

Consolidating Kinematic Models to Promote Coordinated Mobile Manipulations

Ziyuan Jiao^{1*} Zeyu Zhang^{1*} Xin Jiang³ David Han² Song-Chun Zhu¹ Yixin Zhu¹ Hangxin Liu¹

Abstract—We construct a Virtual Kinematic Chain (VKC) that readily consolidates the kinematics of the mobile base, the arm, and the object to be manipulated in mobile manipulations. Accordingly, a mobile manipulation task is represented by altering the state of the constructed VKC, which can be converted to a motion planning problem, formulated and solved by trajectory optimization. This new VKC perspective of mobile manipulation allows a service robot to (i) produce well-coordinated motions, suitable for complex household environments, and (ii) perform intricate multi-step tasks while interacting with multiple objects without an explicit definition of intermediate goals. In simulated experiments, we validate these advantages by comparing the VKC-based approach with baselines that solely optimize individual components. The results manifest that VKC-based joint modeling and planning promote task success rates and produce more efficient trajectories.

I. INTRODUCTION

Mobile manipulation is a core capability for a service robot to function properly in household environments and excel in various tasks. However, since service robots usually come with a bulky mobile base along with a large base footprint, they often struggle in household environments due to three unique challenges: (i) An indoor space is confined and clustered, constraining the robot’s locomotion and posing additional challenges for mobile manipulation. (ii) The majority of tasks involve manipulating objects with diverse structures (*e.g.*, articulated objects like doors and drawers), challenging to generate motion plans for mobile manipulators with a unified approach. (iii) The base and arm’s movements have to be coordinated to ensure efficient and safe operations during mobile manipulation.

There exists research efforts focusing on each of the above challenges, *i.e.*, motion planning in confined space [1–3], control or learning frameworks for opening doors and drawers [4–7], and whole-body planning for foot-arm coordination [8–10]. However, a unified approach that tackles all three challenges altogether in household settings is still mostly missing. Consequently, a service robot’s mobile manipulation skills are far from ideal in terms of efficacy or fluency.

In stark contrast, humans possess fluid manipulation skills and interact with an environment efficiently. Cognitive psychologies and philosophers have proposed a theory of body schema [11]: Humans maintain a body’s representation during their motions and interactions with the environment;



Fig. 1: **Diverse interactions a service robot needs to perform in a household environment.** By abstracting the objects’ kinematic structures and forming a VKC, a service robot can plan and act more efficiently with improved foot-arm coordination.

this representation is malleable and can be extended to incorporate external objects. By treating the manipulated object as part of the extended limb, the theory of body schema provides a plausible account for why humans excel in complex manipulation tasks, from picking and placing an object, to opening doors and drawers, to tool=use [12]. Although the idea of the body schema has been introduced to the robotics community to represent robot structures and guide robot’s behaviors [13], it has left untouched whether the theory of body schema would promote a service robot’s (mobile manipulation in particular) planning and execution skills in complex manipulation tasks. And if it does, what would be a proper representation at a computational level?

To answer this question, we propose to abstract the object being manipulated—rigid objects and constrained mechanisms (doors, drawers, *etc.*)—by its kinematic structure. We integrate the kinematics of the robot and the manipulated object by constructing a single kinematic chain using the idea of Virtual Kinematic Chain (VKC) [14]. Specifically, the kinematic structure of the manipulated object is augmented to the manipulator by connecting the end-effector to an attachable location of the object. A virtual transformation between the mobile base and the virtual base link is also augmented by incorporating its navigation information.

From this new VKC-based perspective, a mobile manipulation task in a household environment is represented by altering the state or the structure of the VKC, which leads to a motion planning problem on VKC, formulated and solved by trajectory optimization. This new perspective enables a service robot to plan and act efficiently by allowing it to

* Ziyuan Jiao and Zeyu Zhang contributed equally to this work.

¹ UCLA Center for Vision, Cognition, Learning, and Autonomy (VCLA) at Statistics Department. Emails: {zyjiao, zeyuzhang, yixin.zhu, hx.liu}@ucla.edu, sczhu@stat.ucla.edu.

² Drexel University, Department of Electrical and Computer Engineering. Email: dkh42@drexel.edu.

³ UCLA ECE Department. Email: jiangxjames@ucla.edu.

The work reported herein was supported by ONR N00014-19-1-2153, ONR MURI N00014-16-1-2007, and DARPA XAI N66001-17-2-4029.

directly incorporate external objects and plan the motion as a whole to achieve better foot-arm coordination; see examples in Fig. 1. In simulations, we validate the proposed VKC perspective in various mobile manipulation tasks. Our experiments show that the consolidated kinematic models are particularly suitable for service robots by alleviating intermediate goal definitions for motion planners; they promote coordinated motions among base, arm, and object.

A. Related Work

The idea of **Virtual Kinematic Chain (VKC)** could be traced back to 1997 by Pratt *et al.* [15] for bipedal robot locomotion [16]. This idea was later adopted to chain serial manipulators to form one kinematic chain [14] and to dual-arm manipulation tasks; for instance, connecting parallel structures via rigid-body objects [17], modeling whole-body control of mobile manipulators [18]. Recently, VKC is also adopted for wheeled-legged robot control [19]. In this paper, we further push the idea of VKC to a mobile manipulator and demonstrate its advantages in modeling and planning complex manipulation tasks in household environments.

Motion planning is among the largest and most fundamental fields in robotics. In essence, methods can be roughly categorized into three major doctrines: search-based (*e.g.*, A* [20], D* [21]), sampling-based (*e.g.*, RRT [22] and its variants [23, 24]), and trajectory optimization (*e.g.*, CHOMP [25], TrajOpt [26]). We formulate the motion planning problem on VKC following the conventions in TrajOpt, as it incorporates kinematic constraints better than sampling-based methods while avoiding searching in large spaces.

Efficiently performing **mobile manipulation** tasks are challenging. Notable efforts have recently been dedicated to algorithms or system implementations, focusing on interactive manipulation tasks. For instance, equilibrium point control [4] and impedance control structure [27] are introduced to open doors and drawers. To improve efficiency, Gochev *et al.* used a heuristic-based method to reduce the search space [28]. Taking advantage of solving the inverse kinematics, Burget *et al.* proposed a whole-body motion planning approach for humanoid's constrained motion [29], and Bodily *et al.* proposed an algorithm for jointly optimizing a robot's base position and joint motions [9]. More recently, Toussaint *et al.* proposed a multi-bounded tree search algorithm to solve multi-step manipulation tasks involving tool-use [30]. Despite their promising results, prior arts primarily focus on a specific problem setup (*e.g.*, opening door and drawer, using tools). In comparison, the proposed approach rethinks mobile manipulation from a more general viewpoint using VKCs and tackles a broader range of tasks.

B. Overview

The remainder of this paper is organized as follows. Section II outlines notations and formally presents the problem definition. Section III details the VKC modeling. In a series of mobile manipulation tasks, we demonstrate the efficacy of VKCs with a high success rate in Section IV. We conclude the paper with discussions in Section V.

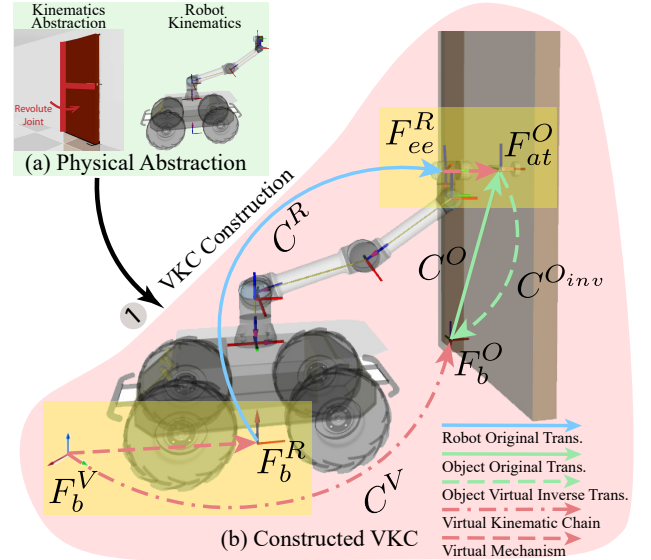


Fig. 2: **Overview of the mobile manipulation planning schematics using the proposed VKC-based approach.** (a) After abstracting out the underlying kinematics of the manipulated object and the mobile manipulator, (b) a VKC is constructed. The yellow boxes denote where the virtual connections are established: (i) One between \mathcal{F}_b^V and \mathcal{F}_b^R , the *virtual* base frame in the world coordinate and the robot's actual base frame, to reflect the navigational information, and (ii) another between \mathcal{F}_{ee}^R and \mathcal{F}_{at}^O , the robot's end-effector frame and the attachable frame of the object, to transfer effects of the manipulator to the manipulated object.

II. NOTATIONS AND PROBLEM DEFINITION

This section introduces the notations throughout the paper and the problem setup describing a mobile manipulation task.

The physical properties and kinematics of links and joints are defined following the Unified Robot Description Format (URDF) in Robot Operating System (ROS) and organized in a tree representation \mathcal{T} . Table I lists all the related notations:

TABLE I: Notations used for constructing VKCs.

Group	Notation	Description
Robot	\mathcal{T}^R	A tree represents the robot kinematic model
	\mathcal{F}_b^R	Robot base link's frame; the root of \mathcal{C}^R
	\mathcal{F}_{ee}^R	Robot end-effector link's frame
	\mathcal{C}^R	$\subset \mathcal{T}^R$, a kinematic chain from \mathcal{F}_b^R to \mathcal{F}_{ee}^R
	\mathcal{F}_i^R	Frame of link i in the kinematic chain \mathcal{C}^R
Object	\mathcal{T}^O	A tree represents the object kinematic model
	\mathcal{F}_b^O	Object base link's frame; the root of \mathcal{T}^O
	\mathcal{F}_{at}^O	Object attachable link's frame
	\mathcal{C}^O	$\subset \mathcal{T}^O$, a kinematic chain from \mathcal{F}_b^O to \mathcal{F}_{at}^O
Others	\mathcal{C}^V	A serial VKC with n Degree of Freedom (DoF)
	\mathbf{q}	$\in \mathbb{R}^n$, the state of VKC in joint space
	\mathbf{g}	$\in \mathbb{R}^k$ ($k \leq n$), the joint goal state
	${}^a_b T$	A homogeneous transformation from \mathcal{F}_a to \mathcal{F}_b
	${}^w_i T_g$	The goal pose of \mathcal{F}_i in the world frame

Below, we further summarize the above notations:

- The group *Robot* refers to notations related to the mobile manipulator, which consists of three components: mobile base, manipulator, and end-effector.
- The group *Object* refers to notations related to the manipulated objects, which could be as simple as a *rigid*

link or be an *articulated* object with two or more links connected by either a prismatic, revolute, or fixed joint. We introduce a *virtual* joint defined as an *attachment*, a local transformation ${}^{at}_{ee}T$ from the object's attachable frame \mathcal{F}_{at}^O (*i.e.*, the link a mobile manipulator can grasp on) to the robot's end-effector frame \mathcal{F}_{ee}^R .

- The group *Others* refers to constructed VKC, its state space, and other related notations in a manipulation task.

Constructing a VKC \mathcal{C}^V requires the inputs of robot kinematic tree \mathcal{T}^R , object kinematic tree \mathcal{T}^O , and transformation from an object attachable frame to the robot end-effector frame ${}^{at}_{ee}T$. The chain's forward kinematics (FK), inverse kinematics (IK), and Jacobians can be effectively solved by existing kinematic solvers (*e.g.*, KDL [31]).

Assuming a rigid connection between the end-effector and the attachable link during manipulation, performing a mobile manipulation task can be regarded as reaching desired VKC poses. As a result, we treat a mobile manipulation task as a motion planning problem on the VKC and solve it by trajectory optimization. Formally, it is equivalent to finding a collision-free path $\mathbf{q}_{1:T}$ from the initial pose \mathbf{q}_{init} to goals \mathbf{g} in joint space and/or goal poses w_iT_g in Euclidean space.

The objective function of the trajectory optimization can be formally expressed as:

$$\min_{\mathbf{q}_{1:T}} \sum_{t=1}^{T-1} \|W_{vel}^{1/2} \delta \mathbf{q}_t\|_2^2 + \sum_{t=2}^{T-1} \|W_{acc}^{1/2} \delta \dot{\mathbf{q}}_t\|_2^2, \quad (1)$$

wherein we penalize the overall weighted squared traveled distance of every joint with the finite forward difference $\delta \mathbf{q}_t \approx \mathbf{q}_{t+1} - \mathbf{q}_t$ and overall smoothness of the trajectory with the second-order finite central difference $\delta \dot{\mathbf{q}}_t \approx \mathbf{q}_{t-1} - 2\mathbf{q}_t + \mathbf{q}_{t+1}$. W_{vel} and W_{acc} are diagonal weight matrices for each joint, respectively. $\mathbf{q}_{1:T}$ represents the trajectory sequence $\{q_1, q_2, \dots, q_T\}$, where \mathbf{q}_t denotes the VKC state at the t^{th} time step.

III. VKC MODELING

The proposed VKC modeling constructs a serial kinematic chain by (i) incorporating both robot and object kinematics via a virtual joint and (ii) augmenting a virtual base to the robot base; see Fig. 2b for a graphic illustration.

A. VKC Construction

Below we formally describe the 4-step procedure of constructing the VKC, \mathcal{C}^V , by consolidating the robot and the object kinematics models.

a) Original Structure: The kinematic models of the mobile manipulator \mathcal{T}^R and the manipulated object \mathcal{T}^O are assumed given by the perception module or by the simulator.

b) Kinematic Inversion: Let us take the task of opening a door as an example. In conventional kinematic notation, the door is the child link, and the door frame is its parent link in the original \mathcal{T}^O . To construct a VKC, this parent-child relationship needs to be inverted before it can be attached to the robot's end-effector, *i.e.*, the door becomes the parent link that “transforms” the door frame. Of note, such an inversion also requires updating the joint connecting the two links,

since a joint (*i.e.*, revolute/prismatic) typically constrains the child link's motion w.r.t. the *child link's* frame.

c) VKC Construction: After inverting the original \mathcal{T}^O , a virtual joint between $\mathcal{F}_b^{O_{inv}}$ and \mathcal{F}_{ee}^R is inserted, whose transformation is denoted as ${}^{ee}_{at}T$. In our application, the transformation of the virtual joint is updated by the actual grasping pose right before the VKC construction to minimize kinematic discrepancies introduced by the execution error. Next, the motion planner will be invoked to plan following motions for the actual VKC. The joint type could also be determined by the grasping type between the gripper and the object (*e.g.*, revolute joint for grasping a cylindrical handle, fixed joint for grasping a rigid ball) to alleviate the inaccuracies during the execution.

d) Virtual Base Frame: A virtual base frame \mathcal{F}_b^V is further added and connected to the mobile base through two perpendicular prismatic joints and a revolute joint, enabling the mobile base's omnidirectional motions on the ground plane.

After the above procedure, the constructed VKC remains in serial and forms an equality constraint to Eq. (1):

$$h_{\text{chain}}(\mathbf{q}_t) = 0, \quad \forall t = 1, 2, \dots, T \quad (2)$$

It specifies the kinematics of the VKC, which includes its forward kinematics and other physical constraints of the manipulated object; *e.g.*, the manipulated object is fixed to the ground, which leads to a closed chain: ${}^w_bT_{1:T}^O - {}^w_bT^O = 0$. Failing to account for this constraint may damage the manipulated object or the mobile manipulator.

B. Goals

The goal of the mobile manipulation can be formulated as an inequality constraint, in addition to the equality constraint introduced by the VKC construction in Eq. (2):

$$\|f_{\text{task}}(\mathbf{q}_T) - \mathbf{g}\|_2^2 \leq \xi_{\text{goal}}, \quad (3)$$

which bounds the squared l_2 norm between the final state in the goal space $f_{\text{task}}(\mathbf{q}_T)$ and the goal state \mathbf{g} with a tolerance ξ_{goal} . The function $f_{\text{task}}: \mathbb{R}^n \rightarrow \mathbb{R}^k$ is a task-dependent function that maps the joint space of a VKC to the goal space that differs from task to task.

Again, let us take the example of opening a door. In the first phase when the robot is reaching the door handle, $f_{\text{task}}(\cdot)$ maps the joint space of a VKC to the robot's end-effector pose. In this case, the goal \mathbf{g} is the robot's end-effector pose ${}^w_{ee}T_g$, and Eq. (3) can be rewritten in a simplified form $\|f_{\text{fk}}(\mathbf{q}_T) - {}^w_iT_g\|_2^2 \leq \xi_{\text{goal}}$. In the second phase when the robot is opening the door, $f_{\text{task}}(\cdot)$ maps VKC's joint space to the joint of the door's revolute axis. Hence, \mathbf{g} is merely the angle θ of the revolute joint, and the trajectory of the other joints in the VKC are implicitly generated by the optimization process, together with obstacle avoidance and trajectory smoothing. Of note, Eqs. (2) and (3) are not the only forms of constraints that a VKC-based approach can incorporate; in fact, it is straightforward to add additional task constraints to the same optimization problem in Eq. (1), depending on various task-specific requirements.

C. Additional Constraints

During the trajectory optimization, we further impose several safety constraints. Without loss of generality, we assume an omnidirectional base and purely kinematic constraints in this paper. However, extra constraints, such as nonholonomic constraints for non-omnidirectional mobile bases or dynamic constraints for arms, could be formulated into the optimization problem by incorporating additional time, first-order, or second-order terms [32].

$$\mathbf{q}^{\min} \leq \mathbf{q}_t \leq \mathbf{q}^{\max}, \quad \forall t = 1, 2, \dots, T \quad (4)$$

$$\|\delta \mathbf{q}_t\|_{\infty} \leq \xi_{\text{vel}}, \|\delta \dot{\mathbf{q}}_t\|_{\infty} \leq \xi_{\text{acc}}, \quad \forall t = 2, 3, \dots, T-1 \quad (5)$$

$$\sum_{i=1}^{N_{\text{link}}} \sum_{j=1}^{N_{\text{obj}}} |\text{dist}_{\text{safe}} - f_{\text{dist}}(L_i, O_j)|^+ \leq \xi_{\text{dist}}, \quad (6)$$

$$\sum_{i=1}^{N_{\text{link}}} \sum_{j=1}^{N_{\text{link}}} |\text{dist}_{\text{safe}} - f_{\text{dist}}(L_i, L_j)|^+ \leq \xi_{\text{dist}}. \quad (7)$$

Eq. (4) is an inequality constraint that defines joint limits, in which \mathbf{q}^{\min} and \mathbf{q}^{\max} specify the lower and upper bound of every joint, respectively. Eq. (5) is an inequality constraint that bounds the joint velocity by ξ_{vel} and the joint acceleration by ξ_{acc} to obtain a feasible trajectory that can be executed without saturation. $\|\cdot\|_{\infty}$ denotes the infinity norm.

Eqs. (6) and (7) are inequality constraints that check link-object collisions and link-link collisions, respectively, where N_{link} and N_{obj} are the number of links and the number of objects, respectively. $\text{dist}_{\text{safe}}$ is a pre-define safety distance, and $f_{\text{dist}}(\cdot)$ is a function that calculates the signed distance [26] between i -th link L_i and j -th object O_j ; the function $|\cdot|^+$ is defined as $|x|^+ = \max(x, 0)$.

The inequality constraints introduced by Eqs. (6) and (7) make the preceding optimization problem highly non-convex and unsolvable by a generic convex solver. In this paper, we approximate it by a sequence of convex problems [26], solved by a sequential convex optimization method.

D. Advantages

As formally derived in the above sections, solving mobile manipulation as trajectory optimization using the proposed VKC-based approach introduces two advantages:

- 1) **Eliminating unnecessary intermediate goals.** Let us use the example of opening a door: Only one goal—the door's angle to be opened to—is required. The final poses of the mobile base and the manipulator are directly produced during the trajectory optimization process without manually specifying unnecessary intermediate goals. Hence, the VKC-based approach provides versatility and simplicity for modeling mobile manipulation tasks.
- 2) **Coordinating locomotion and manipulation.** Using VKCs, the trajectory optimization jointly generates trajectories of the mobile base and the manipulator, producing coordinated locomotion and manipulation, which is oftentimes challenging for conventional methods.

These two advantages are crucial for a service robot operating in a household environment. Below, we demonstrate these advantages in a series of mobile manipulation tasks.

IV. MOTION PLANNING ON VKCS

In this section, we start with simulation setup, followed by the evaluation of the trajectory-optimization-based motion planning on VKC from three perspectives: (i) the necessary trajectory initialization required by the motion planner, (ii) the improvement on base-arm coordination using the VKC approach compared with traditional setups that have to plan their motions separately, and (iii) the capability of operating in a household environment and performing various tasks.

A. Platform and Simulation Setup

The service robot platform we adopted for testing is a Universal Robot UR5e manipulator mounted on a Clearpath Husky A200 UGA; see Fig. 2b for a graphic illustration. The simulation environment is an arena with $10\text{m} \times 10\text{m}$ in size, discretized into 100×100 grids. All the experiments are conducted using a desktop with an Intel i7-9700K CPU, running with ROS-Industrial Tesseract [33].

B. Trajectory Initialization

Although a gradient-descent-based algorithm can effectively solve the trajectory optimization problem on VKCs, it may also be easily stuck at local minima near the given initial trajectory [25, 26]. As a result, a proper trajectory initialization is favored to improve the optimization result. Two primary trajectory initialization methods [34] are:

- 1) **Stationary:** The trajectory $\mathbf{q}_{1:T}$ is initialized by waypoints \mathbf{q}_t that are the same as the initial pose \mathbf{q}_{init} .
- 2) **Interpolated:** The trajectory $\mathbf{q}_{1:T}$ is initialized by waypoints that are linearly interpolated between the initial pose \mathbf{q}_{init} and the goal pose (not \mathbf{g} or ${}^w_{ee}T_g$; see below).

In this paper, we further devise an A*-based trajectory initialization method, which adopts A* to search for a feasible path given the initial and the goal pose of the mobile base. Next, we investigate how different trajectory initialization methods affect the planning results on VKC in three 3D virtual scenarios; see Figs. 3a to 3c. The robot's task is to pick up the rigid stick and use it to reach a target indicated by the red cube. This task consists of three steps: navigate to the stick, manipulate the stick and pick it up, and navigate to and reach the target with the stick. The three scenarios designed for evaluation are in increasing complexity: no obstacle (Fig. 3a), two small obstacles (Fig. 3b), or a much larger one (Fig. 3c). Experimental results reported below are the average of 50 different initial poses, each with 10 times.

A successfully optimized trajectory is a converged result without violating any constraints (e.g., collisions). Fig. 3d compares success rates. When the environment is clean (Scenario 1), even the simplest *Stationary* trajectory initialization method performs well. When there is additional complexity introduced by the obstacles (Scenario 2), the *Stationary* method deteriorates, whereas the *Interpolated* method still maintains a high success rate. When the navigable space is significantly reduced (Scenario 3), only the proposed A*-based initialization method can consistently perform well to generate feasible plans. Taken together, experimental results indicate that combining the proposed A*-based initialization with the optimization-based motion planner can well handle the challenging motions that require combining navigation and manipulation in cluttered space with obstacle avoidance.

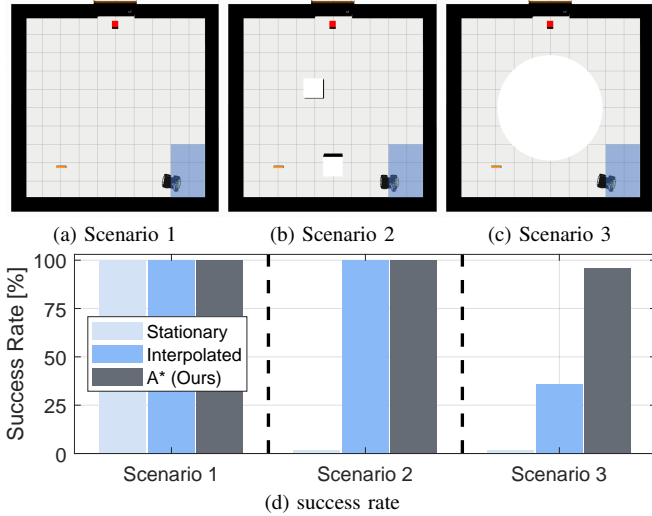


Fig. 3: **Comparisons of motion planning on VKCs by different trajectory initialization methods.** (a)-(c) The experimental scenarios in increasing complexity. The robot’s initial pose is uniformly sampled within the blue region; it is tasked to pick up the stick and use it to reach the red cube. (d) The proposed A*-based trajectory initialization has the highest success rates (almost always) in generating a feasible plan. In comparison, the *Stationary* method fails to generate feasible plans in scenario 2 and 3. Similarly, the *Interpolated* method struggles in scenarios 3 (30% success rate).

C. Comparisons with Baselines

We design two mobile manipulation tasks to validate the advantages of the proposed VKC modeling: (i) rotating the doorknob and pushing to open a door (with two revolute joints), and (ii) pulling to open a drawer (with one prismatic joint); see Figs. 4e and 4m. The robot starts from a uniformly sampled initial pose in the blue-shaded regions.

To benchmark the proposed VKC-based approach, we design two baselines that have to *individually* optimize the mobile base and mobile manipulator’s trajectories. Baseline 1 (**B1**): Use the proposed A*-based trajectory initialization to first plan the mobile base’s motion, and the arm pose is then found by solving the inverse kinematics from the door handle to the mobile base at each way-point. Baseline 2 (**B2**): On top of B1, the manipulated object and the manipulator’s poses are further refined at each time step to avoid collisions between the mobile manipulator and the environment.

Of note, our VKC-based approach only needs to specify one task goal—the desired door angle or the proper drawer length. In contrast, baselines require *additionally* specifying the mobile base’s pose when reaching the doorknob and after opening the door, as the base and the manipulator are planned and optimized individually. We compute these intermediate mobile base’s poses by sampling from feasible regions that are empirically found; see the pink areas in Figs. 4e and 4m for reaching and the green areas for final poses.

We evaluate the planning results by four criteria: (i) the percentage of task completion without violating the constraints as the *success rate*, (ii) the total base traveling distance as the *base’s effort*, (iii) the sum of each joint’s accumulated angular displacement during the entire task execution as the *manipulator’s effort*, and (iv) planning time.

We summarize results in Fig. 4. Planning individually (**B1**) yields a 48% and a 74% success rate for opening door and

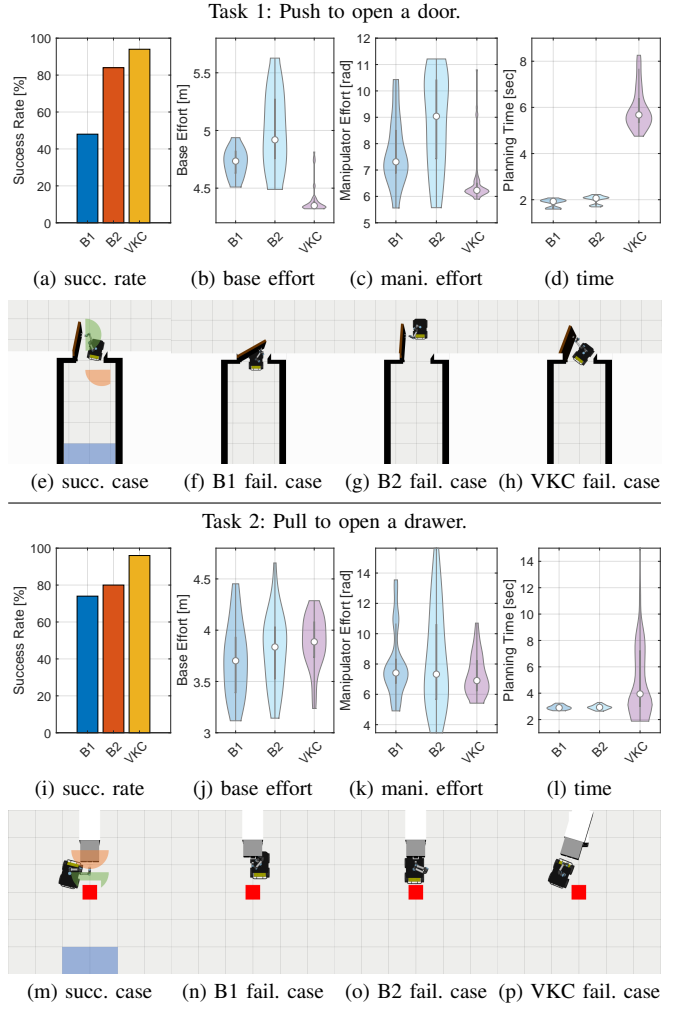


Fig. 4: **Quantitative comparisons between the VKC-based approach and two baselines on two mobile manipulation tasks: push to open a door, and pull to open a drawer.** Modeling a mobile manipulation using VKC leads to (a)(i) higher success rate and more effective trajectories in terms of smaller (b)(j) base and (c)(k) manipulator’s efforts, but requires (d)(l) longer planning time. Of note, baselines require additional definition of intermediate goals, which are often empirically found, (e)(m) shown in the green and pink shaded areas. (f)(n) Typical failure cases of two tasks in B1 are due to collision. Typical failure cases of two tasks in (g)(o) B2 or (h)(p) VKC-based approach are due to constraint violations.

drawer, respectively. The primary reason for failure is the collisions between the mobile manipulator and the door/drawer; see Figs. 4f and 4n. Although introducing a collision check to refine motions (**B2**) improves the success rate, the proposed VKC-based approach significantly outperforms the baselines. The failure cases of **B2** are mainly caused by violating velocity and acceleration constraints; see Figs. 4g and 4o.

Planning on VKCs yields more efficient motions with less base and manipulator efforts; see an instance in Fig. 6. Overall, the proposed VKC-based approach produces smoother trajectories with less fluctuation in the speed profile.

Taken together, our experimental results demonstrate that by using the proposed VKC-based approach, the planning of mobile manipulator avoids contrived definitions of intermediate poses and achieves a higher success rate with more efficient motions, at the cost of a slightly longer planning

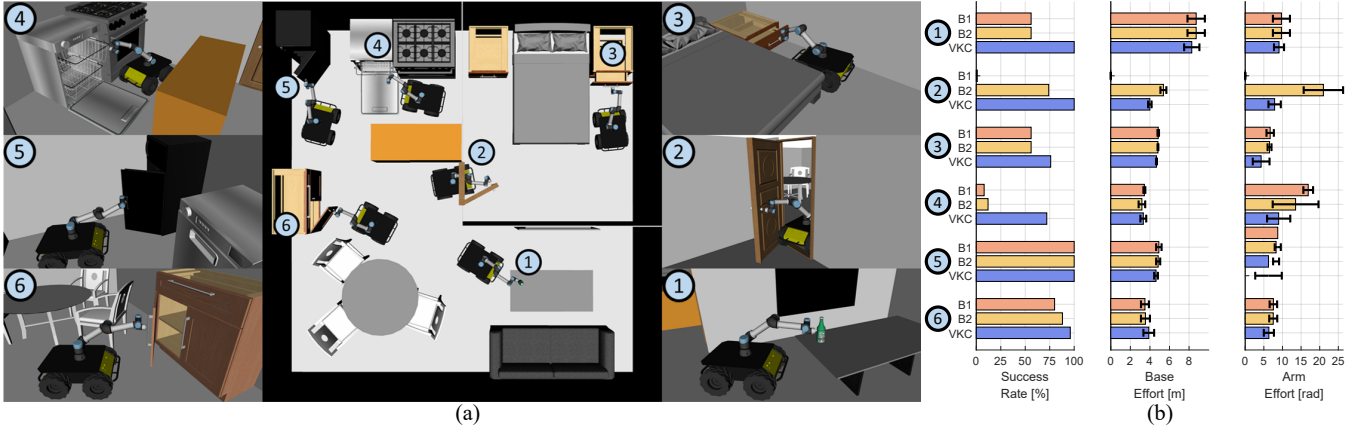


Fig. 5: **The VKC-based approach is advantageous in modeling and planning household mobile manipulation tasks for a service robot.** (a) Our simulated environments and the corresponding task performances. (b) The success rate, and base and arm efforts measured by their travel distance in 6 different mobile manipulation tasks. The black line segments denote the standard deviations. The proposed VKC-based approach generally leads to a higher success rate with less base and arm efforts.

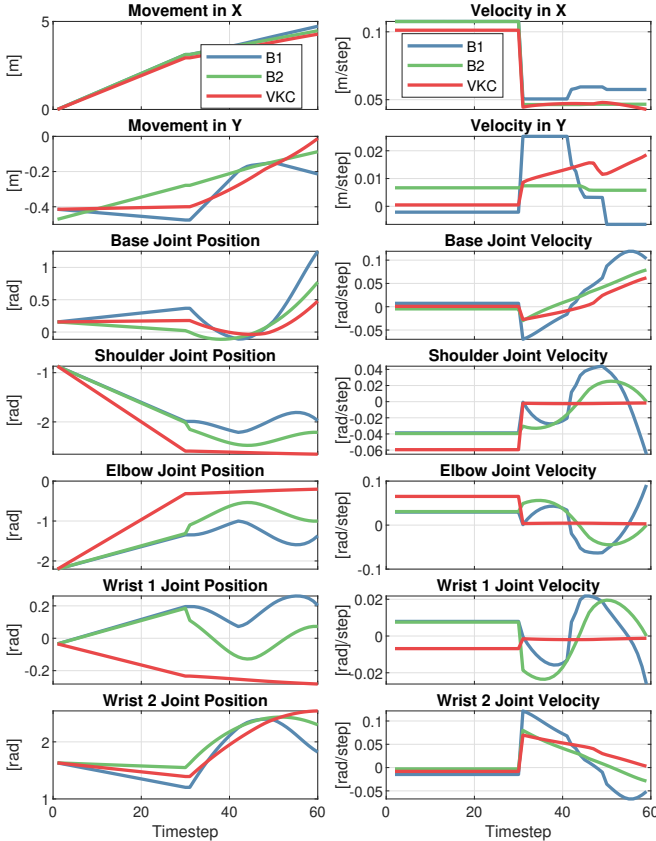


Fig. 6: **The joint position and velocity profiles of the door-opening task in Fig. 4e.** VKC-based approach generates more efficient trajectory with less fluctuation compared to the baselines.

time due to solving the trajectory optimization problem with higher DoFs compared to those in the baselines.

D. Complex Tasks

We further test the VKC modeling in a confined household environment with six daily manipulation tasks; see Fig. 5a.

- 1) Pick up an object randomly sampled on the coffee table (reachable by the robot) and place it on the dining table. **B2** performs the same as the **B1** since the object is fixed and does not collide with the robot.
- 2) Push to open a narrow door. The robot needs to coordinate

its foot and arm movements to open a narrow door (1.1 times wider than the robot) and pass through.

- 3) Open a drawer in the corner. Although space is confined, the robot only needs a backward motion, which should be relatively straightforward.
- 4) Pull out a dishwasher rack, similar to the above task and also in a confined space. However, it is more challenging as the robot has to manipulate to its side.
- 5) Pull to open a refrigerator.
- 6) Pull to open a cabinet, similar to the above task. However, since the cabinet has a larger door size, the chair in the environment will block the robot's movement.

In a successful trial, the robot should produce a collision-free trajectory without violating any kinematic or safety constraints. We compare the results in Fig. 5b with two baselines (**B1** and **B2**). Overall, the proposed VKC-based method demonstrates a higher success rate with less base and arm costs measured by the total distance traveled.

Specifically, in Task ②, **B1** can hardly find a path to go through the door because of the highly constrained free-space, which requires good foot-arm coordination; task ④ reveals a similar result. In Task ③, the low drawer height increases the difficulty for the mobile base to reach a feasible region; hence, both baselines underperform in terms of their success rate. An interesting observation is in Task ⑤ and ⑥: The baseline methods perform well in the pulling action, likely due to the empty space for the furniture doors to swing.

These results provide ample evidence of excellent foot-arm coordination produced by the proposed VKC-based approach, well-suited for confined household environments.

V. CONCLUSION AND DISCUSSION

We presented a modeling method that incorporates the kinematics of a robot's mobile base, arm, and the manipulated object in VKCs. From this new perspective, a mobile manipulation task is regarded as a motion planning problem on VKCs, solved by trajectory optimization. This approach alleviates the definition of intermediate goals and well coordinates base and arm movements, resulting in a higher success rate with more efficient trajectories in various mobile manipulation tasks. Our simulated experiments validate the advantages introduced by the proposed VKC-based modeling

approach, showing its potential in scaling up to complex operations in a household environment [35].

We discuss two issues related to the presented work in greater depth. To deploy the VKC-based method on a physical service robot, **objects' kinematic structures** have to be identified. Although some existing methods can achieve this goal by visual sensors [36–40], results could be error-prone, and a force/torque control framework should be implemented to handle kinematic discrepancies or uncertainties/errors in object positions. Additionally, an effective and **stable grasp** is a prerequisite of constructing a VKC connecting the robot and objects. While grasping remains an unsolved problem, more powerful machine learning-based methods are emerging [41–43], which could shed light on this direction.

REFERENCES

- [1] R. Buchanan, T. Bandyopadhyay, M. Bjelonic, L. Wellhausen, M. Hutter, and N. Kottege, "Walking posture adaptation for legged robot navigation in confined spaces," *Robotics and Automation Letters (RA-L)*, vol. 4, no. 2, pp. 2148–2155, 2019.
- [2] S. H. Cheong, B. Y. Cho, J. Lee, C. Kim, and C. Nam, "Where to relocate?: Object rearrangement inside cluttered and confined environments for robotic manipulation," in *Proceedings of International Conference on Robotics and Automation (ICRA)*, 2020.
- [3] Z. Han, J. Allspaw, G. LeMasurier, J. Parrillo, D. Giger, S. R. Ahmadzadeh, and H. A. Yanco, "Towards mobile multi-task manipulation in a confined and integrated environment with irregular objects," in *Proceedings of International Conference on Robotics and Automation (ICRA)*, 2020.
- [4] A. Jain and C. C. Kemp, "Pulling open doors and drawers: Coordinating an omni-directional base and a compliant arm with equilibrium point control," in *Proceedings of International Conference on Robotics and Automation (ICRA)*, 2010.
- [5] Y. Karayiannidis, C. Smith, F. E. Vina, P. Ogren, and D. Kragic, "'open sesame!': adaptive force/velocity control for opening unknown doors," in *Proceedings of International Conference on Intelligent Robots and Systems (IROS)*, 2012.
- [6] Y. Karayiannidis, C. Smith, F. E. V. Barrientos, P. Ögren, and D. Kragic, "An adaptive control approach for opening doors and drawers under uncertainties," *Transactions on Robotics (T-RO)*, vol. 32, no. 1, pp. 161–175, 2016.
- [7] J. Wang, S. Lin, C. Hu, Y. Zhu, and L. Zhu, "Learning semantic keypoint representations for door opening manipulation," *Robotics and Automation Letters (RA-L)*, vol. 5, no. 4, pp. 6980–6987, 2020.
- [8] K. Shankar, *Kinematics and Local Motion Planning for Quasi-static Whole-body Mobile Manipulation*. PhD thesis, California Institute of Technology, 2016.
- [9] D. M. Bodily, T. F. Allen, and M. D. Killpack, "Motion planning for mobile robots using inverse kinematics branching," in *Proceedings of International Conference on Robotics and Automation (ICRA)*, 2017.
- [10] S. Chitta, B. Cohen, and M. Likhachev, "Planning for autonomous door opening with a mobile manipulator," in *Proceedings of International Conference on Robotics and Automation (ICRA)*, 2010.
- [11] S. Gallagher, *How the body shapes the mind*. Clarendon Press, 2006.
- [12] N. P. Holmes and C. Spence, "Beyond the body schema: Visual, prosthetic, and technological contributions to bodily perception and awareness," *Human body perception from the inside out: Advances in visual cognition*, pp. 15–64, 2006.
- [13] M. Hoffmann, H. Marques, A. Arieta, H. Sumioka, M. Lungarella, and R. Pfeifer, "Body schema in robotics: a review," *IEEE Transactions on Autonomous Mental Development*, vol. 2, no. 4, pp. 304–324, 2010.
- [14] N. Likar, B. Nemec, and L. Žlajpah, "Virtual mechanism approach for dual-arm manipulation," *Robotica*, vol. 32, no. 6, 2014.
- [15] J. Pratt, P. Dilworth, and G. Pratt, "Virtual model control of a bipedal walking robot," in *Proceedings of International Conference on Robotics and Automation (ICRA)*, 1997.
- [16] J. Pratt, C.-M. Chew, A. Torres, P. Dilworth, and G. Pratt, "Virtual model control: An intuitive approach for bipedal locomotion," *International Journal of Robotics Research (IJRR)*, vol. 20, no. 2, pp. 129–143, 2001.
- [17] Y. Wang, C. Smith, Y. Karayiannidis, and P. Ögren, "Cooperative control of a serial-to-parallel structure using a virtual kinematic chain in a mobile dual-arm manipulation application," in *Proceedings of International Conference on Intelligent Robots and Systems (IROS)*, 2015.
- [18] Y. Wang, C. Smith, Y. Karayiannidis, and P. Ögren, "Whole body control of a dual-arm mobile robot using a virtual kinematic chain," *International Journal of Humanoid Robotics*, vol. 13, no. 01, 2016.
- [19] A. Laurenzi, E. M. Hoffman, M. P. Polverini, and N. Tsagarakis, "An augmented kinematic model for the cartesian control of the hybrid wheeled-legged quadrupedal robot centauRO," *Robotics and Automation Letters (RA-L)*, 2019.
- [20] P. E. Hart, N. J. Nilsson, and B. Raphael, "A formal basis for the heuristic determination of minimum cost paths," *IEEE Transactions on Systems Science and Cybernetics*, vol. 4, no. 2, pp. 100–107, 1968.
- [21] A. Stentz, "Optimal and efficient path planning for partially known environments," in *Intelligent unmanned ground vehicles*, pp. 203–220, Springer, 1997.
- [22] S. M. Lavalle and J. J. Kuffner Jr, "Rapidly-exploring random trees: Progress and prospects," in *International Workshop on Algorithmic and Computational Robotics*, 2000.
- [23] J. J. Kuffner and S. M. LaValle, "Rrt-connect: An efficient approach to single-query path planning," in *Proceedings of International Conference on Robotics and Automation (ICRA)*, 2000.
- [24] S. Karaman and E. Frazzoli, "Incremental sampling-based algorithms for optimal motion planning," *Robotics Science and Systems VI*, vol. 104, no. 2, 2010.
- [25] N. Ratliff, M. Zucker, J. A. Bagnell, and S. Srinivasa, "Chomp: gradient optimization techniques for efficient motion planning," in *Proceedings of International Conference on Robotics and Automation (ICRA)*, 2009.
- [26] J. Schulman, Y. Duan, J. Ho, A. Lee, I. Awwal, H. Bradlow, J. Pan, S. Patil, K. Goldberg, and P. Abbeel, "Motion planning with sequential convex optimization and convex collision checking," *International Journal of Robotics Research (IJRR)*, vol. 33, no. 9, pp. 1251–1270, 2014.
- [27] M. Stuede, K. Nuelle, S. Tappe, and T. Ortmaier, "Door opening and traversal with an industrial cartesian impedance controlled mobile robot," in *Proceedings of International Conference on Robotics and Automation (ICRA)*, 2019.
- [28] K. Gochev, A. Safonova, and M. Likhachev, "Planning with adaptive dimensionality for mobile manipulation," in *Proceedings of International Conference on Robotics and Automation (ICRA)*, 2012.
- [29] F. Burget, A. Hornung, and M. Bennewitz, "Whole-body motion planning for manipulation of articulated objects," in *Proceedings of International Conference on Robotics and Automation (ICRA)*, 2013.
- [30] M. Toussaint, K. Allen, K. A. Smith, and J. B. Tenenbaum, "Differentiable physics and stable modes for tool-use and manipulation planning," in *Proceedings of Robotics: Science and Systems (RSS)*, 2018.
- [31] R. Smits, H. Bruyninckx, and E. Aertbeliën, "Kdl: Kinematics and dynamics library," 2011. <https://www.orocos.org/kdl>.
- [32] C. Rösmann, F. Hoffmann, and T. Bertram, "Kinodynamic trajectory optimization and control for car-like robots," in *Proceedings of International Conference on Intelligent Robots and Systems (IROS)*, 2017.
- [33] L. Armstrong, "Optimization motion planning with tesseract and trajopt for industrial applications - ros-industrial."
- [34] B. Magyar, N. Tsiogkas, J. Deray, S. Pfeiffer, and D. Lane, "Timed-elastic bands for manipulation motion planning," *Robotics and Automation Letters (RA-L)*, vol. 4, no. 4, pp. 3513–3520, 2019.
- [35] Z. Jiao, Z. Zhang, W. Wang, D. Han, S.-C. Zhu, Y. Zhu, and H. Liu, "Efficient task planning for mobile manipulation: a virtual kinematic chain perspective," in *Proceedings of International Conference on Intelligent Robots and Systems (IROS)*, 2021.
- [36] P. R. Barragán, L. P. Kaelbling, and T. Lozano-Pérez, "Interactive bayesian identification of kinematic mechanisms," in *Proceedings of International Conference on Robotics and Automation (ICRA)*, 2014.
- [37] A. Paolillo, K. Chappellet, A. Bolotnikov, and A. Kheddar, "Inter-linked visual tracking and robotic manipulation of articulated objects," *Robotics and Automation Letters (RA-L)*, vol. 3, no. 4, pp. 2746–2753, 2018.
- [38] R. Martín-Martín and O. Brock, "Coupled recursive estimation for online interactive perception of articulated objects," *International Journal of Robotics Research (IJRR)*, pp. 1–37, 2019.
- [39] B. Abbatematteo, S. Tellex, and G. Konidaris, "Learning to generalize kinematic models to novel objects," in *Conference on Robot Learning (CoRL)*, 2019.
- [40] M. Han, Z. Zhang, Z. Jiao, X. Xie, Y. Zhu, S.-C. Zhu, and H. Liu, "Reconstructing interactive 3d scenes by panoptic mapping and cad model alignments," in *Proceedings of International Conference on Robotics and Automation (ICRA)*, 2021.
- [41] H. Liang, X. Ma, S. Li, M. Görner, S. Tang, B. Fang, F. Sun, and J. Zhang, "Pointnetgpd: Detecting grasp configurations from point sets," in *Proceedings of International Conference on Robotics and Automation (ICRA)*, 2019.
- [42] A. Mousavian, C. Eppner, and D. Fox, "6-dof grasping: Variational grasp generation for object manipulation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.
- [43] J. Ichnowski, Y. Avigal, V. Satish, and K. Goldberg, "Deep learning can accelerate grasp-optimized motion planning," *Science Robotics*, vol. 5, no. 48, 2020.