization as discrete stages, our framework leverages the strengths of both paradigms to achieve a more cohesive and efficient motion planning and optimal control strategy. This results in CMM that can mimic human-like adaptive behaviors, significantly improving task execution and interaction in dynamic environments.

In the current framework, we assume that the robotic arm of MOCA has a good initial configuration (i.e., with high manipulability). In this way, the second item of the HQP formulation (7) minimizes the joint velocity, resulting in MOCA staying in the good initial joint angles. For a more general setup, the manipulability optimization can also be added to the HQP formulation, which keeps the robotic arm in a reasonable configuration during the loco-manipulation tasks.

Although we showed the trajectory was updated 2 and 4 times in the two tasks separately, it can always be adapted online as long as the timer 1 of the T-KMP is still running. Currently, we validated the proposed framework primarily based on the dynamic pick-and-place task. Other loco-manipulation tasks under more complex environments can also be demonstrated and validated by applying our proposed framework. However, now the object (i.e., the bottle) was detected by the commercial OptiTrack version system with a limited workspace, and the dynamic moving obstacles were not considered. CMM's onboard version sensors (for example, RGB-D camera and LiDAR) can be used, which may increase the overall computation load. In this case, it may be challenging to implement our proposed approach in an online manner in environments requiring rapid decision-making or operating with limited computational resources. To address these issues, more details are presented in the future research discussion in Section 4.

4 Conclusions

This paper aimed to address the issue of reactive planning and control for CMM based on human demonstrations. In particular, the whole-body demonstrated reference trajectories were encoded by T-KMP, and the relationship between EE pose and the related mobile base pose was learned by M-KMP. Therefore, once the desired EE poses changed due to task requirements, the M-KMP updated first and generated the corresponding feasible base pose. Then, the T-KMP updated with inserted new points and generated the whole-body trajectory. The whole-body reference trajectories were sent to the HQP to generate the joint-level command for CMM, where the base pose tracking was essential to imitate the whole-body motion of human locomotion-integrate manipulation behavior. The successful reactive pick-and-place and reactive reaching tasks showed that the proposed ap-

proach could learn and adapt human whole-body behavior to CMM online, even outside the demonstration region.

The OptiTrack system tracked the bottle in the current research, which should be inside the workspace of this visual tracking device. Future work will include mounting an RGB-D camera on MOCA to detect the target object, which can make the learned skills used in larger space. Besides, although the whole-body reference trajectory was generalized online, obstacle avoidance was not considered. Therefore, dynamic obstacle avoidance could be another direction of further research by further exploring the capacity of KMP[28], ensuring CMMs work in more complex unstructured environments. However, the onboard perception will increase the overall computation load. To ensure the real-time perception and online learned skills generalization and execution for CMM, algorithmic optimizations and the utilization of dedicated hardware accelerators should also be considered in the future. Based on the enhanced perception and dynamic obstacle avoidance capacity, another interesting future research direction is to explore multi-object, multi-step loco-manipulation tasks with CMMs, further validating the generality of the proposed framework.

List of abbreviations

CLIK:Closed-loop inverse kinematics
CMM: Collaborative mobile manipulator
DoF: Degree of freedom
EE: End-effector
GMM: Gaussian mixture model
GMR: Gaussian mixture regression
HQP: Hierarchical quadratic programming
IRM: Inverse reachability map
KMP: Kernelized movement primitive
M-KMP: Multi-dimensional kernelized movement primitive
MOCA: Mobile collaborative robotic assistant
MPC: Model predictive control
QP: Quadratic programming
RMSE: Root mean squared error
SoT: Stack of tasks
T-KMP: Time-input kernelized movement primitive

Acknowledgements

This work was supported by the European Research Council's (ERC) starting grant Ergo-Lean (No. GA 850932). Open access funding provided by The Chinese University of Hong Kong, China.

Declarations of conflict of interest

The authors declared that they have no conflicts of interest to this work.

Open Access

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made.

The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

[1] A. Ajoudani, A. M. Zanchettin, S. Ivaldi, A. Albu-Schäffer, K. Kosuge, O. Khatib. Progress and prospects of the human-robot collaboration. *Autonomous Robots*, vol. 42, no. 5, pp. 957–975, 2018. DOI: 10.1007/s10514-017-9677-2.

[2] Y. Wu, P. Balatti, M. Lorenzini, F. Zhao, W. Kim, A. Ajoudani. A teleoperation interface for loco-manipulation control of mobile collaborative robotic assistant. *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3593–3600, 2019. DOI: 10.1109/LRA.2019.2928757.

[3] E. Lamon, M. Leonori, W. Kim, A. Ajoudani. Towards an intelligent collaborative robotic system for mixed case palletizing. In *Proceedings of IEEE International Conference on Robotics and Automation*, Paris, France, pp. 9128–9134, 2020. DOI: 10.1109/ICRA40945.2020.9196850.

[4] T. Teng, X. Wu, Z. Li, D. Caldwell, F. Chen. Robust shared control with stable contact servoing for enhanced object transportation by telerobotic bimanual mobile manipulators. *Journal of Field Robotics*, published online. DOI: 10.1002/rob.22464.

[5] X. Liu, S. S. Ge, F. Zhao, X. Mei. Optimized interaction control for robot manipulator interacting with flexible environment. *IEEE/ASME Transactions on Mechatronics*, vol. 26, no. 6, pp. 2888–2898, 2021. DOI: 10.1109/TMECH.2020.3047919.

[6] C. Pezzato, C. H. Corbato, S. Bonhof, M. Wisse. Active inference and behavior trees for reactive action planning and execution in robotics. *IEEE Transactions on Robotics*, vol. 39, no. 2, pp. 1050–1069, 2023. DOI: 10.1109/TRO.2022.3226144.

[7] L. Zhang, J. Zhao, E. Lamon, Y. Wang, X. Hong. Energy efficient multi-robot task allocation constrained by time window and precedence. *IEEE Transactions on Automation Science and Engineering*, published online. DOI: 10.1109/TASE.2023.3312214.

[8] J. Zhao, A. Giammarino, E. Lamon, J. M. Gandarias, E. D. Momi, A. Ajoudani. A hybrid learning and optimization framework to achieve physically interactive tasks with mobile manipulators. *IEEE Robotics and Automation Let-*

ters, vol. 7, no. 3, pp. 8036–8043, 2022. DOI: 10.1109/LRA.2022.3187258.

[9] J. Fu, I. Burzo, E. Iovene, J. Zhao, G. Ferrigno, E. De Momi. Optimization-based variable impedance control of robotic manipulator for medical contact tasks. *IEEE Transactions on Instrumentation and Measurement*, vol. 73, Article number 4004608, 2024. DOI: 10.1109/TIM.2024.3372209.

[10] A. M. Zanchettin, P. Rocco. Reactive motion planning and control for compliant and constraint-based task execution. In *Proceedings of IEEE International Conference on Robotics and Automation*, Seattle, USA, pp. 2748–2753, 2015. DOI: 10.1109/ICRA.2015.7139572.

[11] F. Tassi, E. De Momi, A. Ajoudani. Augmented hierarchical quadratic programming for adaptive compliance robot control. In *Proceedings of IEEE International Conference on Robotics and Automation*, Xi'an, China, pp. 3568–3574, 2021. DOI: 10.1109/ICRA48506.2021.9561506.

[12] A. Mavrommati, C. Osorio, R. G. Valenti, A. Rajhans, P. J. Mosterman. An application of model predictive control to reactive motion planning of robot manipulators. In *Proceedings of the 17th International Conference on Automation Science and Engineering*, Lyon, France, pp. 915–920, 2021. DOI: 10.1109/CASE49439.2021.9551432.

[13] M. A. Roa, D. Berenson, W. Huang. Mobile manipulation: Toward smart manufacturing. *IEEE Robotics & Automation Magazine*, vol. 22, no. 4, pp. 14–15, 2015. DOI: 10.1109/MRA.2015.2486583.

[14] T. Teng, M. Fernandes, M. Gatti, S. Poni, C. Semini, D. Caldwell, F. Chen. Whole-body control on non-holonomic mobile manipulation for grapevine winter pruning automation. In *Proceedings of the 6th IEEE International Conference on Advanced Robotics and Mechatronics*, Chongqing, China, pp. 37–42, 2021. DOI: 10.1109/ICARM52023.2021.9536083.

[15] T. Sandakalum, M. H. Jr. Ang. Motion planning for mobile manipulators – A systematic review. *Machines*, vol. 10, no. 2, Article number 97, 2022. DOI: 10.3390/machines10020097.

[16] J. Zhao, F. Tassi, Y. Huang, E. De Momi, A. Ajoudani. A combined learning and optimization framework to transfer human whole-body loco-manipulation skills to mobile manipulators, [Online], Available: https://arxiv.org/abs/2402.13915, 2024.

[17] Y. Yang, F. Meng, Z. Meng, C. Yang. RAMPAGE: Toward whole-body, real-time, and agile motion planning in unknown cluttered environments for mobile manipulators. *IEEE Transactions on Industrial Electronics*, vol. 71, no. 11, pp. 14492–14502, 2024. DOI: 10.1109/TIE.2024.3370969.

[18] N. Vahrenkamp, T. Asfour, R. Dillmann. Robot placement based on reachability inversion. In *Proceedings of IEEE International Conference on Robotics and Automation*, Karlsruhe, Germany, pp. 1970–1975, 2013. DOI: 10.1109/ICRA.2013.6630839.

[19] S. Jauhri, J. Peters, G. Chalvatzaki. Robot learning of mobile manipulation with reachability behavior priors. *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 8399–8406, 2022. DOI: 10.1109/LRA.2022.3188109.

[20] J. Pankert, M. Hutter. Perceptive model predictive control for continuous mobile manipulation. *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 6177–6184, 2020. DOI: 10.1109/LRA.2020.3010721.

[21] A. Heins, A. P. Schoellig. Keep it upright: Model predictive control for nonprehensile object transportation with obstacle avoidance on a mobile manipulator. *IEEE Robotics and Automation Letters*, vol. 8, no. 12, pp. 7986–7993, 2023. DOI: 10.1109/LRA.2023.3324520.

[22] J. Haviland, N. Sünderhauf, P. Corke. A holistic approach to reactive mobile manipulation. *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 3122–3129, 2022. DOI: 10.1109/LRA.2022.3146554.

[23] C. Yang, C. Zeng, C. Fang, W. He, Z. Li. A dmps-based framework for robot learning and generalization of human-like variable impedance skills. *IEEE/ASME Transactions on Mechatronics*, vol. 23, no. 3, pp. 1193–1203, 2018. DOI: 10.1109/TMECH.2018.2817589.

[24] J. Zhao, G. J. G. Lahr, F. Tassi, A. Santopaolo, E. De Momi, A. Ajoudani. Impact-friendly object catching at non-zero velocity based on combined optimization and learning. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, Detroit, USA, pp. 4428–4435, 2023. DOI: 10.1109/IROS55552.2023.10341600.

[25] T. Teng, M. Gatti, S. Poni, D. Caldwell, F. Chen. Fuzzy dynamical system for robot learning motion skills from human demonstration. *Robotics and Autonomous Systems*, vol. 164, Article number 104406, 2023. DOI: 10.1016/j.robot.2023.104406.

[26] T. Welschehold, C. Dornhege, W. Burgard. Learning mobile manipulation actions from human demonstrations. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vancouver, Canada, pp. 3196–3201, 2017. DOI: 10.1109/IROS.2017.8206152.

[27] Y. Huang, L. Rozo, J. Silvério, D. G. Caldwell. Kernelized movement primitives. *The International Journal of Robotics Research*, vol. 38, no. 7, pp. 833–852, 2019. DOI: 10.1177/0278364919846363.

[28] J. Silvério, Y. Huang. A non-parametric skill representation with soft null space projectors for fast generalization. In *Proceedings of IEEE International Conference on Robotics and Automation*, London, UK, pp. 2988–2994, 2023. DOI: 10.1109/ICRA48891.2023.10161065.

Jianzhuang Zhao received the B. Eng. degree in mechanical engineering from Zhengzhou University, China in 2018, the M. Eng. degree in mechanical engineering from the Shaanxi Key Laboratory of Intelligent Robots, School of Mechanical Engineering, Xi'an Jiaotong University, China in 2021, and the Ph. D. degree in bioengineering from the Italian Institute of Technology and Polytechnic University of Milan, Italy in 2024. He is currently a postdoctoral researcher with the Human-Robot Interfaces and Interaction (HRI²) Lab at the Italian Institute of Technology (IIT), Italy. He was a finalist for the Best Paper on Mobile Manipulation at IROS 2022.

His research interests include robot learning, mobile manipulation, and human-robot interaction and collaboration.

E-mail: jianzhuang.zhao@iit.it (Corresponding author)

ORCID iD: 0000-0003-0860-4768

Tao Teng received the B.Sc. and the M.Sc. degrees in automation from the South China University of Technology, China in 2019 and 2016, respectively, and the Ph.D. degree in agri-system (robotics) from the Catholic University of the Sacred Heart, Italy in 2023. He is currently a postdoctoral fellow at the Hong Kong Centre for Logistics Robotics and The Chinese University of Hong Kong, China.

His research interests include human–robot interaction, robot learning, and optimal control.

E-mail: tao.teng@foxmail.com (Corresponding author)
ORCID iD: 0000-0002-4064-1043

Elena De Momi received the M.Sc. degree in biomedical engineering and the Ph.D. degree in bioengineering from Polytechnic University of Milan, Italy in 2002 and 2006, respectively, and she is currently an associate professor in the Electronic Information and Bioengineering Department (DEIB), Polytechnic University of Milan, Italy. She co-founded the Neuroengineering and Medical Robotics Laboratory in 2008, where she is responsible for the Medical Robotics section. She is currently a senior editor of the *International Journal of Robotics Research*, an editor of the *IEEE Robotics and Automation Magazine* and an Associate Editor of *IEEE Transaction on Robotics and Robotics and Automation Letters*. She has been Publication Chair and Co-Chair for IEEE ICRA 2019, 2023 and 2024 and is Program Chair of IROS 2025. She is responsible for the lab course in medical robotics, as well as the courses on Clinical Technology Assessment and Smart Hospital within the MSc program in biomedical engineering at Polytechnic University of Milan, Italy. She also serves on the board committee of the PhD program in bioengineering and the National PhD program in robotics and intelligent machines.

Her research interests include medical robotics, computer vision, artificial intelligence, and human robot interaction.

E-mail: elena.demomi@polimi.it
ORCID iD: 0000-0002-8819-2734

Arash Ajoudani received the Ph.D. degree in robotics and automation from the University of Pisa, Italy and Italian Institute of Technology (IIT), Italy in 2014. He is currently a tenured senior scientist with the Italian Institute of Technology (IIT), Italy, where he leads the Human-Robot Interfaces and Interaction (HRI^2) Laboratory. He was a recipient of the European Research Council (ERC) starting grant 2019, the coordinator of the Horizon-2020 project SOPHIA, and the co-coordinator of the Horizon-2020 project CONCERT.

His research interests include physical human–robot interaction, mobile manipulation, robust and adaptive control, assistive robotics, and telerobotics.

E-mail: arash.ajoudani@iit.it
ORCID iD: 0000-0002-1261-737X