Collaborative Bimanual Manipulation Using Optimal Motion Adaptation and Interaction Control

Retargeting Human Commands to Feasible Robot Control References

By Ruoshi Wen, Quentin Rouxel, Michael Mistry, Zhibin Li, and Carlo Tiseo

This article presents a robust and reliable human–robot collaboration (HRC) framework for bimanual manipulation. We propose an optimal motion adaptation method to retarget arbitrary human commands to feasible robot pose references while maintaining payload stability. The framework comprises three modules: 1) *a task-space sequential equilibrium and inverse kinematics optimization* (*task-space SEIKO*) for retargeting human commands and enforcing feasibility constraints, 2) an admittance controller to facilitate compliant human–robot physical interactions, and 3) a low-level controller improving stability during physical interactions. Experimental results show that the proposed framework successfully adapted infeasible and dangerous human commands into continuous motions within safe boundaries and achieved stable grasping and maneuvering of large and heavy objects on a real dual-arm robot via teleoperation and physical interaction. Furthermore, the framework demonstrated the capability in the assembly task of building blocks and the insertion task of industrial power connectors.

Digital Object Identifier 10.1109/MRA.2023.3270222
Date of current version: 16 May 2023

INTRODUCTION

Human-robot collaboration (HRC) allows humans and robots to work closely in a shared workspace, and it plays an important role in Industry 4.0 [1], space exploration [2], and medical applications [3]. Efficient collaboration between human operators and robots combines the advantages of both sides, i.e., human dexterity, flexibility, and adaptability, and robots' high payload, improved accuracy, and 24/7 running.

Compared to single-arm manipulators, multiarm robots have the same strength as parallel manipulators, but they are comparatively more powerful and rigid. The power or size of a gripper does not restrict the grasp of multiarm robots, making them more versatile and efficient in certain applications. For example, dual-arm robots have unique advantages in carrying heavy, bulky, and large objects and performing complex manufacturing tasks, such as part assembly or connector insertion. They also have the advantage of reconfiguring their grasp and are not limited to a fixed or smaller workspace. However, autonomous motion planning for high degree-of-freedom (DoF) robots during long-horizon and multiphase manipulation tasks remains challenging because of the diverse interaction modes as well as the complexity of contact dynamics and

1070-9932/23©2023IEEE

friction. HRC frameworks are worth investigating to resolve the challenges in bimanual manipulation.

Teleoperation has great potential in dexterous manipulation tasks as it enables human operators to remotely control robots using master devices and their own motor skills [4], [5]. Recent development in collaborative robots (cobots) has made it possible to safely operate robots in close proximity to humans [1]. By incorporating the complementary expertise of human operators into the robotic control, the intelligence and capabilities of the robotic system can be significantly enhanced. The assembly task shown in Figure 1 involves a sequential combination of teleoperation and physical interaction by multiple operators. The remote operator, who has a global view of the situation, teleoperates the two robotic arms to reach, maintain contact with, hold, and move one part. The local operator then guides and fine-tunes the manipulation process by physically interacting with the held object, leveraging detailed on-site observations to complete the assembly task.

Teleoperation and physical interaction are two effective ways to realize HRC. However, human commands are inevitably subject to errors because of the lack of direct access and intuitive perception of the robot's physical limits, which could violate the robot's physical constraints and cause task failures and safety issues. Therefore, the robot's local control system should be robust to infeasible commands. To realize safe HRC, it is necessary to adapt these references to feasible motions that satisfy the robot's physical limitations and task constraints.

This article introduces several technical innovations that enhance the robustness of collaborative bimanual manipulation, building upon our previous works [6], [7]. Our framework provides an optimal motion adaptation approach that enables the robot to execute operator commands from both remote teleoperation and local physical interaction, ensuring feasibility and robustness to errors. *Constraint feasibility for robots* refers to ensuring that the physical limits of the robots (joint torque and position limits, and singularity) are respected. *Constraint feasibility for the operating payload* refers to ensuring the stable grasping of the object by the two end effectors during maneuvering and preventing any instances of slipping or falling. Our framework improves the safety of human operators in the vicinity of the robots by reliably holding heavy objects between two end effectors and avoiding breaking the physical limits of the robots.

First, we introduce the task-space SEIKO. This formulation addresses the safety concerns caused by error-prone human commands used to control the poses of a grasped object. The new formulation presented in this study differs from our previous SEIKO work [6] because it considers the equilibrium of the grasped object in addition to the robot limits and task constraints. This allows the robot to complete the challenging grasping task of a significantly tilted box.

Second, our system receives physical interaction commands by measuring external wrenches. An admittance controller provides the robot with compliance to physically interact with local operators while resisting their push by adapting their commands within feasible boundaries. Our approach accepts remote and local commands simultaneously and superimposes the remote position commands and the velocity commands (estimated from the measured external wrenches by the admittance controller) from physical interaction. The task-space SEIKO ensures the feasibility of the combined commands. Our system is flexible and reliable and facilitates seamless collaboration between remote experts and local users in a wide range of human–robot collaborative scenarios.

Finally, the optimal and feasible motions generated by the task-space SEIKO are realized by an interaction controller. This low-level controller, based on fractal impedance control (FIC) [7], provides robust stability during human–robot physical interaction at a low computation cost. The FIC's conservative observer enables multiple controllers' stable superimposition, which is robust to time delay and reduced control bandwidth [3], [8]. This architecture enables the development of a fully operational system capable of real-world bimanual manipulation tasks that are considerably more complex than those of previous works.

The contributions of our work are summarized as follows:

1) The *task-space SEIKO* formulation enables real-time bimanual manipulation of heavy payloads by optimally adapting infeasible Cartesian commands from operators, which ensures that robot limits and task constraints are satisfied in real time, improving the safety and effectiveness of bimanual manipulation (1 kHz).
2) A *collaborative bimanual manipulation framework* facilitates simultaneous teleoperation by a remote operator and physical interaction by a local operator, with the superimposed commands retargeted and adapted by the task-space SEIKO to ensure feasibility and improve the efficiency of multioperator scenarios.

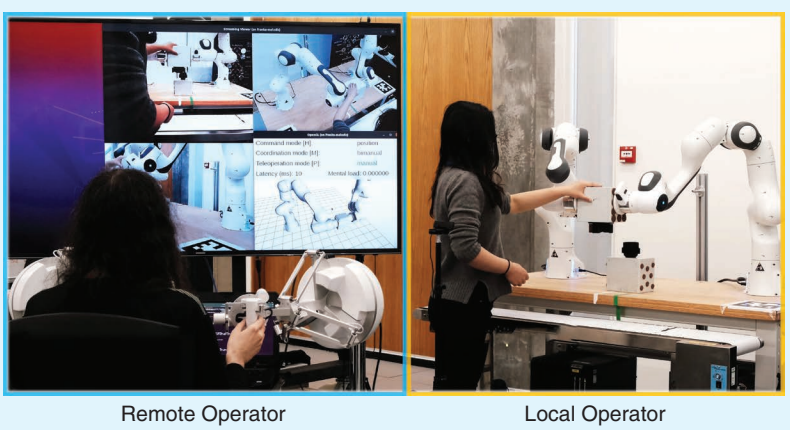

FIGURE 1. Real-world part assembly task performed via teleoperation and direct physical interaction using the proposed collaborative bimanual manipulation framework.

3) *Complex and realistic bimanual manipulation tasks* involve maneuvering heavy objects in arbitrary ways, guiding the robot through physical interactions with stacks of unknown objects, collaborative part assembly among multiple operators, and insertion of industrial connectors, highlighting the potential impact of our work in real-world applications.

Through challenging experiments, we have demonstrated the capability of our framework to ensure constraint feasibility for robots, human operators, and the grasped payload in real-world collaborative bimanual manipulation tasks. The results demonstrate a high level of flexibility to fuse commands from multiple intelligent agents, each possessing unique motor skills and feedback mechanisms, thus facilitating decision making. The contributions of our work pave the way for more robust multiagent cooperation, reducing physical and cognitive demands on human operators and significantly improving productivity and efficiency.

The article is organized as follows. The section "Related Works" summarizes the related works on human–robot collaborative bimanual manipulation. The section "Collaborative Manipulation Framework" overviews the framework's architecture and formulates the task-space SEIKO and the interaction controller. The sections "Experiments and Results" and "Discussion" detail our experimental setups and evaluate the performance of the proposed method, and the section "Conclusions and Future Work" draws the conclusions.

RELATED WORKS

Inverse kinematics (IK) has been extensively used to compute whole-body joint positions to get the desired end effectors' poses while considering kinematic constraints in teleoperation. For example, the quasi-static equilibrium of humanoids on flat ground is formulated as an IK problem where the center of mass (CoM) projection is constrained within the feet support polygon. However, contact wrenches, which are important for bimanual manipulation tasks, are not considered in IK-based schemes. For loco-manipulation tasks in multicontact settings, a motion retargeting framework—SEIKO—is proposed in [6] and has been validated by the teleoperation of high-DoF robots, such as humanoids and quadrupeds. The retargeting of multicontact motions is formulated as sequential quadratic programming, which can optimize the whole-body configurations and contact wrenches in real time. Although originally formulated in joint space for floating-base robots, SEIKO can be rewritten to address bimanual manipulation by specifically considering the task-space constraints of the grasped object and the kinematic loop of the two arms.

HRC through physical interaction requires either direct contact between humans and robots or indirect force exchange through an object. The applications of physical human–robot interaction range from collaborative manufacturing, such as assembly [9], to emotional support, such as hugging robots [10]. Impedance control [11] and admittance control [12] are typically used for physical human–robot interaction to realize the desired dynamic behavior at its ports of interaction with the environment, and their improved control for reactive interaction has enabled close and efficient collaboration between robots and humans.

Multiple admittance and impedance controllers have been proposed to modulate the contact port performances to adapt to different task requirements over the years by exploiting both optimization and learning methods [13], [14]. Although it is theoretically possible to combine the admittance controller with any passive impedance controller, it has limited responsiveness and tracking accuracy. The FIC controller [3] has the advantages of a passive impedance controller while improving responsiveness and tracking accuracy. Because of its highly nonlinear stable dynamics, the system can accurately track trajectories while still maintaining the intrinsically soft behavior of impedance controllers [7]. FIC can be combined with admittance controllers because it allows the superimposition of multiple independent controllers while retaining the system's stability.

However, the ability to generate robust interaction with high tracking accuracy is not sufficient to enable the success of bimanual manipulation tasks. The challenges of these tasks include the coordination of the two arms and the retargeting of human commands that violate the robot's physical limits [15], [16], [17], [18]. The coordination of multiple robotic arms requires the formulation of dynamic-aware motion optimization that deals with intrinsic nonlinearity and nonholonomic constraints in concave domains. There are multiple ways to deal with these constraints, such as hierarchical quadratic programming (HQP) [15], [16], [17] and nonlinear programming (NLP) [18].

HQP solves this problem by dividing the action into a hierarchical set of tasks. The decomposition of the problem into subproblems simplifies the constraints by decoupling them while regarding the codomain of the higher task as the domain of the current optimization [15], [16]. NLP uses a general formulation to tackle nonlinear optimizations [18]. HQP and NLP can be formulated in both inverse and forward dynamics [15]. It is shown that the formulation of the inverse dynamics has faster convergence when using simultaneous methods (e.g., direct transcription) for NLP. These methods decompose the problem into a discrete set of subproblems enforcing boundary constraints during the transitions between subsequent domains, guaranteeing a smooth transition. The parallel processing of the optimization problem solves complex mathematical problems faster, which allows for modeling the mechanism capabilities more accurately.

These methods have become increasingly more efficient over the last decade, and the computational time has been reduced drastically. Nevertheless, the coordinated dynamic interaction of multiple robotic arms is still an open problem because of the intrinsic variability in the interaction with the environment, which is either unknown or difficult to model.

The gap between simulation and reality will often reduce the robustness of interaction.

COLLABORATIVE MANIPULATION FRAMEWORK

The control framework of the collaborative bimanual manipulation is shown in Figure 2. The operator commands can be given via teleoperation or physical interaction. The proposed framework consists of three parts: the admittance controller for compliant physical interaction and modulating the contact forces; the task-space SEIKO, which optimizes the target pose of the object $\boldsymbol{X}_O^{\text{target}} \in SE(3)$ subject to the contact constraints and the physical limits of the robot; and the interaction controller, which generates the joint torque commands for robots' stability during the interaction.

We propose a task-space formulation of SEIKO for the motion adaptation of infeasible human commands. It runs one iteration per control cycle to optimize the feasible state for the next time step, with a computation time of less than 1 ms. The optimal commands generated by the task-space SEIKO are realized by the interaction controller, which has three FIC controllers and two compensatory terms to generate the desired joint torques for each arm.

OPERATOR INTERFACE

We use three cameras as multicamera feeds to capture live streams of the robot for the remote operators; the interface is shown in Figure 2(a). The remote operators can teleoperate the robot via two 6-DoF haptic devices (Force Dimension Sigma 7) using two operation modes: the independent and the bimanual modes. These modalities also include two submodalities, namely pose and twist, that can be chosen by the operator online. The pose submodality uses the displacement of the current pose from the nominal pose measured by the Sigma 7 haptic device as the control reference of the robot, which is more intuitive to operators. The twist submodality has better use of the limited workspace of the haptic device as it can accumulate the displacement of the operator's pose from the nominal pose as the desired end-effector pose. In the independent mode, the remote operator teleoperates the left end effector via the left haptic device and the right end effector via the right device in Cartesian space, using either command mode (pose/twist) or a hybrid of both. After placing the object between the two hands, the operator can enable the bimanual mode, in which the human commands are sent to control the pose of the CoM of the object via one haptic

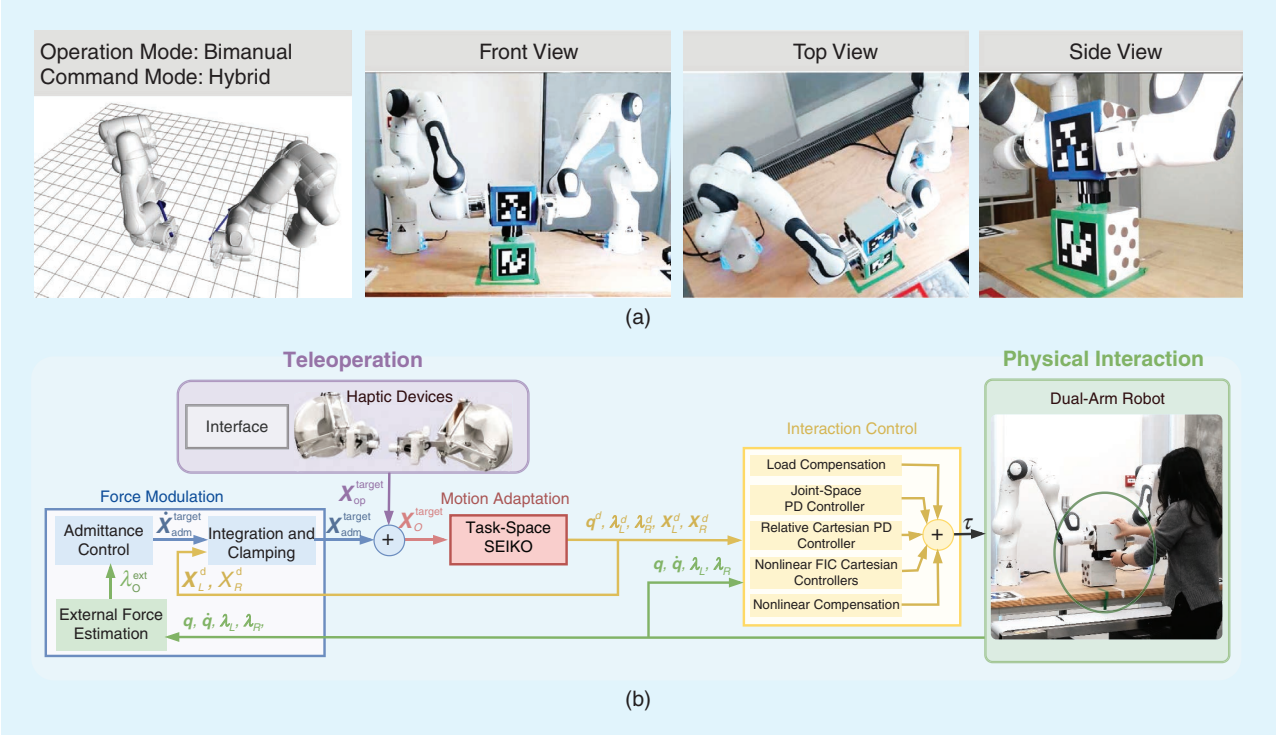

FIGURE 2. The overall architecture of the collaborative bimanual manipulation framework. (a) The teleoperation interface with multicamera views. The interactive 3D scene is rendered from the robot state for the remote operator to switch mode. The images show that the remote user is operating in the bimanual mode with hybrid command mode. (b) Both the pose commands from the teleoperation interface and the physical interaction commands are simultaneously superimposed to generate the target object pose. Pose commands are expressed relative to a reference pose; external forces are processed by an admittance controller to velocity commands, which are integrated into the reference pose [see (1) for details]. The task-space SEIKO then adapts this combined target pose to satisfy the physical limits of the robotic system and the task constraints to produce the desired object poses as well as the desired contact wrenches. The interaction controller implements the superimposed controllers, including the passive impedance controllers (joint-space PD, Cartesian coordination, and nonlinear FIC) to provide stability during human–robot interaction, and the compensatory controllers for the load compensation and the nonlinear dynamics (gravity, Coriolis and centrifugal effects). PD: proportional derivative.

device. In bimanual mode, the target pose commands from the operator are optimized by the task-space SEIKO to satisfy the task constraints and physical limits of the robots.

The controller used for the haptic feedback on the Sigma 7 is adopted from our previous work, and its technical details can be found in [19]. The target end-effector pose $\boldsymbol{X}_{\text{op}}^{\text{target}}$ for the two command modes in the bimanual mode can be computed using (1) from human commands, which will be sent to the task-space SEIKO for motion adaptation.

$$\begin{aligned} \boldsymbol{X}_{\text{op}}^{\text{target}} &= \boldsymbol{X}_{\text{ref}}^{\text{target}} + \Delta\boldsymbol{X}_{S\text{-}7}, \quad \text{pose mode} \\ \boldsymbol{X}_{\text{ref}}^{\text{target}} &= \boldsymbol{X}_{\text{ref0}}^{\text{target}} + \Delta\boldsymbol{X}_{S\text{-}7}\Delta t, \quad \text{twist mode} \end{aligned} \tag{1}$$

where $\boldsymbol{X}_{\text{ref}}^{\text{target}}$ is the reference rest pose of the end effector, $\Delta\boldsymbol{X}_{S\text{-}7}$ is the displacement of the current pose from the nominal pose measured by the Sigma 7 haptic device, and $\boldsymbol{X}_{\text{ref0}}^{\text{target}}$ is the end effector's reference pose of the arm at the last time step.

END-EFFECTOR FORCE MODULATION

The model-based state estimation uses an admittance controller to track the desired interaction force at the end effector, which takes the force tracking error as input and generates the end effectors' pose commands. The admittance controller and the task-space SEIKO handle the task performance independently and drive the impedance controller to generate the desired behavior. The admittance controller filters the measured force tracking error and generates the desired robot end-effector pose, which can be adaptive to different tasks by producing the desired interaction dynamics at the end effector, as follows:

$$\ddot{\boldsymbol{X}}_{\text{adm}}^{\text{target}} = \boldsymbol{M}^{-1}\left(\boldsymbol{\lambda}_O^{\text{ext}} - \boldsymbol{\lambda}_{\text{adm}}^{d}\right) \tag{2}$$

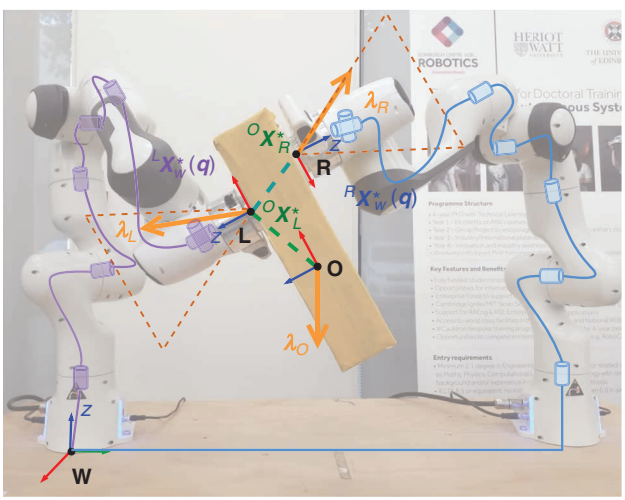

FIGURE 3. Illustration of notation of reference frames, force, and spatial transformations formulated in the optimal motion adaptation. The end-effector frames of the left hand (L) and the right hand (R) and the object frame (the origin O is located at its CoM) are depicted. The world frame W is at the base of the left arm. The contact wrenches between the object and the left and right arms are denoted by $\boldsymbol{\lambda}_L, \boldsymbol{\lambda}_R$, respectively, while $\boldsymbol{\lambda}_O$ denotes the gravitational wrench. The friction cones of the two contact wrenches are depicted by the orange dashed lines.

where $\boldsymbol{M}^{-1}$ is the desired inertia at the end effector, $\boldsymbol{\lambda}_O^{\text{ext}}$ is the estimated end-effector effort, and $\boldsymbol{\lambda}_{\text{adm}}^{d}$ is the desired interaction wrench. $\boldsymbol{\lambda}_{\text{adm}}^{d}$ can be treated either as an external input or modeled as a preassigned spring–damper system as follows:

$$\boldsymbol{\lambda}_{\text{adm}}^{d} = \boldsymbol{K}_d\left(\boldsymbol{X}^{d} - \boldsymbol{X}\right) - \boldsymbol{D}_d\dot{\boldsymbol{X}}, \tag{3}$$

where $\boldsymbol{K}_d$ and $\boldsymbol{D}_d$ are the desired damping and stiffness, respectively. The desired acceleration of the two arms is integrated twice to get the desired end-effector pose $(\boldsymbol{X}_{\text{adm}}^{\text{target}})$, as follows:

$$\begin{cases} \dot{\boldsymbol{X}}_{\text{adm}}^{\text{target}}(t+\Delta t) = \dot{\boldsymbol{X}}_{\text{adm}}^{\text{target}}(t) + \Delta\dot{\boldsymbol{X}}_{\text{adm}}^{\max}(t) \\ \boldsymbol{X}_{\text{adm}}^{\text{target}}(t+\Delta t) = \boldsymbol{X}_{\text{adm}}^{\text{target}}(t) + \dot{\boldsymbol{X}}_{\text{adm}}(t)\Delta t \end{cases} \tag{4}$$

where

$$\Delta\dot{\boldsymbol{X}}_{\text{adm}}^{\max} = \text{sign}(\ddot{\boldsymbol{X}}_{\text{adm}})\max(|\ddot{\boldsymbol{X}}_{\text{adm}}|\Delta t, \Delta\dot{\boldsymbol{X}}_{\max})$$

where $\Delta\dot{\boldsymbol{X}}_{\max} = \left\| \boldsymbol{X}_O^{\text{target}}(t) - \boldsymbol{X}_O^{d}(t) \right\| / \Delta t$ is the feasibly maximum twist change of the selected arm at the current time step, bounding the end-effector twist to the feasibility sphere around the current end-effector state effectively.

The state estimator in Figure 2(b) estimates the external interaction with the manipulated object by subtracting the weight of the held object from the wrench computed from the superimposition of the measured wrenches at the robot end effectors $\boldsymbol{\lambda}_O^{\text{ext}} = {}^{O}\boldsymbol{X}_L^*\boldsymbol{\lambda}_L + {}^{O}\boldsymbol{X}_R^*\boldsymbol{\lambda}_R$, where ${}^{O}\boldsymbol{X}_L^*$ and ${}^{O}\boldsymbol{X}_R^* \in \mathbb{R}^{6\times 6}$ are the wrench transformation matrix between the robot end-effector frames of the two arms (L, R) and the object frame (O). $\boldsymbol{\lambda}_L$ and $\boldsymbol{\lambda}_R \in \mathbb{R}^6$ are the measured end-effector wrenches for the left and the right arm, respectively.

MOTION ADAPTATION

The motion adaptation method takes the end-effector pose commands as input and solves the posture and contact forces $(q, \boldsymbol{\lambda})$ simultaneously as a nonlinear optimization problem. We formulate a task-space SEIKO in which the equilibrium of the external object held between the two arms is expressed in Cartesian space. The reference frames used to estimate the object's static equilibrium from the robot state are shown in Figure 3. SEIKO transforms an input target pose of the CoM of the manipulated object $\boldsymbol{X}_O^{\text{target}} \in SE(3)$ expressed in the world frame to the optimized state of the arms that guarantees feasibility. The task-space SEIKO executes a single iteration of the optimization for each control cycle. At each iteration, the state change $(\Delta\boldsymbol{q}, \Delta\boldsymbol{\lambda})$ is computed to update the previous state to a new desired state $\boldsymbol{q}^{d} \in \mathbb{R}^n$, where n is the total number of the left and right robotic arms, and the desired wrenches applied on the object are $\boldsymbol{\lambda}_L^{d}, \boldsymbol{\lambda}_R^{d} \in \mathbb{R}^6$ by the two arms. The Cartesian poses $\boldsymbol{X}_L^{d}, \boldsymbol{X}_R^{d} \in SE(3)$ of the two end effectors are then computed by forward kinematics.

The task-space SEIKO is formulated as a constrained nonlinear optimization and solved by a sequence of QP problems.

Both the cost function and the constraints are linearized and approximated at the first order and rely on analytical derivatives for computational speed and stability. The constrained least squares optimization problem to be solved at each control loop is as follows:

$$\begin{aligned} \min_{\Delta \boldsymbol{x}} \; & \| \boldsymbol{C}_{\text{cost}}(\boldsymbol{x})\,\Delta\boldsymbol{x} - \boldsymbol{c}_{\text{cost}}(\boldsymbol{x}) \|^2_{\boldsymbol{w}} \\ & \text{s.t. } \boldsymbol{C}_{\text{eq}}(\boldsymbol{x})\,\Delta\boldsymbol{x} + \boldsymbol{c}_{\text{eq}}(\boldsymbol{x}) = \boldsymbol{0}, \\ & \boldsymbol{C}_{\text{ineq}}(\boldsymbol{x})\,\Delta\boldsymbol{x} + \boldsymbol{c}_{\text{ineq}}(\boldsymbol{x}) \geqslant \boldsymbol{0}, \\ & \text{where } \boldsymbol{x} = \begin{bmatrix} \boldsymbol{q}^d \\ \boldsymbol{\lambda}_L^d \\ \boldsymbol{\lambda}_R^d \end{bmatrix}, \Delta\boldsymbol{x} = \begin{bmatrix} \Delta\boldsymbol{q} \\ \Delta\boldsymbol{\lambda}_L \\ \Delta\boldsymbol{\lambda}_R \end{bmatrix} \end{aligned} \tag{5}$$

where $\boldsymbol{x} \in \mathbb{R}^{n+6+6}$ is the current desired state, and the incremental change $\Delta\boldsymbol{x}$ is the decision variable. $\boldsymbol{q}^d \in \mathbb{R}^n$ is the vector containing all joint positions of the two robotic arms, and $\boldsymbol{C}_{\text{cost}}, \boldsymbol{c}_{\text{cost}}, \boldsymbol{C}_{\text{eq}}, \boldsymbol{c}_{\text{eq}}, \boldsymbol{C}_{\text{ineq}}$, and $\boldsymbol{c}_{\text{ineq}}$ are the matrices and vectors defining the cost, equality, and inequality constraints, respectively. Our decision variable here is the incremental change $\Delta\boldsymbol{x}$, and the desired state is updated at each iteration by $\boldsymbol{x}(t+\Delta t) = \boldsymbol{x}(t) + \Delta\boldsymbol{x}(t)$. The equality and inequality constraints are detailed as follows.

EQUALITY CONSTRAINTS

To adapt the motions within the stability region for the success of bimanual manipulation tasks, we introduce a constraint term based on the grasp matrix and the robot state to the given optimization formulation. The frames used to formulate the optimization problem include the world frame $\boldsymbol{W}$, the object frame $\boldsymbol{O}$, and the left and right end-effector frames $\boldsymbol{L}$, $\boldsymbol{R}$, respectively, depicted in Figure 3. We align the object frame $\boldsymbol{O}$ to the frame $\boldsymbol{L}$ and set its origin at the CoM of the object for a simplified static equilibrium equation. The position of $\boldsymbol{O}$ can be estimated from the contact state.

The static equilibrium equation of an object in the bimanual manipulation task is as follows:

$$ {}^{O}\boldsymbol{X}^*_{W}(\boldsymbol{q})\,\boldsymbol{\lambda}_{\text{o}} = \begin{bmatrix} {}^{O}\boldsymbol{X}^*_{L} & {}^{O}\boldsymbol{X}^*_{R} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}_{L} \\ \boldsymbol{\lambda}_{R} \end{bmatrix} \tag{6}$$

where $\boldsymbol{\lambda}_L, \boldsymbol{\lambda}_R \in \mathbb{R}^6$ are the wrenches expressed in the left and right end-effector frames $\boldsymbol{L}$, $\boldsymbol{R}$, while ${}^{O}\boldsymbol{X}^*_{L}$ and ${}^{O}\boldsymbol{X}^*_{R} \in \mathbb{R}^{6\times 6}$ are the spatial wrench transformation matrix [20] from $\boldsymbol{L}$ and $\boldsymbol{R}$ to the object frame $\boldsymbol{O}$, respectively. The z-axes of $\boldsymbol{L}$ and $\boldsymbol{R}$ are parallel to the normal direction of the contact surface. $\boldsymbol{\lambda}_O = [\boldsymbol{\tau}_O\; \boldsymbol{f}_O]^\top = [0\;0\;0\;0\;0\;-mg]^\top$ is the gravitational wrench of the object expressed in the world frame $\boldsymbol{W}$, where m is the object's mass and g the gravitational acceleration. ${}^{O}\boldsymbol{X}^*_{W}(\boldsymbol{q}) \in \mathbb{R}^{6\times 6}$ is the wrench transformation from the world frame $\boldsymbol{W}$ to the object frame $\boldsymbol{O}$. Because $\boldsymbol{\lambda}_O$ does not have a torque component ($\boldsymbol{\tau}_O = \boldsymbol{0}$) since it acts on the CoM of the object, ${}^{O}\boldsymbol{X}^*_{W}(\boldsymbol{q})$ can be simplified as $\begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & {}^{O}\boldsymbol{R}_{W}(\boldsymbol{q}) \end{bmatrix}$, where ${}^{O}\boldsymbol{R}_{W}(\boldsymbol{q}) \in SO(3)$ is the rotational term of ${}^{O}\boldsymbol{X}^*_{W}(\boldsymbol{q})$, which only depends on the joint positions.

The equality constraint for static equilibrium in (6) is differentiated with respect to $[\Delta\boldsymbol{q}\;\Delta\boldsymbol{\lambda}_L\;\Delta\boldsymbol{\lambda}_R]^\top$, resulting in the following equation:

$$\begin{aligned} & \left[\begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{H}(\boldsymbol{q}, \boldsymbol{f}_O) \end{bmatrix} {}^{O}\boldsymbol{X}^*_{L} \;\; {}^{O}\boldsymbol{X}^*_{R} \right] \begin{bmatrix} \Delta\boldsymbol{q} \\ \Delta\boldsymbol{\lambda}_L \\ \Delta\boldsymbol{\lambda}_R \end{bmatrix} + \\ & \quad {}^{O}\boldsymbol{X}^*_{W}(\boldsymbol{q})\,\boldsymbol{\lambda}_O - {}^{O}\boldsymbol{X}^*_{L}\boldsymbol{\lambda}_L + {}^{O}\boldsymbol{X}^*_{R}\boldsymbol{\lambda}_R = \boldsymbol{0}, \\ & \text{where } \boldsymbol{H}(\boldsymbol{q}, \boldsymbol{f}_O) \in \mathbb{R}^{3\times n} = \\ & \quad -\frac{1}{2}\,{}^{O}\boldsymbol{R}_{L}\boldsymbol{S}({}^{L}\boldsymbol{R}_{W}(\boldsymbol{q})\boldsymbol{f}_O)\,{}^{L}\boldsymbol{J}^{\text{rot}}_{W}(\boldsymbol{q}) \\ & \quad -\frac{1}{2}\,{}^{O}\boldsymbol{R}_{R}\boldsymbol{S}({}^{R}\boldsymbol{R}_{W}(\boldsymbol{q})\boldsymbol{f}_O)\,{}^{R}\boldsymbol{J}^{\text{rot}}_{W}(\boldsymbol{q}) \end{aligned} \tag{7}$$

and where ${}^{L}\boldsymbol{J}^{\text{rot}}_{W}(\boldsymbol{q}), {}^{R}\boldsymbol{J}^{\text{rot}}_{W}(\boldsymbol{q}) \in \mathbb{R}^{3\times n}$ are the angular parts of the Jacobians of the frames $\boldsymbol{L}$ and $\boldsymbol{R}$ expressed in world frame $\boldsymbol{W}$, respectively. ${}^{L}\boldsymbol{R}_{W}(\boldsymbol{q}), {}^{R}\boldsymbol{R}_{W}(\boldsymbol{q}), {}^{O}\boldsymbol{R}_{L}(\boldsymbol{q}), {}^{O}\boldsymbol{R}_{R}(\boldsymbol{q}) \in SO(3)$ are the spatial rotational transformations between the frames depicted in Figure 3. $\boldsymbol{S}(\cdot): \mathbb{R}^3 \to \mathbb{R}^{3\times 3}$ is the operator generating the skew matrix associated to a 3D vector. We rearranged the terms in $\boldsymbol{H}(\boldsymbol{q}, \boldsymbol{f}_O)$ so that (7) is linear with respect to the decision variable $\Delta\boldsymbol{x}$ and is suitable for the QP formulation $\boldsymbol{C}_{\text{eq}}$ and $\boldsymbol{c}_{\text{eq}}$ in (5).

INEQUALITY CONSTRAINTS

The inequality constraints used in the optimization enforce the feasibility of the bimanual manipulation. Two sets of different constraints are defined. In joint space, we enforce the physical limits of the robotic arms by setting constraints for each joint: the maximum torque ($\boldsymbol{\tau}^d$), the maximum velocity ($\dot{\boldsymbol{q}}^d$), and the angular position range ($\boldsymbol{q}^d$). In task space, we constrain the contact wrenches ($\boldsymbol{\lambda}_L^d, \boldsymbol{\lambda}_R^d$) to enforce the stability of the contact, e.g., minimal and maximal normal contact force and limits on the friction pyramid and center of pressure. All of these constraints can be written in the linear form $\boldsymbol{C}_{\text{ineq}}, \boldsymbol{c}_{\text{ineq}}$ with respect to the decision variable $\Delta\boldsymbol{x}$ (see [6]), spawning $4n$ rows for the joint space and 36 rows for the task-space contact.

INTERACTION CONTROLLER

The interaction controller adopted from FIC, which is robust to time delay and stably superimposes multiple control efforts, is used to generate desired joint torques and control the dual-arm robot. As shown in Figure 2(b), the interactive controller is composed of five independent controllers.

The nonlinear FIC Cartesian controller ($\mathsf{NLPD}(\boldsymbol{X}^d, \boldsymbol{X}, \boldsymbol{\nu})$) independently drives the arms' end effectors toward their respective desired poses. The linear task-space proportional derivative (PD) controller ($\mathsf{PD}(\boldsymbol{X}^d_{\text{rel}}, \boldsymbol{X}_{\text{rel}}, \boldsymbol{\nu}_{\text{rel}})$) generates a wrench command to maintain the desired relative pose between the two arms. The linear joint-space PD controller ($\mathsf{PD}(\boldsymbol{q}^d, \boldsymbol{q}, \dot{\boldsymbol{q}})$) drives the arms toward the desired poses optimized by the task-space SEIKO.

The remaining components of the controller compensate for the arm nonlinear dynamics (Coriolis matrix and gravity com-

pensation) and the load weight ($\boldsymbol{\lambda}^d$). The torque command sent to the robotic arms is as follows:

$$\begin{aligned}\boldsymbol{\tau} = &\boldsymbol{C}(\boldsymbol{q},\dot{\boldsymbol{q}}) + \boldsymbol{G}(\boldsymbol{q}) + \mathrm{PD}(\boldsymbol{q}^d,\boldsymbol{q},\dot{\boldsymbol{q}}) \\ &+ \boldsymbol{J}_{\mathrm{local}}(\boldsymbol{q})^\top \boldsymbol{\lambda}^d \\ &+ \boldsymbol{J}_{\mathrm{world}}(\boldsymbol{q})^\top \mathrm{NLPD}(\boldsymbol{X}^d,\boldsymbol{X},\boldsymbol{\nu}) \\ &+ \boldsymbol{J}_{\mathrm{rel}}(\boldsymbol{q})^\top \mathrm{PD}(\boldsymbol{X}^d_{\mathrm{rel}},\boldsymbol{X}_{\mathrm{rel}},\boldsymbol{\nu}_{\mathrm{rel}})\end{aligned} \tag{8}$$

where $\boldsymbol{C}(\boldsymbol{q},\dot{\boldsymbol{q}}) \in \mathbb{R}^n$ is the vector containing the Coriolis and centrifugal torques, while $\boldsymbol{G}(\boldsymbol{q}) \in \mathbb{R}^n$ is the vector containing the gravitational torques; $\boldsymbol{q} \in \mathbb{R}^n$ and $\dot{\boldsymbol{q}} \in \mathbb{R}^n$ are the vectors containing the measured joint positions and velocities of the two arms; $\boldsymbol{q}^d \in \mathbb{R}^n$ is the desired joint positions computed by SEIKO; $\boldsymbol{J}_{\mathrm{local}}$ and $\boldsymbol{J}_{\mathrm{world}} \in \mathbb{R}^{(6+6)\times n}$ are the stacked Jacobian matrices of the two arms expressed in the local and world frames, respectively; and $\boldsymbol{J}_{\mathrm{rel}} \in \mathbb{R}^{6\times n}$ is the relative Jacobian between the two hands; $\boldsymbol{\lambda}^d \in \mathbb{R}^{6+6}$ is the stacked vector of the desired contact wrenches of the two arms computed by SEIKO, applied as a feedforward term; $\boldsymbol{X}$, $\boldsymbol{X}^d \in SE(3) \times SE(3)$ and $\boldsymbol{\nu} \in \mathbb{R}^{6+6}$ are the measured and desired Cartesian poses and measured twist of the two hands expressed in the world frame; $\boldsymbol{X}_{\mathrm{rel}}$ and $\boldsymbol{X}^d_{\mathrm{rel}} \in SE(3)$ and $\boldsymbol{\nu}_{\mathrm{rel}} \in \mathbb{R}^6$ are the measured and desired relative poses and measured relative twist between the two hands, respectively.

The passive PD controllers are implemented using six decoupled dimensions of the following monodimensional controller:

$$\begin{aligned}&\mathrm{PD}(\alpha^d,\alpha,\dot{\alpha}) = F_k - k_d\dot{\alpha}, \\ &F_k = \begin{cases} k_p(\alpha^d - \alpha), & \text{if } |\alpha^d - \alpha| < d \\ \mathrm{sign}(\alpha^d - \alpha)f, & \text{else} \end{cases}\end{aligned} \tag{9}$$

where α is the position (i.e., the monodimensional pose), α^d is the desired position, and $\dot{\alpha}$ is the velocity (i.e., the monodimensional twist). The gains $k_p = f/d \in \mathbb{R}$ and $k_d^{\mathrm{joint}} = 2\zeta\sqrt{k_p} \in \mathbb{R}$, which allow an intuitive tuning of the controllers.

Table 1. Control parameters for the joint space, and linear and angular control in Cartesian space.

PARAMETER	VALUE	UNIT	PARAMETER	VALUE	UNIT
f^{lin}	10–40	N	f^{ang}	2	N m
d^{lin}	0.08	m	d^{ang}	8	degree
ζ^{lin}	0.8		ζ^{ang}	0.2	
$f^{\mathrm{rel,lin}}$	50	N	$f^{\mathrm{rel,ang}}$	5	N
$d^{\mathrm{rel,lin}}$	0.05	m	$d^{\mathrm{rel,ang}}$	5	degree
$\zeta^{\mathrm{rel,lin}}$	0.4		$\zeta^{\mathrm{rel,ang}}$	0.1	
f^{joint}	0.3	N m			
d^{joint}	10	degree			
ζ^{joint}	0				

The NLPD is implemented using six decouple dimensions of the following mono-dimensional controller:

$$\begin{aligned}&\mathrm{NLPD}(\alpha^d,\alpha,\dot{\alpha}) = F_k - k_d\dot{\alpha}, \\ &F_k = \begin{cases} E(\tilde{\alpha}), & \text{Divergence} \\ \dfrac{2E(\tilde{\alpha}_{\max})}{\tilde{\alpha}_{\max}}\left(\tilde{\alpha} - \dfrac{\tilde{\alpha}_{\max}}{2}\right) & \text{Convergence} \end{cases} \\ &E = \begin{cases} k_p\tilde{\alpha}, & |\tilde{\alpha}| \le \xi d \\ \dfrac{\Lambda}{2}\left(\tanh\left(\dfrac{\tilde{\alpha} - \tilde{\alpha}_b}{S\tilde{\alpha}_b} + \pi\right) + 1\right) & \\ +E_0, & \text{else} \end{cases}\end{aligned} \tag{10}$$

where k_p is the constant stiffness, $\tilde{\alpha} = \alpha_d - \alpha$ is the position error (i.e., the monodimensional pose error), and d is the tracking error when the force saturation occurs. $\Lambda = E_{\max} - E_0$, $E_0 = \xi k_p d$, $S = (1-\xi)(d/(2\pi))$ controls the saturation speed, and $\xi = 0.9$ controls the starting of the saturation behavior while approaching d. Further details on the theoretical background of these controllers can be found in [3]. The parameters used for all of the controllers are reported in Table 1.

EXPERIMENTS AND RESULTS

EXPERIMENTAL SETUP

We set up a dual-arm robot using two Franka Emika robotic arms to validate the performance of the proposed collaborative bimanual manipulation framework. The external wrenches at the two end effectors are estimated from the measured joint torques. We validated the performance of our framework against two different baselines. We initially compared against a typical task-space impedance controller by teleoperating the robotic arms to violate the joint position limits. Later, we also compared with and without the motion adaptation capability by teleoperating the robot to violate the joint torque limits.

First, as a baseline, we implemented a typical impedance controller in Cartesian space with a force feedforward term for bimanual tasks. We disabled the contact force optimizer and the motion adaptation module, both realized by the task-space SEIKO. We set the parameters of the impedance controller to be the same as the ones used by the relative Cartesian PD controller in the section "Interaction Controller." We designed a series of teleoperation experiments using a heavy box (approximately 3 kg). Because the contact force optimizer was disabled, we had to set the value of the force feedforward term manually. Referring to the previously recorded contact force (between 40 and 50 N), we set the contact force to 45 N. We compared the performance of the proposed framework and the baseline impedance controller in experiments where the operator teleoperated the dual-arm robot to move both the position and the orientation of the heavy box. During a large range of commanded motions, the operator's commands for the dual-arm robot to move in a posture will violate the joint position limits at times.

Second, to validate how the task-space SEIKO affects the motion adaptation performance, we conducted teleoperation

experiments to move the same heavy box with SEIKO enabled and disabled. We misaligned the CoM of the box with the center of the two end effectors to create a large gravitational torque at each contact point. The torque will increase when rotating the box around the *x*-axis, making it easy to reach the robots' joint torque and friction cone limits. We manually set the torque limits of both shoulder (joint 2) joints to 90% of their maximum torques to protect the robot from damage caused by the violation of the joint torque limits. The operator teleoperated the dual-arm robot to move the box with translational and rotational motions. We initially tested the proposed method in simulation and visualized the adapted motions of moving the box along the *x*-axis and rotating the box around the *x*-axis (in the world frame), as shown in Figure 4. The green boxes indicate the object's original references calculated from arbitrary human commands, which were adapted to the feasible motions indicated by the transparent box.

Furthermore, we validated the extensive capability of the proposed method with collaborative bimanual manipulation tasks, including maneuvering a stack of objects via direct physical interaction, a multioperator part assembly task, and an industrial connector insertion task.

COMPARISON RESULTS

The snapshots in Figure 5 for the first comparison and in Figure 6 for the second comparison show the motions of the box when the operator teleoperated the robotic arms to move the box via two haptic devices.

The performance of our framework compared to the impedance controller baseline is shown in Figure 5. Using our framework, the system still held the object when the operator tried to move the robotic arms near the joint position limits. However, when the operator commanded the robotic arms near the joint position limits with the impedance controller, the end effectors failed to hold the object, and the system stopped working. This happened because our proposed framework adapted the commanded pose from the operator using the task-space SEIKO to satisfy the robot's task constraints and physical limits. Please find more details of these experiments in the accompanying video in the supplementary materials available at https://doi.org/10.1109/MRA.2023.3270222. It also shows that the impedance controller works for slow teleoperated motions within the feasibility boundary but becomes less robust for faster motions and fails, and the operator's commands are infeasible.

The performance of the motion adaptation module is validated on the real dual-arm robot with the same heavy box. For the comparison, we disabled the constraint enforcement module of the task-space SEIKO, preventing motions from being adapted to the physical limits of the robot. Without motion adaptation, while everything else was kept the same, slippage and crashing happened to the box during teleoperation, which can be seen in Figure 6(a) and (b). The robotic arms failed to apply enough contact forces when the shoulder joint torque limits were violated. In comparison, when the task-space SEIKO was fully enabled in the framework, the robot succeeded in holding the box, even when the commands from the operator violated the joint torque limits. The Cartesian trajectories of the CoM of the box including the human commands and the adapted motions are shown in Figure 6(c) and (d).

During the first 2 s of the translational experiment with adaption enabled, the right shoulder joint had reached its torque limit, while the joint torque of the left shoulder continued to increase. The Cartesian trajectories show that the robotic arms still followed human commands, even though one joint had reached its torque limit. This happens because the proposed motion adaptation can optimize and adjust the

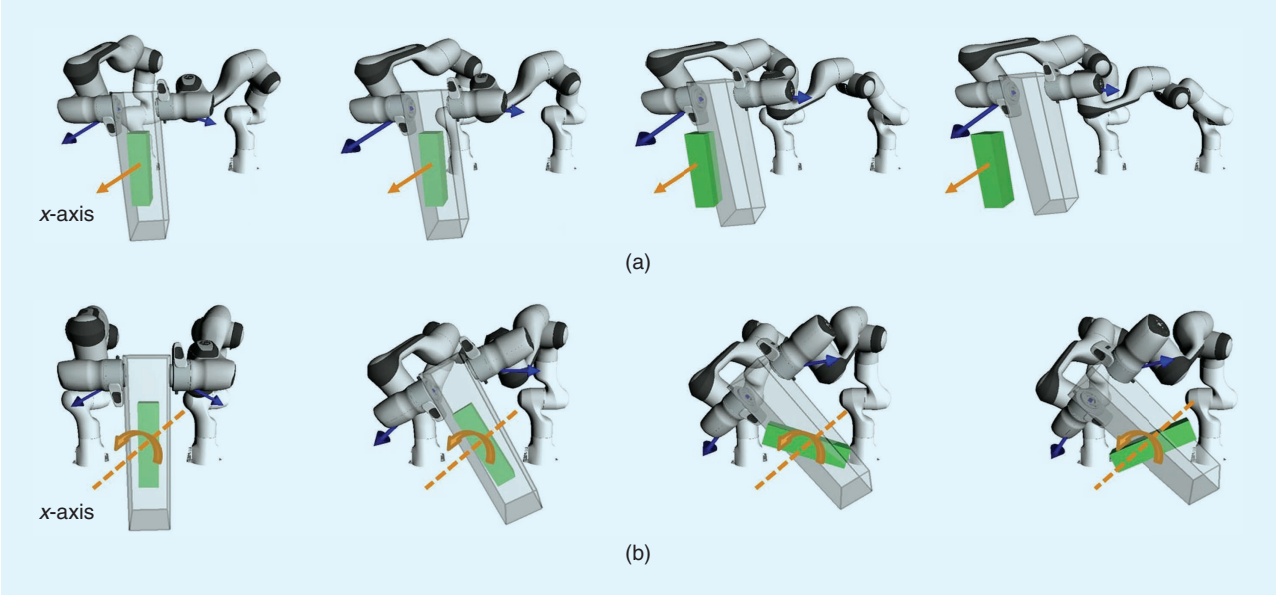

FIGURE 4. Validation of feasibility adaptation visualized by the original references (green boxes) of arbitrary human commands versus the optimized adaptations (transparent boxes) with contract forces (blue arrows) in (a) translational and (b) rotational tasks, respectively. The orange arrows indicate the directions of the translational and rotational motions.

joint-space configuration within the physical constraints of the robots. After 2 s, the shoulder joints of both arms reached their joint torque limits, and the joint-space configuration could not be further adjusted to follow human commands. The right two snapshots in Figure 6(c) and (d) show that the robot stopped moving the box, even though the operator was still sending forward commands through haptic devices.

During the first 3 s of the rotational experiment with adaption enabled, the robot could rotate the box to follow human commands, shown by the trajectories of the Euler angles in Figure 6(d). However, the human commands were adapted after the right shoulder joint reached its torque limit. The right two snapshots show that the box rotation was stopped, even though the human operator continued giving rotational commands. The success of this task validated that SEIKO is capable of real-time motion adaptation by optimizing the maximum rotational motions of the end effectors subject to joint torque limits and the static equilibrium by generating enough contact forces while keeping them within the friction cones.

FURTHER EXPERIMENTAL VALIDATIONS

An operator and the robot collaborated to move a stack of books via direct physical interaction, as shown in Figure 7. The tradeoff of the admittance and impedance behaviors is critical to the success and robustness of coordinated tasks. This adaptive tradeoff increases the robustness of the coordinated motions in the presence of model inaccuracy.

Two long-horizon complex manipulation tasks: multioperator part assembly and teleoperated industrial connector insertion, validate the extensive capability of the proposed collaborative framework.

Figure 8(a) shows the multioperator part assembly task completed by the collaboration of a remote operator, a local operator, and a dual-arm robot. The remote operator first teleoperated the two robotic arms using the independent mode to approach the part on the conveyor belt and then switched to the bimanual mode to maneuver it over the other part placed on the table. Then the task was handed over to the local operator, who interactively guided and adjusted the fine manipulation process of the pose of the part by

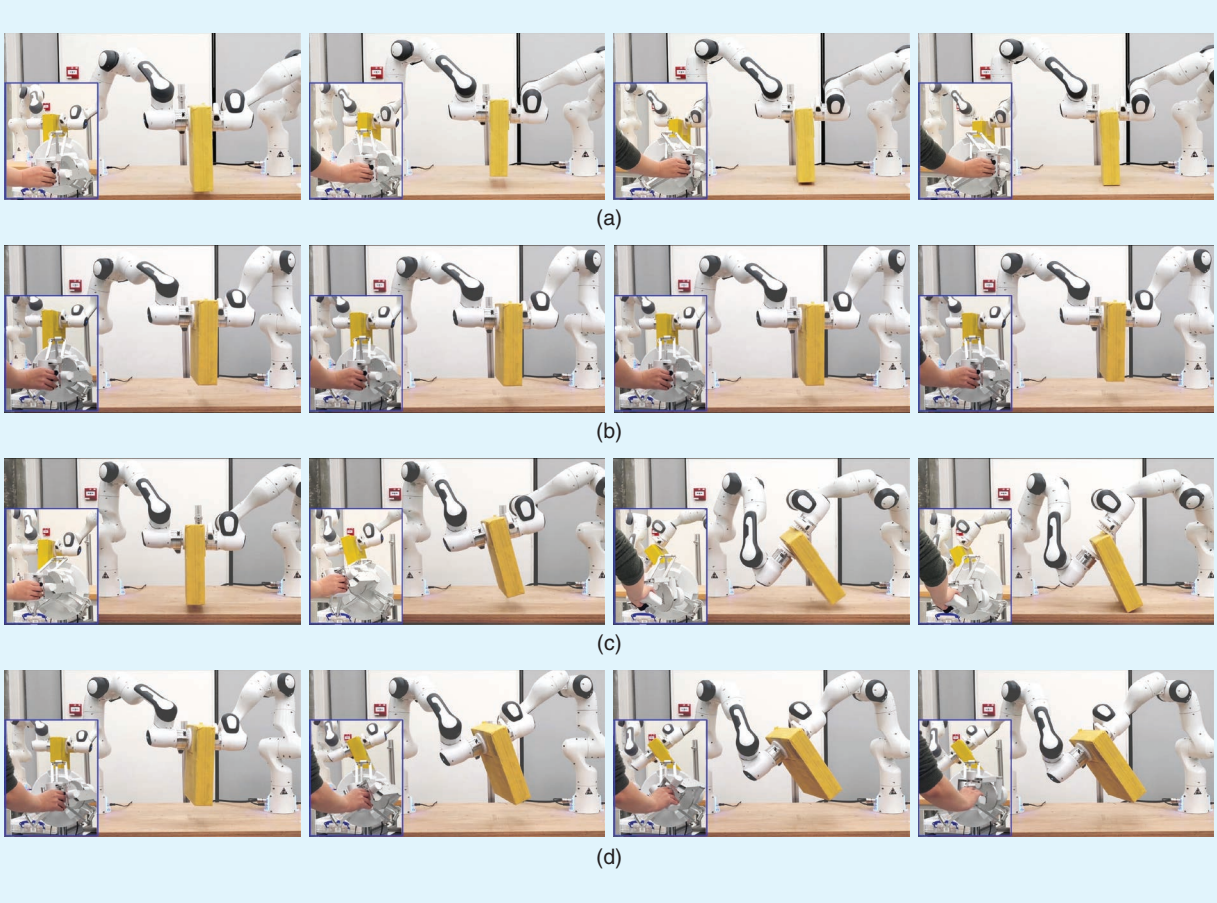

FIGURE 5. Proposed framework with the task-space SEIKO can successfully adapt the original Cartesian commands from the human operator to ensure the satisfaction of the joint position limits, while the baseline impedance controller fails at enforcing the equilibrium of contact wrenches of the box. (a) *Baseline*: The box fell when teleoperated to move along the *y*-axis and reached right arm joint position limits. (b) *Proposed framework with SEIKO*: The box remained stable when teleoperated to move along the *y*-axis and reached right arm joint position limits. (c) *Baseline*: The box fell when teleoperated to rotate along the *x*-axis and friction cone limits were reached. (d) *Proposed framework with SEIKO*: The box remained stable when teleoperated to rotate along the *x*-axis at friction cone limits.

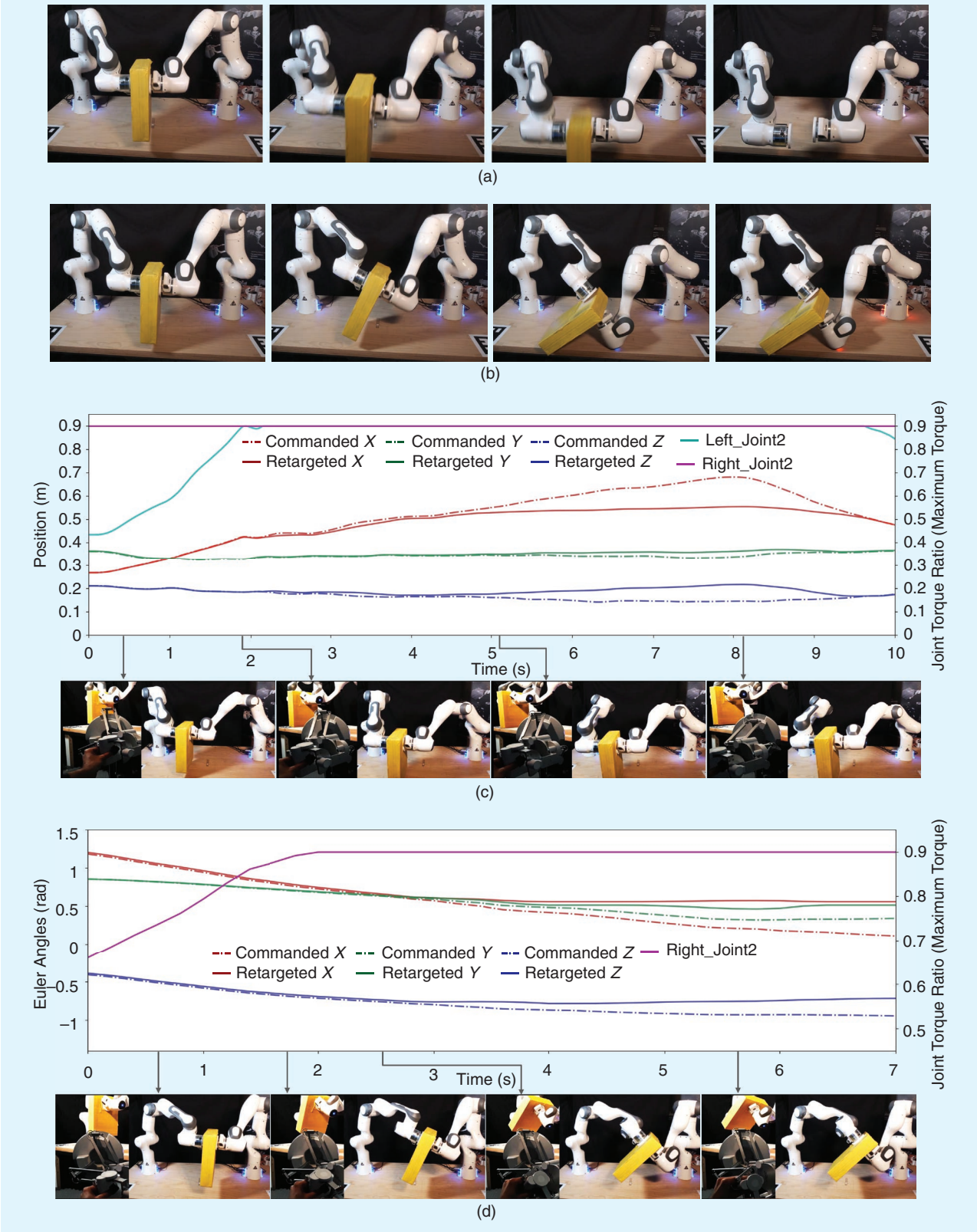

FIGURE 6. The Cartesian trajectories of the box CoM in the world frame along with the snapshots of the robotic arms and haptic devices during the teleoperation of translational and rotational motions. The data plots of the commanded (dashed lines) and the retargeted motions (solid lines) show that the proposed task-space SEIKO can adapt the human Cartesian commands to satisfy the joint torque limits and the box's equilibrium constraint. (a) Without motion adaptation, the box dropped when the arms were teleoperated to stretch forward and reached the shoulder joints' maximum torque limits. (b) Without motion adaptation, the box dropped when teleoperated to rotate, and shoulder joint torque limits prevented application of the correct contact force. (c) With motion adaptation, commands were retargeted not to stretch the arms beyond torque limits when teleoperated to move forward along the *x*-axis. (d) With motion adaptation, commands were retargeted not to tilt the box beyond torque limits when teleoperated to rotate the box around the *x*-axis.

physically interacting with it using the more detailed contact information.

Figure 8(b) shows the connector insertion task, where two customized end effectors were designed to manipulate a cylindrical connector. The connector insertion task includes three subtasks: opening the lid of the socket, which is partially covered by pebbles; unplugging the connector from an existing socket; and inserting the connector to the new socket. First one robotic arm was teleoperated to reach the socket covered by pebbles and open its lid. The dual-arm robot was then teleoperated to reach and grasp the connector and unplug it from the socket on the table. At last, the robot was teleoperated to maneuver the connector to approach the socket in the pebbles and complete the insertion task.

The success of the part assembly task validated that the proposed framework has achieved reliable HRC in teleoperation and direct physical interaction. The connector insertion task further validated that the proposed framework is compatible with dual-arm robots and customized effectors for more dexterous grasping and manipulation.

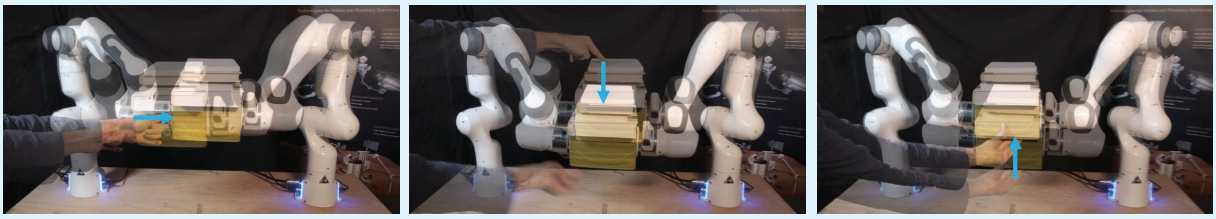

FIGURE 7. A local operator and dual-arm robot collaborated to move a stack of books translationally via physical interaction.

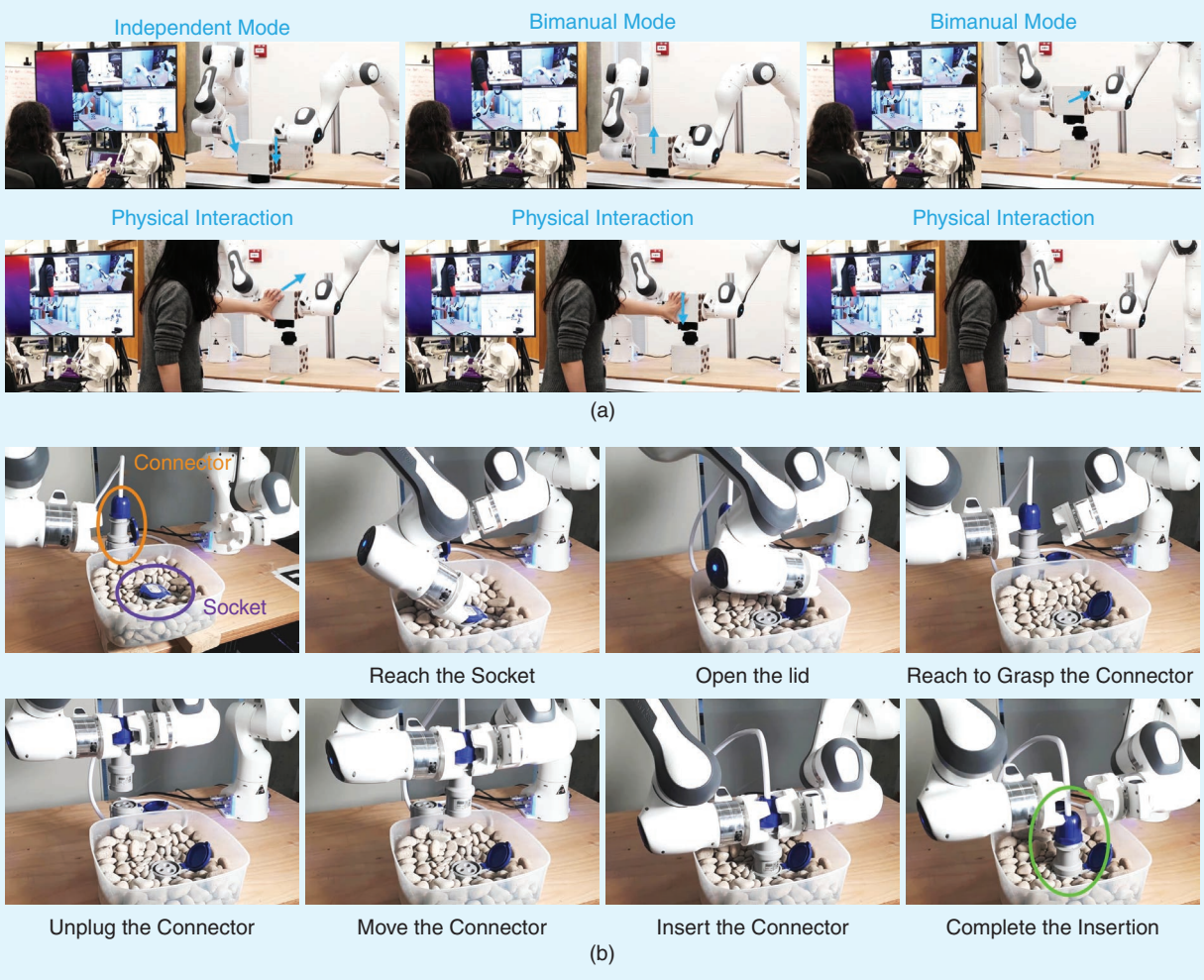

FIGURE 8. Part assembly and connector insertion tasks using collaborative bimanual manipulation. (a) A multioperator part assembly task. (b) A teleoperated industrial connector insertion task.

DISCUSSION

Our framework ensures feasibility for robots, objects, and human operators by respecting two types of constraints: equality and inequality constraints. The equality constraint is the static equilibrium of the object, and the inequality constraints include the robots' physical constraints in the joint space and the contact wrenches in the task space (more details are provided in the section "Motion Adaptation"). When given a lower bound of the friction coefficient and an upper bound of the object's mass, our model-based motion adaptation framework optimizes the desired joint positions, end-effector poses, and wrenches with respect to the previous constraints.

Uncertainties in real-world systems include 1) human commands, which are not always guaranteed to be realized by real robots; 2) model uncertainties, the model errors of the robot and the manipulated object from the ideal model used for optimization; 3) sensor noises, which affect the estimation and accuracy of the proprioceptive state of the system; and 4) actuation, referring to how accurately the mechatronics of the robot can execute the desired actions.

The proposed task-space SEIKO was formulated under the assumption of quasi-static motions. To ensure the validity of this assumption, we use a lowpass filter (Butterworth first order, cutoff frequency at 2 Hz) to remove jerky motion artifacts in human operator commands. We also use a velocity and acceleration filter to limit the maximum Cartesian linear velocity to 0.2 m/s and acceleration to 2 m/s^2 (the same applies to the angular components). A human operator's original commands will not always be exactly tracked by robots. Rather, the task-space SEIKO adapts human commands, e.g., those that exceed the workspace, to feasible motions, satisfying the task constraints and the robots' physical limits.

Our proposed interaction controller combines higher level admittance control with lower level impedance control, enabling online adjustment of the tradeoff between the wrench and pose tracking accuracy. Admittance controllers track the desired interaction wrench using an equivalent dynamics model, but accurate pose control cannot be guaranteed. By contrast, impedance controllers enable control of the desired pose but do not accurately track the interaction wrench [3], [7]. Our approach allows for online tuning of the admittance and impedance ports, thereby modulating their behavior. It enables the modulation of the dominant behavior by controlling the admittance port, while maintaining softness in the robot even when the velocity command, estimated from the physical interaction, is fully saturated due to retargeting constraints. This is possible because the impedance controller takes over in such cases.

An experiment using the linear passive Cartesian controller available in the Franka application programming interface was also performed, and the outcome is shown in the attached video in the supplementary materials available at https://doi.org/10.1109/MRA.2023.3270222. The passive impedance controller can perform the task in most conditions; however, the system fails when reaching challenging configurations (e.g., singularities) where the model errors reduce the motion accuracy of the robots. By contrast, the FIC can rely on its nonlinear behavior to preserve motion accuracy in the presence of model error, ensuring higher robustness compared to the linear controller. It is also worth mentioning that the FIC is an extension of the passive impedance controller to nonlinear force profiles. Its behavior converges to the linear controller when a linear force profile is chosen [3], [7], [8], [19].

CONCLUSIONS AND FUTURE WORK

This article presents a collaborative bimanual manipulation framework using optimal motion adaptation. The human Cartesian commands both from teleoperation and physical interaction are optimized by the task-space SEIKO to satisfy the physical limits of the robots and the contact constraints. The desired motions are realized by the interaction controller, which combines independent controllers to generate the desired joint torque commands to control the two robotic arms. The experimental results validate that the proposed task-space formulation of SEIKO creates robust and feasible motions during interactive and collaborative bimanual manipulation.

Our system ensures feasibility constraints during simultaneous remote and local human–robot interactions and is not limited to the completion of grasping and manipulation tasks. Because of its flexibility in handling multiple inputs simultaneously, our system facilitates seamless collaboration between remote experts and local users, enabling the latter to provide additional assistance based on their on-site observations, despite potentially having lower skill levels. Our system is more flexible and versatile in challenging scenarios that require multiple operation modalities and multiple operators. In the future, we will explore more use cases that would benefit from our system's capability of handling simultaneous inputs from teleoperation and physical interaction. For example, in teleoperation scenarios, such as remote surgery and space maintenance, experts who are located remotely can only operate through a teleoperation system, while local personnel who have less experience can provide basic adjustments or guidance as needed through physical interactions using better on-site observations.

The usage of the customized end effectors of the Franka robotic arms has achieved the success of bimanual manipulation tasks for the insertion task of a cylindrical connector. We will also explore the possibility of applying the motion adaption method with customized and/or commercial grippers for dexterous manipulation tasks for manufacturing and medical applications.

ACKNOWLEDGMENT

This research is supported by the Engineering and Physical Sciences Research Council Future AI and Robotics for Space (EP/R026092), ORCA (EP/R026173), National Centre for Nuclear Robotics (EP/R02572X), EU Horizon 2020 project THING (ICT-2017-1), and EU Horizon 2020 project Harmony (101017008). This article has supplementary

downloadable material available at https://doi.org/10.1109/MRA.2023.3270222, provided by the authors.

AUTHORS

Ruoshi Wen, Institute for Perception, Action, and Behaviour, School of Informatics, University of Edinburgh, EH8 9AB Edinburgh, U.K. E-mail: ruoshiwen@gmail.com.

Quentin Rouxel, Institute for Perception, Action, and Behaviour, School of Informatics, University of Edinburgh, EH8 9AB Edinburgh, U.K. E-mail: leph.quentinrouxel@gmail.com

Michael Mistry, Institute for Perception, Action, and Behaviour, School of Informatics, University of Edinburgh, EH8 9AB Edinburgh, U.K. E-mail: mmistry@inf.ed.ac.uk.

Zhibin Li, Department of Computer Science, University College London, WC1E 6BT London, U.K. E-mail: alex.li@ucl.ac.uk

Carlo Tiseo, Institute for Perception, Action, and Behaviour, School of Informatics, University of Edinburgh, EH8 9AB Edinburgh, U.K., and School of Engineering and Informatics, University of Sussex, BN1 9RH Brighton, U.K. E-mail: carlotiseo@gmail.com.

REFERENCES

[1] A. Weiss, A.-K. Wortmeier, and B. Kubicek, "Cobots in industry 4.0: A roadmap for future practice studies on human–robot collaboration," *IEEE Trans. Human-Mach. Syst.*, vol. 51, no. 4, pp. 335–345, Aug. 2021, doi: 10.1109/THMS.2021.3092684.

[2] D. Szafir, B. Mutlu, and T. Fong, "Designing planning and control interfaces to support user collaboration with flying robots," *Int. J. Robot. Res.*, vol. 36, nos. 5–7, pp. 514–542, Jun. 2017, doi: 10.1177/0278364916688256.

[3] C. Tiseo et al., "HapFIC: An adaptive force/position controller for safe environment interaction in articulated systems," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1432–1440, Jul. 2021, doi: 10.1109/TNSRE.2021.3098062.

[4] R. Wen, K. Yuan, Q. Wang, S. Heng, and Z. Li, "Force-guided high-precision grasping control of fragile and deformable objects using SEMG-based force prediction," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 2762–2769, Apr. 2020, doi: 10.1109/LRA.2020.2974439.

[5] M. Selvaggio, J. Cacace, C. Pacchierotti, F. Ruggiero, and P. Giordano, "A shared-control teleoperation architecture for nonprehensile object transportation," *IEEE Trans. Robot.*, vol. 38, no. 1, pp. 569–583, Feb. 2022, doi: 10.1109/TRO.2021.3086773.

[6] Q. Rouxel, K. Yuan, R. Wen, and Z. Li, "Multicontact motion retargeting using whole-body optimization of full kinematics and sequential force equilibrium," *IEEE/ASME Trans. Mechatron.*, vol. 27, no. 5, pp. 4188–4198, Oct. 2022, doi: 10.1109/TMECH.2022.3152844.

[7] K. K. Babarahmati, C. Tiseo, J. Smith, H. C. Lin, M. S. Erden, and M. Mistry, "Fractal impedance for passive controllers: A framework for interaction robotics," *Nonlinear Dyn.*, vol. 110, no. 3, pp. 2517–2533, Nov. 2022, doi: 10.1007/s11071-022-07754-3.

[8] C. Tiseo, W. Merkt, W. Wolfslag, S. Vijayakumar, and M. Mistry, "Safe and compliant control of redundant robots using superimposition of passive task-space controllers," 2020, *arXiv:2002.12249*.

[9] A. Cherubini, R. Passama, A. Crosnier, A. Lasnier, and P. Fraisse, "Collaborative manufacturing with physical human–robot interaction," *Robot. Comput.-Integr. Manuf.*, vol. 40, pp. 1–13, Aug. 2016, doi: 10.1016/j.rcim.2015.12.007.

[10] A. Kaplish and K. Yamane, "Motion retargeting and control for teleoperated physical human-robot interaction," in *Proc. IEEE-RAS 19th Int. Conf. Humanoid Robots (Humanoids)*, 2019, pp. 723–730, doi: 10.1109/Humanoids43949.2019.9035060.

[11] N. Hogan and S. P. Buerger, "Impedance and interaction control," in *Robotics and Automation Handbook*. Boca Raton, FL, USA: CRC Press, 2018, pp. 375–398.

[12] A. Q. Keemink, H. van der Kooij, and A. H. Stienen, "Admittance control for physical human–robot interaction," *Int. J. Robot. Res.*, vol. 37, no. 11, pp. 1421–1444, Sep. 2018, doi: 10.1177/0278364918768950.

[13] M. Khoramshahi and A. Billard, "A dynamical system approach to task-adaptation in physical human–robot interaction," *Auton. Robots*, vol. 43, no. 4, pp. 927–946, Apr. 2019, doi: 10.1007/s10514-018-9764-z.

[14] D. Rakita, B. Mutlu, M. Gleicher, and L. M. Hiatt, "Shared control–based bimanual robot manipulation," *Sci. Robot.*, vol. 4, no. 30, May 2019, Art. no. eaaw0955, doi: 10.1126/scirobotics.aaw0955.

[15] N. Mansard, O. Khatib, and A. Kheddar, "A unified approach to integrate unilateral constraints in the stack of tasks," *IEEE Trans. Robot.*, vol. 25, no. 3, pp. 670–685, Jun. 2009, doi: 10.1109/TRO.2009.2020345.

[16] A. Escande, N. Mansard, and P.-B. Wieber, "Hierarchical quadratic programming: Fast online humanoid-robot motion generation," *Int. J. Robot. Res.*, vol. 33, no. 7, pp. 1006–1028, Jun. 2014, doi: 10.1177/0278364914521306.

[17] N. Dehio, J. Smith, D. L. Wigand, P. Mohammadi, M. Mistry, and J. J. Steil, "Enabling impedance-based physical human–multi–robot collaboration: Experiments with four torque-controlled manipulators," *Int. J. Robot. Res.*, vol. 41, no. 1, pp. 68–84, Nov. 2021, doi: 10.1177/02783649211053650.

[18] L. Yan, T. Stouraitis, and S. Vijayakumar, "Decentralized ability-aware adaptive control for multi-robot collaborative manipulation," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 2311–2318, Feb. 2021, doi: 10.1109/LRA.2021.3060379.

[19] C. Tiseo, Q. Rouxel, Z. Li, and M. Mistry, "Fine manipulation and dynamic interaction in haptic teleoperation," 2021, *arXiv:2109.04524*.

[20] R. Featherstone, *Rigid Body Dynamics Algorithms*. New York, NY, USA: Springer-Verlag, 2014.