Tac-Man: Tactile-Informed Prior-Free Manipulation of Articulated Objects

Zihang Zhao[1,2 ⋆] zhaozihang@stu.pku.edu.cn
Yuyang Li[1,2,3 ⋆] liyuyang20@mails.tsinghua.edu.cn
Wanlin Li[2] liwanlin@bigai.ai
Zhenghao Qi[1,3] qi-zh21@mails.tsinghua.edu.cn
Lecheng Ruan[2 ✉] ruanlecheng@ucla.edu
Yixin Zhu[1 ✉] yixin.zhu@pku.edu.cn
Kaspar Althoefer[4] k.althoefer@qmul.ac.uk

⋆ equal contributors ✉ corresponding authors [1] Institute for Artificial Intelligence, Peking University
[2] Beijing Institute for General Artificial Intelligence [3] Department of Automation, Tsinghua University
[4] School of Engineering and Materials Science, Queen Mary University of London

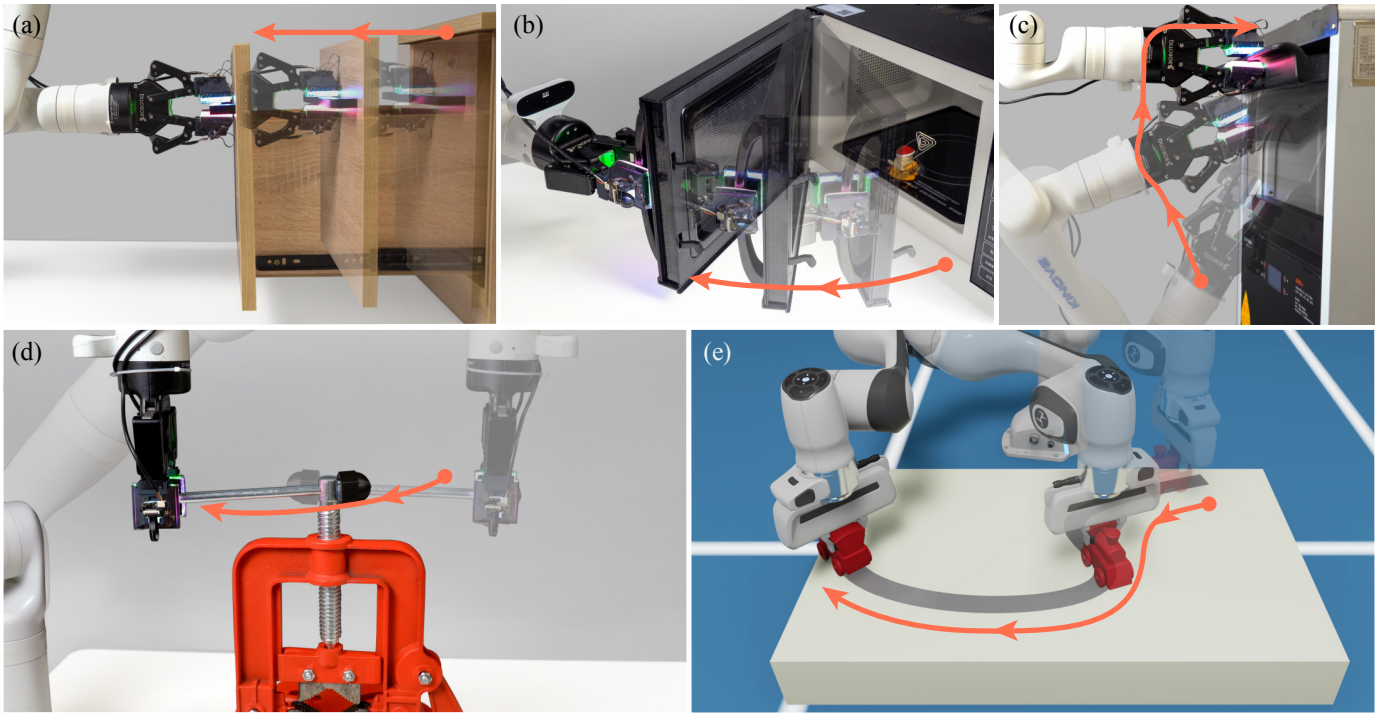

Fig. 1: **Tactile-informed prior-free manipulation of articulated objects.** Integrating GelSight-inspired tactile sensors into the robot's gripper enables precise manipulation of various articulated objects, eliminating the need for prior kinematic model knowledge. The capabilities are showcased by (a) manipulating objects with prismatic joints, such as drawers; (b) handling objects with revolute joints, such as microwave ovens; and managing complex mechanisms that involve simultaneous translational and rotational movements, as exemplified by (c) an ice machine and (d) a vise. Additionally, the robot proficiently performs (e) simulated arbitrary trajectories. Detailed trajectories and additional information are available in Supplementary Video S0. Complete URLs for all supplementary videos are listed in Tab. A1.

***Abstract*—Integrating robotics into human-centric environments such as homes, necessitates advanced manipulation skills as robotic devices will need to engage with articulated objects like doors and drawers. Key challenges in robotic manipulation are the unpredictability and diversity of these objects' internal structures, which render models based on priors, both explicit and implicit, inadequate. Their reliability is significantly diminished by pre-interaction ambiguities, imperfect structural parameters, encounters with unknown objects, and unforeseen disturbances. Here, we present a prior-free strategy, Tac-Man, focusing on maintaining stable robot-object contact during manipulation. Utilizing tactile feedback, but independent of object priors, Tac-Man enables robots to proficiently handle a variety of articulated objects, including those with complex joints, even when influenced by unexpected disturbances. Demonstrated in both real-world experiments and extensive simulations, it consistently achieves near-perfect success in dynamic and varied settings, outperforming existing methods. Our results indicate that tactile sensing alone suffices for managing diverse articulated objects, offering greater robustness and generalization than prior-based approaches. This underscores the importance of detailed contact modeling in complex manipulation tasks, especially with articulated objects. Advancements in tactile sensors significantly expand the scope of robotic applications in human-centric environments, particularly where accurate models are difficult to obtain.**

***Index Terms*—Articulated object manipulation, prior-free, tactile sensing, contact regulation.**

I. Introduction

ARTICULATED objects, defined by interconnected parts that move in conjunction with each other, include a vast array of everyday items, such as doors, drawers, and appliances, as illustrated in Fig. 1. Robots' capabilities in the manipulation of various articulated objects are essential for effective integration into human-centric environments, assisting humans in diverse tasks while representing a significant challenge within robotics [1, 2].

To interact efficiently with these objects, robots traditionally rely on prior knowledge of their kinematics. This knowledge is often embedded either explicitly, as in kinematic models used for designing manipulation strategies via control or planning techniques [3–20], or implicitly, within policy models developed through machine learning techniques [21–38]. However, acquiring precise prior knowledge about such objects presents several spatial and temporal challenges:

- **Ambiguity:** Objects with similar external appearances may have vastly different internal structures, complicating the unique determination of their kinematics through perception. This issue is extensively explored in Zhu *et al.* [39].
- **Imperfectness:** Accurate acquisition of an object's structural parameters is challenging due to perception noise or deviation from standard models, such as the off-axis doors.
- **Unknown:** Certain objects feature unique and sophisticated kinematics, making them difficult to model and generalize, as exemplified by the objects in Fig. 1(c) and (d).
- **Obsolescence:** Kinematic models can become outdated due to active or passive changes over time, leading to a gap in knowledge when executing manipulation tasks.

These challenges prompt us to question the hitherto assumed indispensability of prior knowledge in the manipulation of articulated objects and explore alternative approaches.

Drawing inspiration from human tactile interaction, we propose a novel prior-free approach, Tac-Man. Tac-Man is centered around stable robot-object contact during manipulation, achieved through high-resolution tactile sensing. By integrating GelSight-inspired sensors [40, 41] into the robot's gripper, Tac-Man adapts according to the tactile feedback during the interaction, allowing for dynamic adjustment during object handling. Without any prior-based modeling, Tac-Man addresses the challenges posed by ambiguity, imperfectness, unknowns, and obsolescence in object kinematics, which are prevalent in dynamic human-centric environments.

Our comprehensive validation includes a series of real-world experiments as well as extensive simulation studies. The experiments demonstrate Tac-Man's proficiency in handling objects with different joint types, from basic prismatic and revolute joints to complex mechanisms with simultaneous translational and rotational movements. We conducted further simulations to scale up these real-world experiments and probe Tac-Man's adaptability and robustness. Collectively, these studies confirm the effectiveness of Tac-Man, especially in scenarios where traditional reliance on priors is unreliable or inadequate.

The present paper makes three major contributions:

- We introduce Tac-Man, a novel tactile-informed prior-free manipulation framework for articulated objects. Unlike traditional methods that depend on prior knowledge of object kinematics, Tac-Man leverages tactile feedback to dynamically maintain stable contact during the manipulation process. This strategy ensures robust interaction with objects by adapting to real-time tactile data, offering a novel perspective on robotic manipulation.
- Through comprehensive real-world experiments, we validate the superior performance of Tac-Man over conventional methods in scenarios characterized by ambiguous, imperfect, unknown, and obsolescent priors. These experiments not only underscore Tac-Man's adaptability and precision in handling complex manipulations but also demonstrate its potential in enhancing robotic autonomy in unpredictable environments.
- We conduct an extensive simulation study to evaluate the scalability and generalization capabilities of Tac-Man across a diverse range of articulated objects and trajectories, including those arbitrarily generated. This study, one of the largest of its kind, confirms Tac-Man's effectiveness in a broad spectrum of settings, illustrating its versatility and the feasibility of its application in real-world scenarios beyond the confines of controlled experiments.

The paper is organized to provide a thorough exploration of Tac-Man and its implications. Sec. II reviews relevant literature, positioning our work within the broader context of robotic manipulation of articulated objects. Sec. III details Tac-Man, our novel tactile-informed, prior-free methodology. Sec. IV describes the setup and results of our real-world experiments, while Sec. V presents our simulation studies, showcasing the scalability and generalizability of Tac-Man. In Sec. VI, we present a detailed discussion on the strengths, limitations, and potential applications of Tac-Man, as well as considerations for future research. Finally, Sec. VII concludes the paper, summarizing our key findings and contributions to this field.

II. Related Work

A. Manipulation with Explicit Priors

Traditional robotic manipulation strategies have predominantly utilized manually predefined kinematic models, supplied externally, for interaction with articulated objects [3–5]. While this approach has been effective in certain controlled scenarios, it inherently limits the autonomy of the robotic system due to its reliance on pre-established, accurate kinematic information.

Seeking to enhance system autonomy, recent developments in robotic manipulation have shifted focus toward autonomously deriving object models via robot perception. This trend is exemplified by strategies employing visual inputs to predict kinematic models [9–11, 13–15, 20]. Nonetheless, this reliance on visual perception introduces ambiguity, as objects with similar appearances can harbor different internal articulations, challenging the accuracy of model predictions [13,

39]. To address these issues, several studies have integrated multi-frame analysis and wrist-mounted force/torque sensors for more accurate kinematic model estimation [6–8, 12, 16–19]. Despite these improvements, the dependency on being able to identify specific joint types, such as revolute or prismatic, limits the efficacy of such methods.

Our work introduces a novel prior-free approach to address these limitations inherent in manipulation methods with explicit priors. Tac-Man, validated by both real-world experiments and simulations, demonstrates superior adaptability and efficacy, especially in scenarios where traditional methods, reliant on explicit priors, are found to be unreliable or insufficient.

B. Manipulation with Implicit Priors

Robotic manipulation utilizing implicit priors typically relies on extensive data to infer appropriate actions from the observed state of an object. There are two prominent approaches: imitation learning and reinforcement learning.

Imitation learning enables robots to acquire implicit priors through the observation of human demonstrations [21–30], sometimes facilitated by tele-operational systems [42]. One major challenge in this approach is the collection of diverse and rich data from humans [43]. Although transfer methods [44–47] have improved the applicability to similar objects, extending these priors to a wider range of objects remains a substantial hurdle.

In contrast, reinforcement learning allows robots to develop implicit priors through iterative interactions in simulated settings [31–34]. The advent of extensive datasets of articulated objects [35–38] has significantly propelled this approach forward. However, the variance in these datasets often falls short in terms of replicating the complexity encountered in real-world situations, leading to incomplete prior acquisition, in turn constraining the robots' manipulation capabilities.

While both approaches show promise in articulated object manipulation, they grapple with the issue of incomplete priors due to limited data coverage. Tac-Man, sidestepping the need for laborious data collection and extensive training, addresses these challenges and has demonstrated its efficacy in manipulating a diverse array of articulated objects.

C. Robot-Object Contact Modeling

Current research tends to oversimplify and overlook the contact between robots and objects. The assumption of a fixed contact point is prevalent in numerous models [3, 4, 7, 12, 14, 16–19], yet this is often not feasible in practical applications. [5] have introduced the concept of a virtual joint [48] to represent this contact, but in actual implementation, they still rely on the assumption of a fixed joint. Works by [6, 8] have relaxed this constraint to some extent, permitting only specific types of movement between the robot and the object. [49] estimates contact using tactile sensors. However, it is limited to a specific type of handling based on rational motion and overlooks the potential for slippage.

Our research, however, uncovers that the data obtained from the contact site itself is adequate to support the task of articulated object manipulation, enabling a prior-free approach. By utilizing advanced tactile sensors to regulate contact, Tac-Man successfully demonstrates the ability to manipulate articulated objects under a diverse array of scenarios. This prior-free approach significantly enhances robots' manipulation skills in comparison to existing methodologies.

D. Tactile Sensors

Compared to 6D force/torque sensors commonly mounted at robots' wrists, tactile sensors play a crucial role in directly measuring the state of robot-object contact [50, 51]. Commercially available solutions like the BioTac sensor [52] are known for their high sensitivity, though they often lack spatial resolution. Alternative technologies, such as force-sensitive resistors, piezoresistive materials, and strain gauges, offer flexibility in mounting and are useful for single-point force sensing [53, 54]. However, they do not provide detailed spatial information about contact. E-skins, recognized for their adaptability and multi-modal sensing capabilities, deliver high resolution and are a promising development [55–60]. Yet, they are prone to manufacturing complexities, durability issues, and higher costs.

GelSight-type sensors, on the other hand, present a compelling alternative [61–69]. Characterized by pixel-level spatial resolution and cost-effective production, these sensors represent significant progress in tactile sensing technology. In our research, we have developed two GelSight-inspired tactile sensors, each uniquely mounted on the Robotiq 2F-85 gripper's opposing sides. These sensors are custom-designed with distributed markers [41], enabling efficient feature correspondence between frames. This design improves the gripper's ability to perceive fine details of contact, facilitating more nuanced manipulation of articulated objects.

III. The Tac-Man Method

This section outlines our problem formulation for prior-free manipulation of articulated objects. At the core of Tac-Man is the ability to regulate stable contact throughout the manipulation process. This involves guiding the object from its initial to its final state using just tactile signals. We begin by introducing essential notations and preliminaries in Sec. III-A, as these provide the foundational basis for our problem formulation and the subsequent derivations detailed in Sec. III-B. We then describe the contact representation in Sec. III-C. This representation is crucial in determining whether a contact is stable and this is discussed in Sec. III-D. Finally, Sec. III-E presents a computational method for calculating the robot's pose adjustments. These adjustments are essential for maintaining stable contact, thereby enabling the successful manipulation of articulated objects.

A. Notation and Preliminaries

Before delving into the detailed formulation of Tac-Man, we start by defining the following notations:

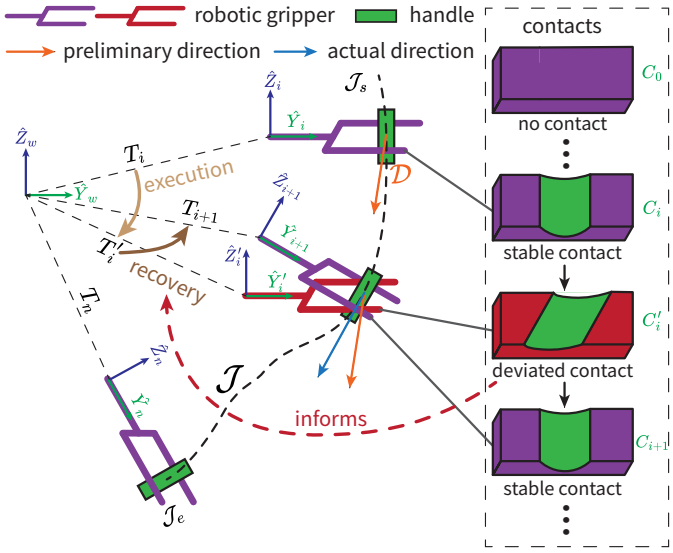

Fig. 2: **Schematic of Tac-Man.** Tac-Man starts from an initial pose and an estimated preliminary direction. Thereafter, the pose is dynamically adjusted so as to maintain stable contact throughout the manipulation process. This continuous adjustment is crucial to ensure the correct interaction direction is also maintained. This culminates in the successful manipulation of the articulated object.

- Bold small letters represent vectors (*e.g.*, $\mathbf{v}$), while capital letters denote matrices (*e.g.*, M). The subscript $[i]$, where $i = 1, \ldots, n$, on a bold small letter, refers to the i-th element of the vector. $I_{n\times n}$ is used to represent the identity matrix of dimension n.
- The symbol $\hat{\cdot}$ indicates the axis of a frame. The generalized pose of a frame $\{i\}$ relative to another frame $\{j\}$ is described by a position vector $\mathbf{p}_i^j \in \mathbb{R}^3$ and a rotation matrix $R_i^j \in SO(3)$. For simplicity, when frame $\{j\}$ is equivalent to the world frame $\{w\}$, the superscript is omitted. The homogeneous transformation matrix T_i^j, integrating the position vector and rotation matrix, is given by:

$$T_i^j = \begin{bmatrix} & R_i^j & & \mathbf{p}_i^j \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

- $\mathcal{K}_{ij}$ refers to a set of corresponding points from point clouds C_i and C_j, formulated as:

$$\mathcal{K}_{ij} = \{(\mathbf{u}, \mathbf{v}) \mid \mathbf{u} \in C_i, \mathbf{v} \in C_j\}.$$

The notation $\#\cdot$ is used to denote the number of elements in a set.

B. Problem Formulation

In this subsection, we formulate the process of tactile-informed prior-free manipulation of articulated objects. Consider an articulated object with a movable part, which is assumed to be graspable through a rigidly attached "handle." We postulate that the object's articulation restricts the part and the handle to a 1-Degree-of-freedom (DoF) trajectory. Let us denote the handle's trajectory as $\mathcal{J}$, beginning at point $\mathcal{J}_s$ and concluding at $\mathcal{J}_e$. Initially, in the absence of priors, only $\mathcal{J}_s$ is known. A robotic gripper, positioned in a global pose T_1 is grasping the handle. The objective is to navigate the gripper through a series of poses, $T_1, T_2, T_3, \ldots$, to maneuver the handle towards $\mathcal{J}_e$. At any given pose T_i, the contact state C_i between the object handle and the robotic gripper can be ascertained via a mapping function f_c:

$$C_i = f_c(\mathcal{J}, T_i). \tag{1}$$

In practical terms, C_i is directly measurable through tactile sensors. Importantly, the state prior to establishing any contact is denoted as C_0.

Stable contact: To ensure successful manipulation, it is crucial that the contacts $C_{1:n}$ meet certain constraints, guaranteeing that the gripper can consistently grasp the handle throughout the process. In this paper, we characterize a state that meets these constraints as maintaining *stable* contact.

Stable contact has two features. First, it must ensure that the gripper maintains an intact grip on the handle throughout the manipulation process and prevents slippage. This concept is quantified by referencing the initial contact state C_0 and assessing subsequent contact states C_i. We employ the function $f_d(C_i, C_0)$ to calculate the deformation of C_i; this deformation must stay within the elastic deformation limits D of the gripper material. Additionally, the shear deformation, as determined by the function $f_s(C_i, C_0)$, should not surpass S, a threshold dictated by the friction ratio between the handle and gripper material. The mathematical expression, therefore, is formulated as:

$$\begin{bmatrix} f_d(C_i, C_0) \\ f_s(C_i, C_0) \end{bmatrix} \leqslant \begin{bmatrix} D \\ S \end{bmatrix}. \tag{2}$$

Furthermore, a *stable* contact necessitates the gripper's engagement with the object's handle to ensure synchronized movement with the robot's actions. To establish a reference for *stable* contact C_1, an initial predefined pose T_1 is set. This practice is common in the manipulation of articulated objects [8]. A contact C_i is considered stable also requires its deviation from C_1 stays within a specified threshold E:

$$f_e(C_i, C_1) \leqslant E, \tag{3}$$

where $f_e(C_i, C_1)$ computes the difference between C_i and C_1.

Robot manipulation: Upon establishing stable contact, the manipulation process commences with a preliminary direction relative to frame $\{1\}$ for operating the articulated object. This direction, represented as $\mathcal{D}$, is encapsulated by a rotation matrix R^1 in $SO(3)$, detailed in Sec. VI-B. The homogeneous transformation matrix $T_{\mathcal{D}}$, resulting from moving along $\mathcal{D}$ by a distance y, is defined as:

$$T_{\mathcal{D}} = \begin{bmatrix} & I_{3\times3} & & R^1\mathbf{t} \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{4}$$

where $\mathbf{t}$, the translation vector, corresponds to the manipulation length along the negative y-axis as:

$$\mathbf{t} = \begin{bmatrix} 0 \\ -y \\ 0 \end{bmatrix}. \tag{5}$$

The robot follows the preliminary direction, executing a series of gripper frame poses $T_{1:n}$ to guide the handle's

transition from $\mathcal{J}_s$ to $\mathcal{J}_e$ while ensuring continuous stable contact. For each incremental step $i = 1, \ldots, n-1$, the next pose T_{i+1} is computed as follows:

$$T_{i+1} = T_i T_{i+1}^i = T_i \underbrace{T_{\mathcal{D}}}_{\text{execution}} \underbrace{\underset{T_{R_i} \in \mathbb{R}^{4\times4}}{\arg\min} \| f_c(T_i T_{R_i}, \mathcal{J}) - C_1 \|_2}_{\text{recovery}}. \quad (6)$$

Execution: The execution term in Eq. (6) involves proceeding to move along the preliminary direction as described in Eq. (4). This navigation progresses under the assumption of no prior knowledge about the object handle's trajectory, posing a risk of exceeding the tolerance limits necessary for maintaining stable contact. To balance this risk and maintain efficiency, it is crucial to carefully select the magnitude of y in Eq. (4). The objective is to maximize y while ensuring that the constraints for stable contact are met. This optimization can be mathematically represented as:

$$\begin{aligned} \underset{y}{\text{maximize}} \quad & y \\ \text{subject to} \quad & f_d(f_c(T_i T_{\mathcal{D}}, \mathcal{J}), C_0) \leqslant D, \\ & f_s(f_c(T_i T_{\mathcal{D}}, \mathcal{J}), C_0) \leqslant S, \\ & f_e(f_c(T_i T_{\mathcal{D}}, \mathcal{J}), C_1) \leqslant E. \end{aligned} \quad (7)$$

Recovery: The recovery term in Eq. (6) implies that at each step, the robot adjusts its pose to minimize any deviation from the stable contact state C_1, which may arise due to the execution step.

C. Contact Representation

For the effective computation of Eq. (6) and optimization of Eq. (7), a suitable representation of C_i is crucial. To ensure compatibility with the homogeneous transformation matrix and to match the capabilities of current tactile sensors, we represent the contact C_i as a set of positions of all discretized points at the contact site, relative to the reference frame $\{i\}$. The position of each contact point within C_i is denoted by $\mathbf{p}^i$. To facilitate alignment with the homogeneous transformation matrix, the column vector representation of $\mathbf{p}^i$ is defined in an augmented format as follows:

$$\mathbf{p}^i = \begin{bmatrix} p_x^i \\ p_y^i \\ p_z^i \\ 1 \end{bmatrix}, \quad (8)$$

where p_x^i, p_y^i, and p_z^i represent the coordinates relative to the gripper frame $\{i\}$.

Significantly, C_0 includes all points on the gripper's outer surface. Assuming the total number of points is m, C_0 is represented as:

$$C_0 = \{\mathbf{p}_i^0 \mid i = 1, \ldots, m\}. \quad (9)$$

For simplicity, we assume a flat gripper surface, a common assumption in current robotic grippers and easily achievable with tactile sensor fabrication. Consequently, p_x^0 is consistent across all points and set to 0, attainable through coordinate transformation. The position of a contact point at C_i is then characterized by the condition:

$$|\mathbf{p}_{[1]}^i| \geqslant \epsilon \quad (10)$$

where ϵ is a small positive number, indicating that deformation in the $\hat{x}$ direction must exceed ϵ for it to be considered contact.

Therefore, C_i for $i = 1 \ldots n$ can be defined as:

$$C_i = \{\mathbf{p}^i \mid |\mathbf{p}_{[1]}^i| \geqslant \epsilon\}. \quad (11)$$

D. Stable Contact

With the established representation of contact, we now turn to the computational process for determining stable contact.

The computation of stable contact involves addressing the gripper's maximum elastic deformation, encompassing both normal and shear components. These are captured in the expressions for $f_d(C_i, C_0)$ and D within the context of stable contact. Ensuring that the contact deformation remains within the material's elastic limits is crucial for maintaining a stable grip on the object:

$$f_d(C_i, C_0) = \begin{bmatrix} \max\{|\mathbf{u}_{[1]}| \mid \mathbf{u} \in C_i\} \\ \max\{\| \begin{bmatrix} \mathbf{u}_{[2]} - \mathbf{v}_{[2]} \\ \mathbf{u}_{[3]} - \mathbf{v}_{[3]} \end{bmatrix} \|_2 \mid (\mathbf{u}, \mathbf{v}) \in \mathcal{K}_{0i}\} \end{bmatrix} \quad (12)$$

and

$$D = \begin{bmatrix} d_n \\ d_s \end{bmatrix}, \quad (13)$$

where d_n and d_s denote the maximum normal and shear elastic deformation, respectively. For notation simplicity, we denote $|\mathbf{u}_{[1]}|$ as ΔN and $\| \begin{bmatrix} \mathbf{u}_{[2]} - \mathbf{v}_{[2]} \\ \mathbf{u}_{[3]} - \mathbf{v}_{[3]} \end{bmatrix} \|_2$ as ΔS henceforth.

Similarly, we formulate $f_s(C_i, C_0)$ and S to address concerns related to slipping in Eq. (2). Practically, to prevent sliding, the friction f between the object and the gripper should not surpass the maximum static friction f_M. In our model, the friction at C_i is positively correlated with the shear deformation observed from C_0:

$$f \sim \Delta S. \quad (14)$$

Thus, the shear deformation must adhere to the following constraints:

$$\forall (\mathbf{u}, \mathbf{v}) \in \mathcal{K}_{0i}, \Delta S \leqslant \delta^M(f_N), \quad (15)$$

where δ^M represents the deformation associated with the maximum static friction f_M, proportional to the normal force f_N. Assuming linear elasticity in the normal direction, δ^M can be modeled as:

$$\delta^M(\Delta S) = \delta_0 \Delta S, (\mathbf{u}, \mathbf{v}) \in \mathcal{K}_{0i}, \quad (16)$$

with δ_0 being a constant determined experimentally:

$$\delta_0 = \frac{\delta^M(\Delta S_0)}{\Delta S_0}, (\mathbf{u}, \mathbf{v}) \in \mathcal{K}_{0i}. \quad (17)$$

The no-sliding constraint is thus defined as:

$$\forall (\mathbf{u}, \mathbf{v}) \in \mathcal{K}_{0i}, \frac{\Delta S}{\Delta N} \leqslant \delta_0, \quad (18)$$

leading to the expression for $f_s(C_i, C_0)$:

$$f_s(C_i, C_0) = \max\{\frac{\Delta S}{\Delta N} \mid (\mathbf{u}, \mathbf{v}) \in \mathcal{K}_{0i}\}, \tag{19}$$

and setting

$$S = \delta_0. \tag{20}$$

Additionally, in the context of Eq. (3), we intuitively compute it as the difference between the corresponding points of two contacts, considering our established contact representation. To quantify this difference, we apply the loss function outlined in Besl *et al.* [70]. However, to ensure fairness in cases where a contact has a larger number of contact points, we normalize this loss by the number of corresponding pairs in $\mathcal{K}_{1i}$:

$$f_e(C_i, C_1) = \frac{\sum_{(\mathbf{u},\mathbf{v})\in\mathcal{K}_{1i}} \|\mathbf{u}-\mathbf{v}\|_2}{\#\mathcal{K}_{1i}}. \tag{21}$$

E. Robot Pose Update

Having established the contact representation and computation methods for stable contact constraints, we should address the optimization in Eq. (7) to compute the value of y for the determination of $T_{\mathcal{D}}$. However, in the absence of knowledge about the handle's trajectory $\mathcal{J}$, determining y explicitly remains challenging. We, therefore, adopt an exploration mechanism by continuously executing with small increments in y until any term in the constraints of optimization Eq. (7) reaches its upper bound, scaled by a safety margin α $(0 < \alpha < 1)$. This process is concluded as

$$T_i' = T_i T_{\mathcal{D}}, \tag{22}$$

which is referred to as the *execution stage*.

To recover the contact to within the specified bounds after the execution stage, we employ an optimization algorithm [70] to compute the optimal transformation T_{R_i} for the recovery term in Eq. (6):

$$T_{R_i} = \underset{T_{R_i}\in\mathbb{R}^{4\times4}}{\arg\min} \sum_{(\mathbf{u},\mathbf{v})\in\mathcal{K}_{1i}} \|T_{i+1}^{i}\mathbf{u}-\mathbf{v}\|_2. \tag{23}$$

The updated robot pose is then given by:

$$T_{i+1} = T_i' T_{R_i}. \tag{24}$$

This phase is referred to as the *recovery stage*. The update of the robot pose follows an iteration between the two stages.

IV. Real-world Experiments

To assess the effectiveness of Tac-Man, our proposed approach to prior-free manipulation of articulated objects, we conducted a series of real-world experiments. These were designed with meticulous attention to detail, as outlined in Sec. IV-A. They aim to thoroughly evaluate key constraints and determine essential hyper-parameters for the task, as discussed in Sec. III. Our experiments showcase the strengths of Tac-Man, particularly its robustness and adaptability under various conditions. These include scenarios where priors for object articulation are ambiguous (Sec. IV-B), imperfect (Sec. IV-C), unknown (Sec. IV-D), or obsolescent (Sec. IV-E), as previously discussed in Sec. I.

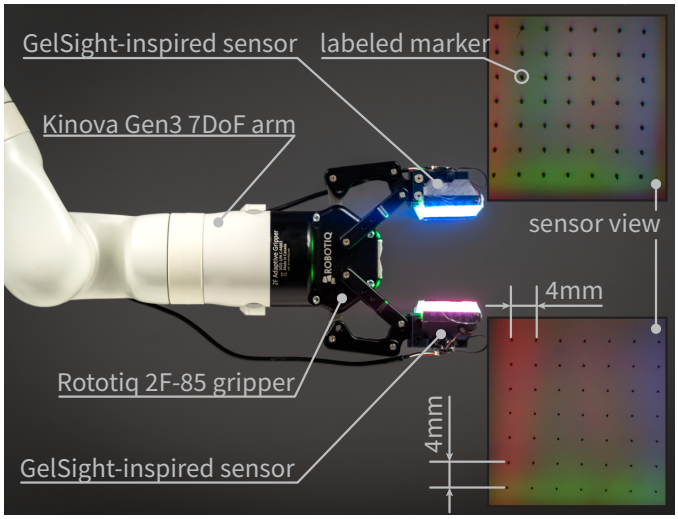

Fig. 3: **Real-world experiment setup.** The experiments are conducted in a real-world environment using a Gen3 7DoF robotic arm paired with a Robotiq 2F-85 gripper. Integral to the setup are two GelSight-inspired tactile sensors, each equipped with a 7×7 grid of labeled markers to enhance tactile sensing capabilities.

A. Experimental Setup

Robot system setup: Tac-Man is designed for broad compatibility with robotic systems capable of facilitating 6D movement in the gripper. In experiments, we utilized a Kinova Gen3 7DoF robotic arm equipped with a Robotiq 2F-85 gripper (Fig. 3). To assess the contact constraints as outlined in Sec. III, we integrated two GelSight-inspired tactile sensors into the gripper.

GelSight-inspired tactile sensor: Further enhancing our ability to validate these constraints, we based our sensors on the standard GelSight design, which provides high-resolution, pixel-level data at the contact surface. Each sensor was specially fabricated with a unique 7×7 grid of labeled markers (Fig. 4). This grid aids in accurately identifying point correspondences between two contacts. Although this modification slightly reduces the resolution and depth reconstruction accuracy as compared to conventional GelSight sensors, its effectiveness in practical scenarios and in validating the contact constraints detailed in Sec. III is notable.

System hyper-parameters: The determination of values for the defined parameters is crucial. The following outlines the chosen values, with a summary provided in Tab. I:

- The threshold for identifying contact points, ϵ in Eq. (10), is set to be adaptable. This is due to the material's flexibility, as shown in Fig. 5, which results in deformation not only in the contact region but also in the surrounding area upon contact with the object. This deformation relates to ΔN and the handle's shape, among other factors. For adaptability, ϵ is defined as the minimal ΔN that ensures at least 8 contact points, exceeding the necessary 6 for accurate 6D transformation solving in Eq. (23) and surpassing the number of markers in a line (7 in our sensors), thus ensuring a 3D spread of contact points.
- The normal elastic deformation limit, d_n in Eq. (13), is set

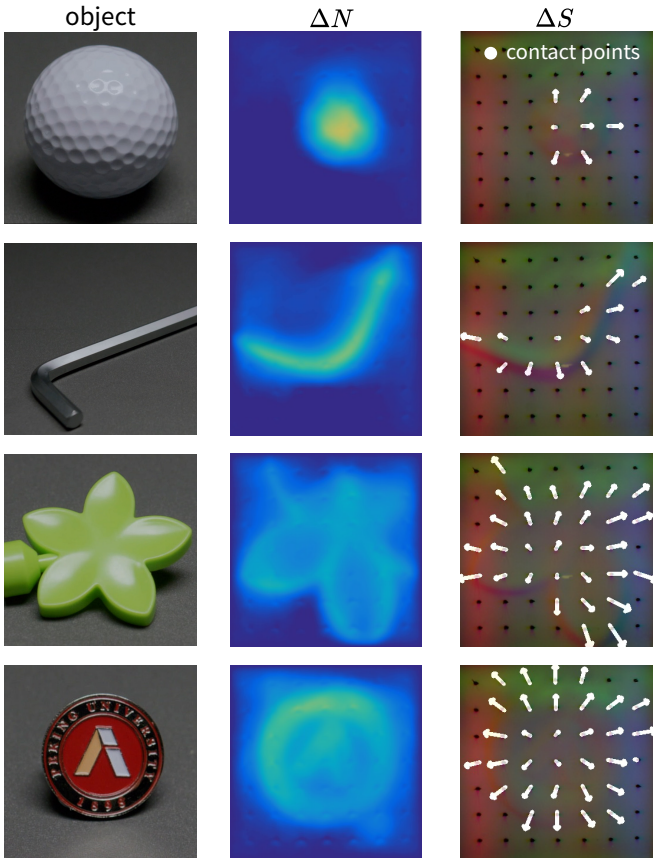

Fig. 4: **Visualization of tactile sensor reading.** The GelSight-inspired sensors are adept at providing detailed contact information during object interactions. These sensors excel in capturing ΔN in high spatial resolution, essential for precise detection of contact points. To further enhance their effectiveness, labeled markers are used for accurate point correspondence. Utilizing these features, ΔS is calculated and visualized using white arrows in the illustration. For better visibility, the magnitude of these arrows is exaggerated by a factor of 5, ensuring a clearer visualization of the contact dynamics.

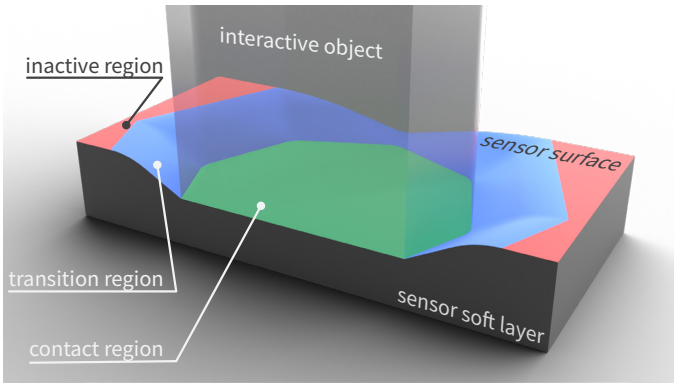

Fig. 5: **Selection concerns for the ϵ value.** The flexibility inherent in the tactile sensor material results in the activation of both the direct contact region and the adjacent transitional areas when contacting the handle. This characteristic underlines the need for a carefully adaptable ϵ value. Such adaptability ensures the precise identification of true contact points, thereby reducing the likelihood of detecting unreliable points.

at 2 mm. This value is chosen based on observations that

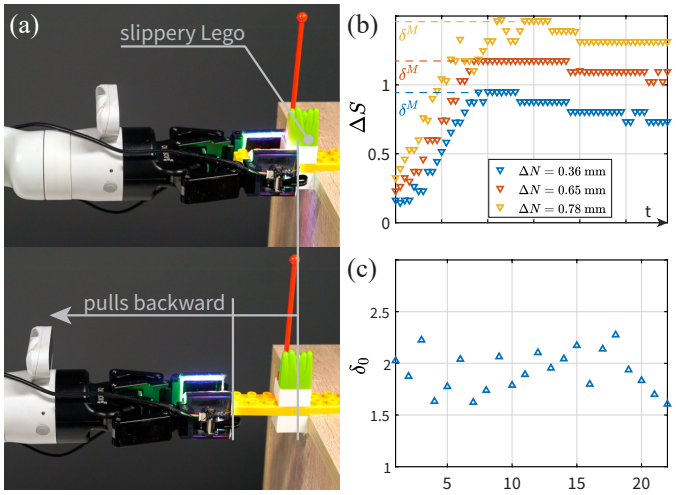

Fig. 6: δ_0 **value acquisition.** The experiment to obtain hyper-parameter δ_0, as defined in Eq. (17): (a) The experiment begins with the arm pulling backward after the gripper has firmly grasped a slippery Lego piece. (b) Multiple trials are conducted at different values of ΔN. Three representative trials are visualized to illustrate the process. The results follow the friction model well and δ^M can be easily identified. The results align well with the friction model, allowing for the straightforward identification of δ^M. (c) The δ_0 obtained from these multiple trials is visualized, and a value of 1.5 is selected for safety considerations.

within this range, the system can revert to its initial state after releasing contact. It is important to note that this value represents a conservatively reduced normal elastic deformation for enforcing stricter constraints.

- The parameter δ_0 in Eq. (17) is linked to handle material properties. To set a lower bound, an experiment with a slippery Lego piece, depicted in Fig. 6(a), identifies the δ^M values at different ΔN. This is done by tracking the peak shear displacement for each marker, as shown in Fig. 6(b). The results from multiple trials, illustrated in Fig. 6(c), lead to setting δ_0 at 1.5 for safety.
- The shear elastic deformation limit, d_s in Eq. (13), is calculated as $d_n\delta_0$. Though smaller than the actual one, this value is deemed acceptable for stricter constraint enforcement.
- The upper bound for another stable contact constraint, E in Eq. (3), is experimentally determined to be 0.4 mm, ensuring it does not exceed d_n, d_s, and δ_0 during manipulation.
- The movement magnitude in the *execution stage*, y in Eq. (7), is set as $\epsilon\delta_0$ to preserve stable contact during execution.

TABLE I: **Values of system hyper-parameters**

hyper-parameter	ϵ	d_n	d_s	δ_0	E	y	α
value	adaptive	2 mm	3 mm	1.5	0.4 mm	$\epsilon\delta_0$	0.6

The experiments described below are conducted in accordance with the setup and hyper-parameters detailed above. This ensures consistency and relevance in evaluating the efficacy of Tac-Man in real-world scenarios.

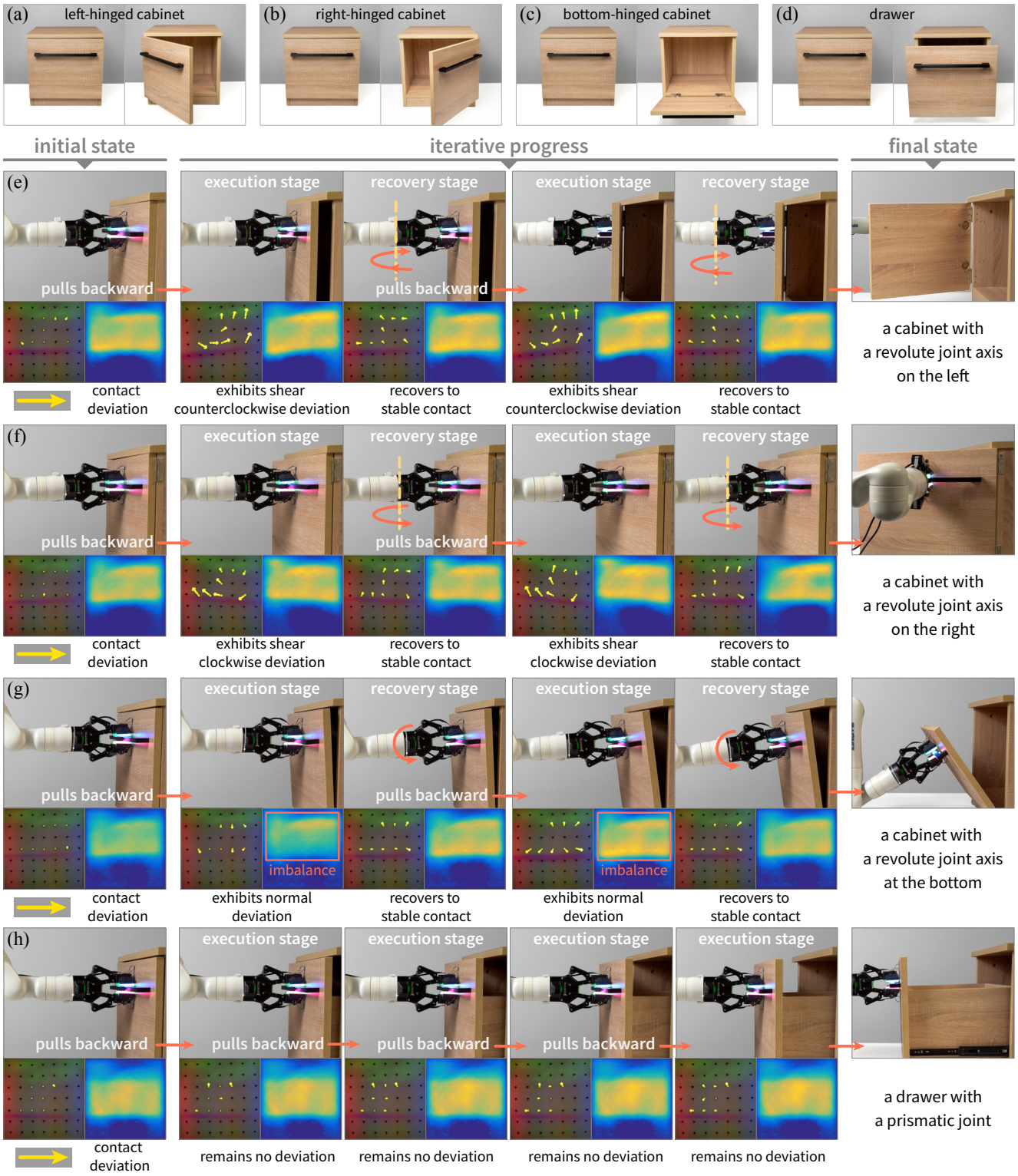

Fig. 7: **Manipulation under ambiguous priors.** Tac-Man showcases its proficiency in addressing the challenge caused by ambiguous priors, characterized by objects that, while visually indistinguishable, feature distinct articulation mechanisms. Our evaluation involves four articulated objects that look identical but differ in articulation: (a) a cabinet with a revolute joint on the left; (b) another cabinet with a revolute joint on the right; (c) a cabinet with a revolute joint at the bottom; and (d) a drawer with a prismatic joint. In scenarios (e)–(h), corresponding to objects (a)–(d), Tac-Man starts from the same position and in the same direction. The initial backward pull by the arm reveals variations in contact loss due to the differing articulations. Upon reaching a certain deviation threshold, the system triggers an adjustment in the arm's pose. This adjustment, optimized by Eq. (23), aims to recover stable contact and re-align to the correct interaction direction. The arm then continues its backward pull. Through this iterative process of adjustment and execution, each object, starting from the same initial state, is maneuvered to a distinct final state reflective of its unique articulation. For video demonstrations, refer to Supplementary Video S1.

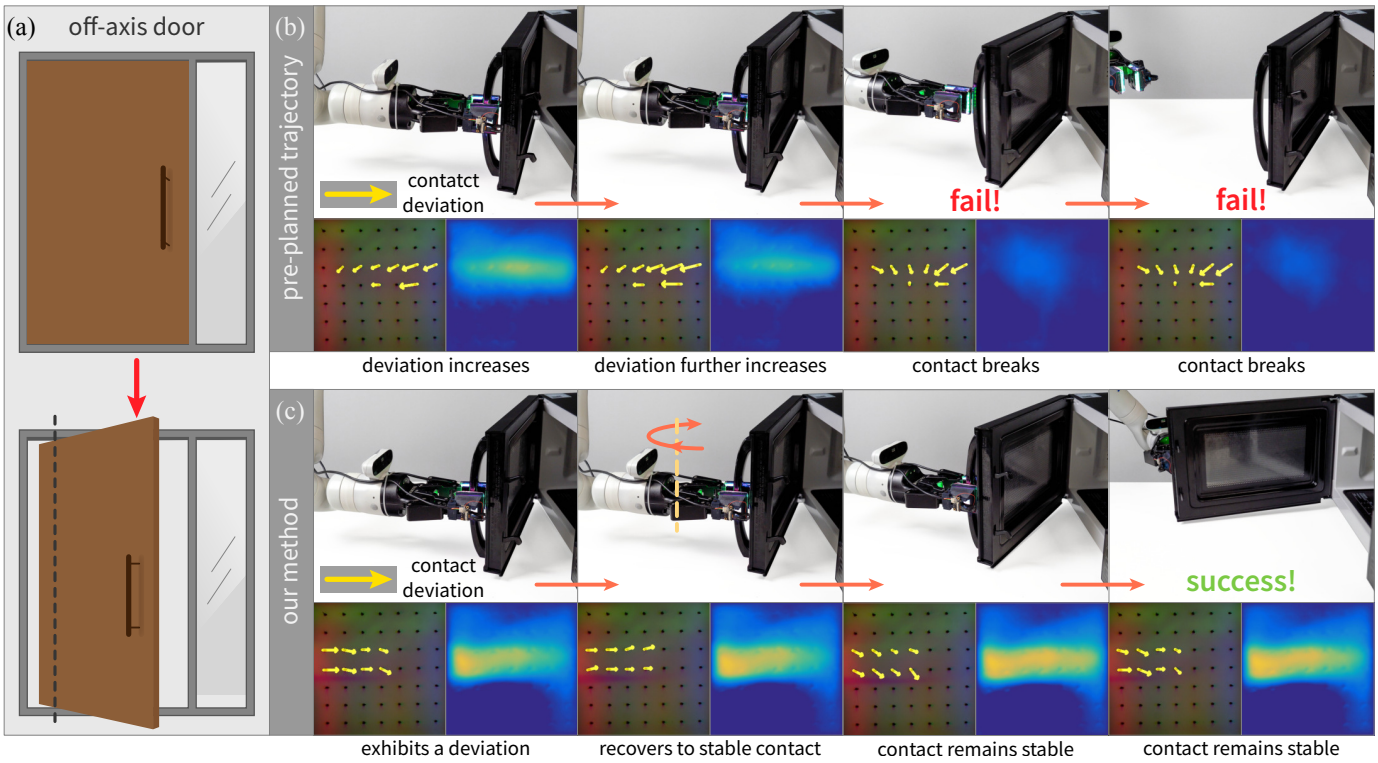

Fig. 8: **Manipulation under imperfect priors.** Tac-Man adeptly addresses challenges posed by imperfect priors, which accurately identify the model type but contain erroneous parameters. (a) An example is an off-axis door, designed for space and energy efficiency, diverging from traditional models by not placing the hinge directly on one side. This variation can mislead prior assumptions regarding the arc trajectory's radius. (b) We mirror this scenario by assigning the robot a pre-programmed arc trajectory to open a microwave oven with a revolute joint, incorporating a 10 % radius error. This slight misalignment between the priors and actual parameters leads to a gradual loss, eventually causing the gripper to lose contact with the handle, resulting in a failed manipulation attempt. (c) In contrast, Tac-Man dynamically re-calibrates the arm's pose to correct any unexpected errors, ensuring alignment with the correct interaction direction, going on to successfully open the microwave oven. This demonstration of adaptability underscores Tac-Man's ability to complete tasks even with imperfect priors. For video demonstrations, refer to Supplementary Video S2.

B. Manipulation under Ambiguous Priors

Ambiguous visual priors pose a significant challenge in robotic manipulation, especially when objects with visually indistinguishable appearances harbor varying articulation mechanisms. This dilemma is discussed in Sec. II-A and exemplified by Zhu *et al.* [39]. To illustrate, Fig. 7(a)–(d) features four objects that, despite their identical appearances, are equipped with different articulations: a cabinet with a revolute joint on the left (a), another with the joint on the right (b), a third with a joint at the bottom (c), and a drawer with a prismatic joint (d). This diversity underscores the complexity of inferring object kinematics solely from visual information.

Tac-Man effectively demonstrates proficient manipulation of distinct objects in all four cases, as shown in Fig. 7(e)–(h), each corresponding to the configurations in Fig. 7(a)–(d). Starting from the same initial state with stable contact C_1, Tac-Man employs an iterative two-stage cycle, beginning with an identical preliminary direction R^1. During the *execution stage*, tactile feedback is utilized to discern deviations from the correct interaction direction by observing changes in the contact points. For example, the contact in (e) exhibits a shear counterclockwise deviation, indicating a need for an additional clockwise rotation in the robot's movement, and conversely for the shear clockwise deviation observed in (f). In the scenario depicted in (g), a deviation in the normal direction ΔN signals the requirement for an added rotation for recovery. Conversely, in (h), the contact remains stable as the arm's movement is congruent with the correct direction. When a predetermined deviation threshold is reached, the system enters the *recovery stage*, employing the optimization outlined in Eq. (23) to compute the optimal arm pose that compensates for the observed deviation, thereby realigning the gripper to the correct interaction direction. Following this adjustment, the process resumes along the preliminary direction R^1. Through continual refinement to preserve stable contact, the arm successfully manipulates the objects into their intended states, as depicted in the last images in Fig. 7.

C. Manipulation under Imperfect Priors

Imperfect priors present a significant challenge in robotic manipulation, even when the articulation structure is known. Errors in model parameters can make these priors unreliable, as demonstrated by the example of an off-axis door, depicted in Fig. 8(a). Unlike conventional doors with hinges located along one edge, the off-axis door features a hinge positioned away from the edge, complicating the acquisition of accurate priors.

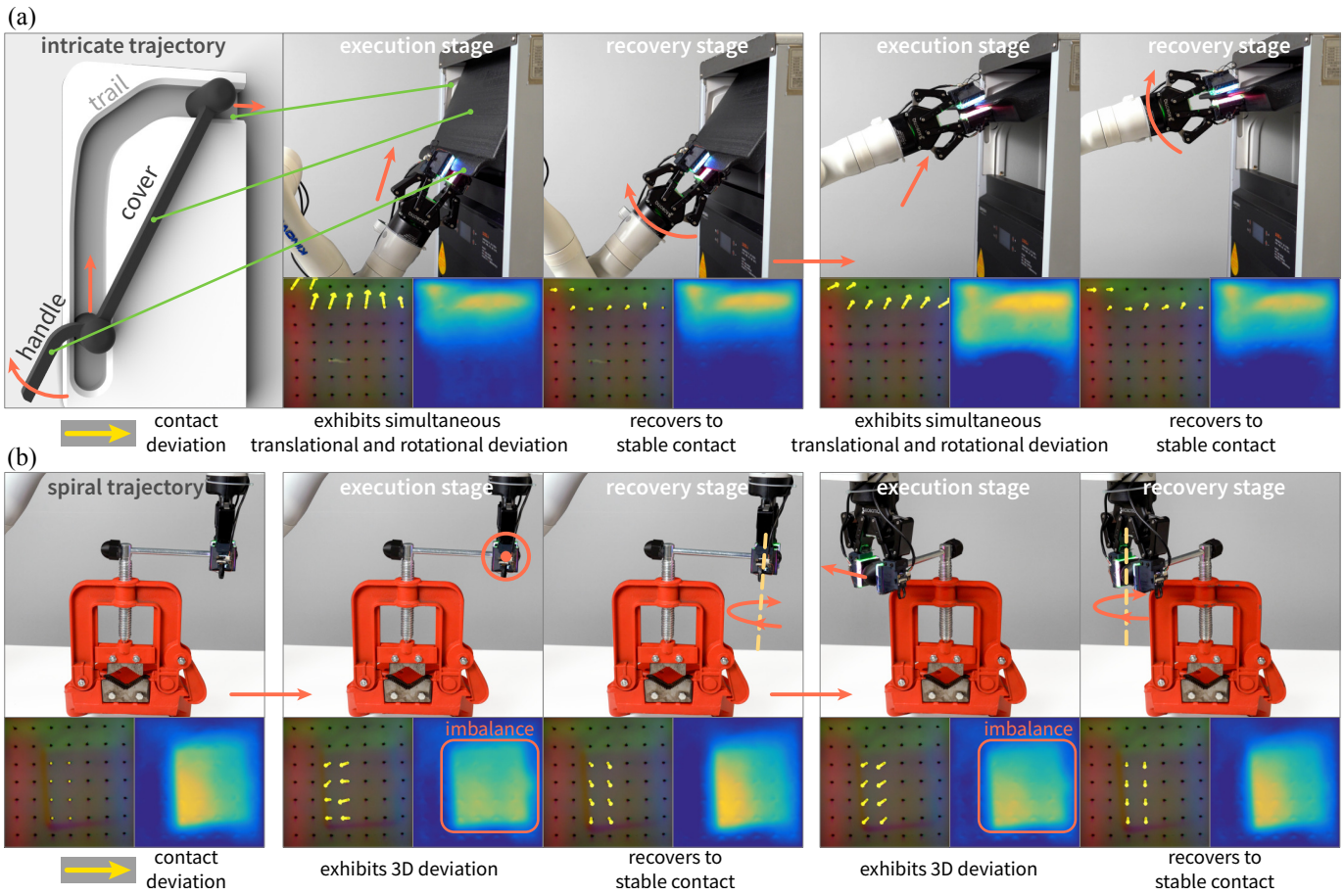

Fig. 9: **Manipulation under unknown priors.** Tac-Man showcases exceptional proficiency in managing scenarios characterized by unknown priors, particularly when object kinematics defy prior-based prediction or exceed the limits of current robotic perception. We illustrate this capability with two distinct examples. In scenario (a), the object's articulation mechanism, which is concealed beneath a cover, involves motions that combine translation and rotation—a complexity not easily captured by standard prismatic or revolute joint models. Scenario (b) presents a handle executing a helical trajectory, with the precise determination of its pitch posing a substantial challenge. Through these examples, Tac-Man proves its adeptness at navigating the intricacies of complex kinematic patterns, underscoring its effectiveness in scenarios where traditional perception models may falter. For video demonstrations, refer to Supplementary Video S3.

The impact of imperfect priors is illustrated through a challenging manipulation task, in which the robot is programmed to follow an arc trajectory to open a microwave oven with a revolute joint, as shown in Fig. 8(b). However, a 10 % error in the trajectory's radius introduces a small but critical imperfection in the priors. This error, without any real-time corrective adjustments, accumulates a loss that eventually leads to the gripper losing contact with the handle, resulting in task failure.

In contrast, Fig. 8(c) demonstrates Tac-Man's approach, employing adaptive adjustments during the manipulation to effectively counteract the inaccuracies in the priors. By dynamically compensating for the unforeseen deviation, Tac-Man ensures alignment with the correct interaction direction. This adaptability allows Tac-Man to successfully complete the task, fully opening the microwave oven door, showcasing its efficacy in overcoming the limitations imposed by imperfect priors.

D. Manipulation under Unknown Priors

Encountering objects whose articulation mechanisms defy accurate perception introduces a scenario where priors on kinematics remain unknown. Such conditions, characterized by complex movements that cannot be succinctly represented by conventional kinematic models like prismatic or revolute joints, present a significant challenge. Existing research has scarcely addressed the manipulation of objects requiring combined translational and rotational motions due to the difficulty in modeling these actions explicitly. For example, Fig. 9(a) illustrates an object designed compactly, necessitating that its handle executes both linear and rotational movement simultaneously. Another example is shown in Fig. 9(b), where a vise handle follows a spiral path. The ability to accurately identify and model such complex motion patterns remains beyond the reach of current perception methods.

Tac-Man circumvents this challenge, as demonstrated with two examples of objects whose handles exhibit complex movements combining translational and rotational motions. Through iterative processing, the Tac-Man adjusts the robot

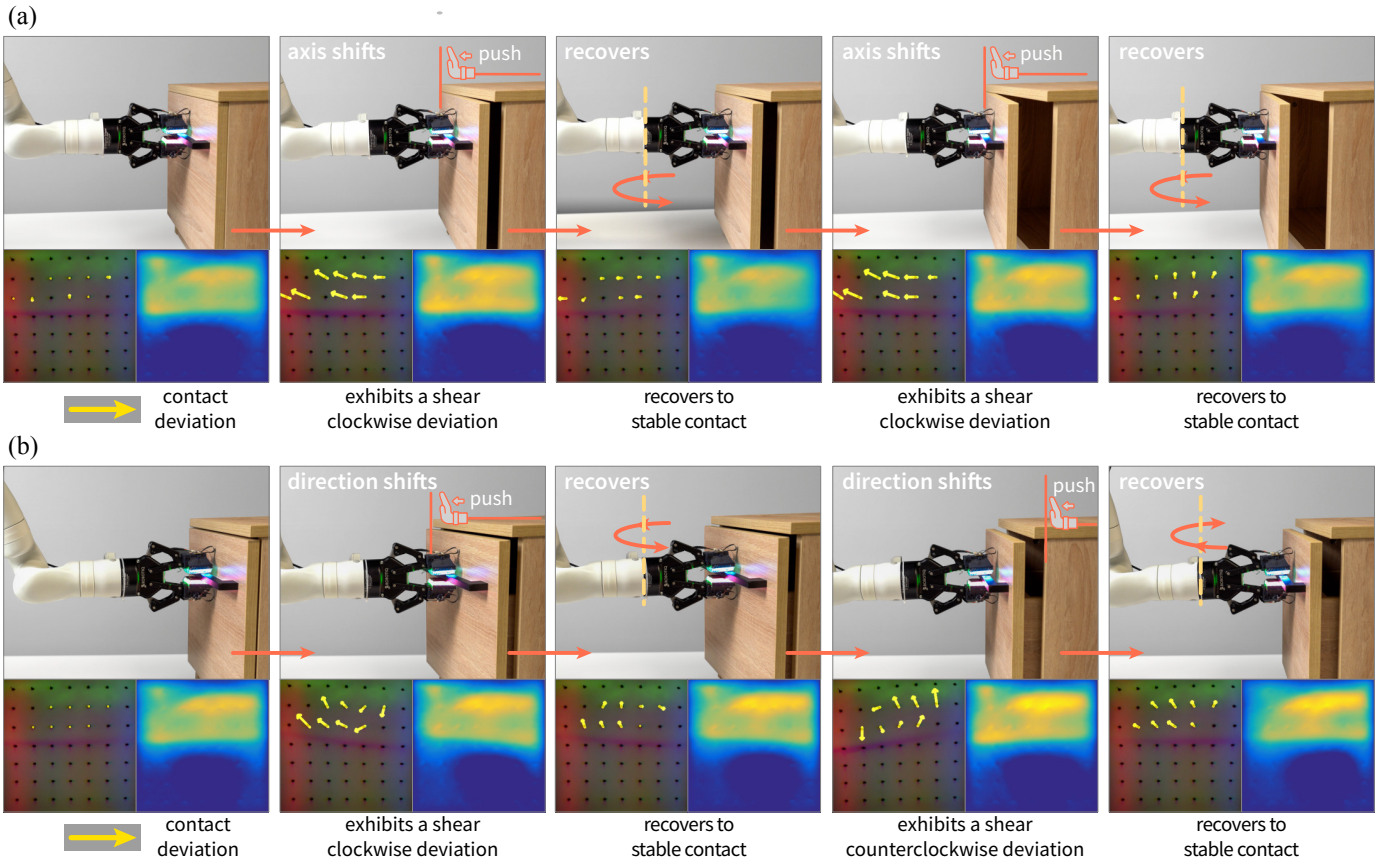

Fig. 10: **Manipulation under obsolescent priors.** Tac-Man demonstrates exceptional performance in scenarios characterized by obsolescent priors, often resulting from unpredictable perturbations that modify object kinematics. Such perturbations are common in dynamic, human-centric environments, frequently emerging from unanticipated human interactions. To assess Tac-Man's resilience against these perturbations, we introduced random forces to one side of the object in scenarios (a) and (b), effectively altering the cabinet's rotation axis in (a) and the drawer's opening direction in (b). Despite the introduction of errors through these human-induced disturbances, Tac-Man adeptly adjusts, preserving the correct interaction direction and ensuring successful manipulation outcomes. For video demonstrations, refer to Supplementary Video S4.

to the correct interaction direction and successfully manipulates the object in both examples.

E. Manipulation under Obsolescent Priors

Dynamic environments often introduce unpredictable perturbations, such as human interventions or environmental turbulence, rendering initially accurate priors for object kinematics obsolescent. To evaluate Tac-Man's performance in this context, we conducted experiments depicted in Fig. 10 in which the position of the articulated object was deliberately altered during manipulation. This was achieved by applying force to one side of the object, leading to a shift in either the rotation axis of a cabinet (Fig. 10(a)) or the opening direction of a drawer (Fig. 10(b)), as indicated by the hand icons.

Despite encountering such unpredictable disturbances, Tac-Man demonstrates remarkable adaptability. It swiftly compensates for the changes by correctly adjusting the robot's interaction strategy, thus enabling successful manipulation. This resilience underscores Tac-Man's ability to effectively handle real-time perturbations, ensuring consistent manipulation performance even when initial kinematic priors are no longer valid due to unforeseen environmental interactions.

V. Large-scale Verification in Simulation

Given the logistical challenges of physically experimenting with a comprehensive variety of objects, we look to simulation studies to further validate the generalization capability of Tac-Man. Utilizing NVIDIA Isaac Sim, our simulations cover a wide range of objects, mirroring the setup of our real-world experiments as outlined in Sec. V-A. For a thorough assessment, we incorporate objects with prismatic and revolute joints from the PartNet-Mobility dataset [36], reflecting the common articulations found in existing datasets.

Acknowledging the limitation of current datasets of objects with simpler joint types, we further extend our verification to include articulated objects designed with complex manipulation trajectories. These objects are uniquely characterized by trajectories that follow randomly sampled curves, introducing additional challenges to the manipulation task. In Sec. V-B, we elaborate on the preparation of these objects and present both qualitative and quantitative findings from our simulations. The results compellingly illustrate Tac-Man's adeptness

at navigating objects with a broad range of articulations, underscoring its effective adaptability and generalization across diverse scenarios.

A. Environmental Setup

Our simulation environment, depicted in Fig. 11, features a robot arm equipped with a gripper poised to engage an object's handle. Initially, the robot is programmed to follow a preliminary direction R^1, applying Tac-Man to articulate the object. A critical aspect of Tac-Man is its reliance on the compliance of silica gel material to accurately track contact between the gripper and the handle. To mimic this in a rigid body simulation environment, we employ specialized designs for contact measurement, detailed in Appx. B.

The simulation divides test objects into two categories: standard objects with prismatic or revolute joints from existing datasets, and randomly created custom objects with complex trajectories. From the PartNet-Mobility dataset [36], we select objects with detailed articulation annotations provided by GAPartNet [38], as outlined in Appx. C. For the evaluation, 25 objects with prismatic joints and another 25 with revolute joints were pre-processed. A manipulation is deemed successful if a revolute joint rotates beyond 60.0°or a prismatic joint extends over 250.0 mm. Fig. 11(a) showcases a subset of these objects and their initial grasps.

To assess Tac-Man's adaptability beyond basic joint types, we also evaluate it on articulated objects designed with randomly generated trajectories. These are constructed by sampling Bézier curve segments of various orders (2, 3, 4, and 5) on a 2D plane, ensuring no self-intersections. Each curve inspires a 3D playboard that defines a manipulation path, with a toy train at the start point acting as the handle. Physical collisions are configured to restrict the handle's movement along the trajectory, with the construction process detailed in Appx. D. Fig. 11(b) illustrates an example of a 3-order Bézier curve, its corresponding play-board, and the simulation's initial state. Additional examples of generated play-boards are presented in Fig. 11(c), highlighting the increased complexity and manipulation challenge with higher-order curves. For the evaluation, 25 curves of orders 3 to 6 are sampled, each tested once. Success is achieved when the toy train is moved from one end of the curve to the other by the gripper.

B. Results

Our simulation results substantiate the efficacy of the setup in evaluating Tac-Man. In Fig. 12, we highlight five representative examples: (a–c) illustrate the manipulation of objects with standard prismatic and revolute joints, while (d–e) detail scenarios involving intricate trajectory manipulations. The direction of pull and subsequent adjustments are indicated by the orange arrows. The simulated tactile feedback is visualized in the grey area beneath each figure, in which inactive, reference and current contact points are depicted as grey dots, blue circles, and red dots, respectively. Orange arrows, enlarged five-fold for clarity, depict the displacement from reference to current markers, with their size increasing in proportion to the error magnitude.

Consistent with our real-world experiments, deviations in the manipulation direction lead to deviation (Fig. 12(a-c)) in the tactile pattern, signaling the need for corrective movements beyond the preliminary direction. Once the contact deviation surpasses $\alpha\epsilon$ (Fig. 12(b–e)), Tac-Man adeptly adjusts to re-establish stable contact.

Quantitative evaluations for objects with prismatic and revolute joints are presented in Tab. II (columns 2-3), showcasing a 100 % success rate across all 50 objects (25 per joint type). Additionally, results for objects following intricate manipulation trajectories are tabulated in Tab. II (columns 4-7), with PB-n denoting performance on trajectories of order n. Remarkably, Tac-Man attains a 100 % success rate in all 200 trials (25 curves × 4 orders × 2 endpoints), underlining its proficiency and reliability in handling a broad spectrum of articulation patterns on a large scale without reliance on specific priors.

TABLE II: **Simulation quantitative results**

Category	**Pri. and Rev.**		**Intricate Trajectories**			
	Pri.	Rev.	PB-2	PB-3	PB-4	PB-5
Succ. (%)	100	100	100	100	100	100

VI. Discussion

A. Advancing Robotic Manipulation with Tac-Man

Our study has validated the efficacy of Tac-Man, a tactile-informed, prior-free approach for manipulating articulated objects. By leveraging tactile feedback, Tac-Man discerns discrepancies between the preliminary direction and actual interaction directions, allowing dynamic adjustments to ensure precise manipulations. A standout feature of Tac-Man is its operational independence from pre-existing knowledge about object kinematics, presenting a considerable advantage over traditional methods reliant on priors. This capability is particularly beneficial in environments where priors may be ambiguous (Sec. IV-B), imperfect (Sec. IV-C), unknown (Sec. IV-D), or obsolescent (Sec. IV-E)—conditions prevalent in dynamic, human-centric settings.

Tac-Man enhances robotic autonomy by obviating the need for manually inputting accurate object kinematic models, a common requirement in earlier strategies [3–5]. It sidesteps the limitations of methods that depend on visual perception for determining object kinematics [9–11, 13–15, 20], which are vulnerable to the ambiguities and inaccuracies intrinsic to visual data [39]. Unlike approaches that utilize multi-frame observations to clarify ambiguities [6–8, 12, 16–19], Tac-Man does not presuppose the existence of specific articulation types, thus demonstrating its versatility in managing novel articulation mechanisms as evidenced in Sec. IV-D.

Moreover, Tac-Man circumvents the exhaustive data collection process typically required by learning-based methods to tackle articulated object manipulation challenges [21–38]. The diversity and complexity of articulated objects, ranging from basic prismatic and revolute joints to more intricate

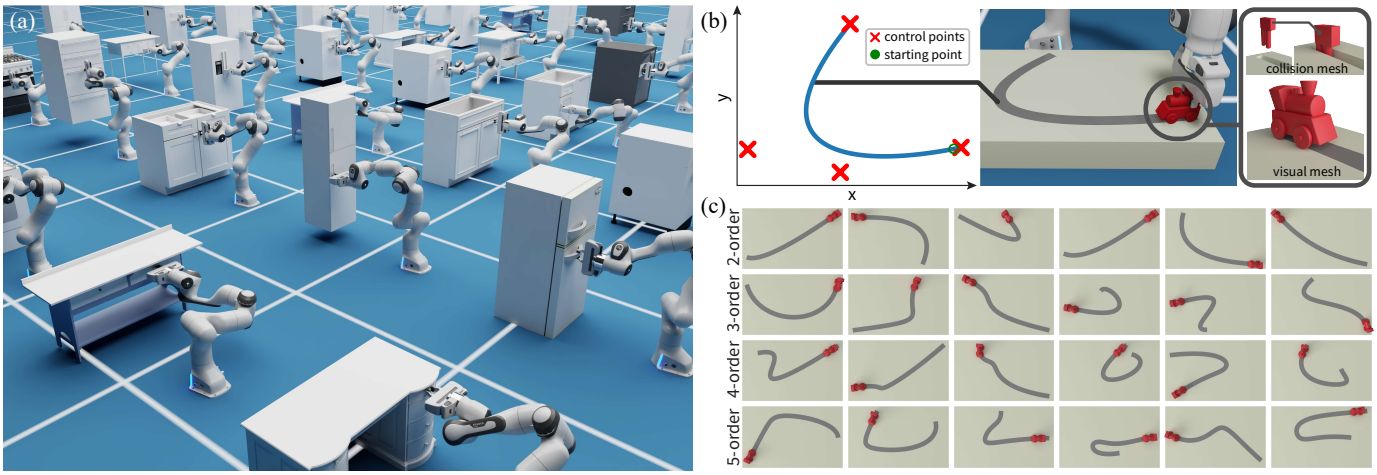

Fig. 11: **Simulation setup.** (a) The setup includes objects with prismatic and revolute joints from the PartNet-Mobility dataset [35], shown in their initial simulation states. (b) We further introduce a playboard featuring an intricate manipulation trajectory based on a randomly sampled Bézier curve; control points of the curve are marked with red "x"s. At the starting point of this curve, a toy train is placed to act as the manipulation handle (right), with a detailed collision setup (below). (c) Further examples of generated playboards, where rows 1-4 correspond to playboards with 2nd, 3rd, 4th, and 5th-order Bézier curves as the trajectories, showcasing the range of complexity in manipulation paths.

mechanisms featuring simultaneous translation and rotation, pose a significant challenge for data-driven approaches. Yet, Tac-Man adeptly addresses these challenges without relying on extensive datasets, thereby facilitating autonomous exploration and data collection for previously unseen objects. This autonomous data-gathering capability enhances the data collection process's versatility and substantially reduces human labor.

Additionally, Tac-Man showcases remarkable adaptability to dynamic changes within environments, as explored in Sec. IV-E. Its ability to adjust to unforeseen human interventions underscores a vital strength for robots operating in human-centric environments, where static assumptions about object kinematics can lead to manipulation failures. This adaptability is essential for ensuring seamless human-robot coexistence, enabling robots to effectively navigate the inherent unpredictability of real-world settings.

B. Limitations and Future Work

Preliminary direction acquisition: The efficacy of Tac-Man hinges on the availability of a preliminary direction for manipulating an articulated object. This foundational assumption is supported by advances in research, such as the work by Mo *et al.* [71], which aims to identify critical interaction points and predict preliminary directions for interaction, like pulling or pushing, for engaging with articulated objects. Although these directions might not always be accurate, Tac-Man is specifically engineered to accommodate and correct such preliminary direction, effectively handling any associated uncertainties. To showcase Tac-Man's resilience in this context, we experiment with a drawer, providing the robot arm with two intentionally erroneous directions. Despite these deviations, as depicted in Fig. 13, Tac-Man demonstrates its robust capability to adapt and successfully complete the drawer opening task.

This exploration into the reliance on preliminary direction underscores a limitation that also opens avenues for future research. Enhancing the accuracy of these directions or developing mechanisms within Tac-Man to autonomously generate and refine interaction directions could further improve its applicability and success rate across a wider array of articulated objects and scenarios. Investigating these enhancements would be a valuable direction for future work, potentially leading to even more adaptable and autonomous robotic manipulation systems.

Time efficiency: While Tac-Man demonstrates notable capabilities in the prior-free manipulation of articulated objects, an area for potential enhancement is its time efficiency. As indicated in Tab. III, the completion time for manipulation tasks using Tac-Man tends to exceed that of human performance. This disparity is largely attributed to the current implementation of Tac-Man, which operates without the benefits of closed-loop control.

In its present form, Tac-Man employs a cautious, step-by-step exploratory approach. This method, devoid of pre-established object priors or learning, requires careful evaluation of, and response to, tactile feedback. The task duration figures reported in Tab. III primarily reflect this initial phase of exploration, critical for the robot to accurately interpret and interact with the object.

Following this initial phase, the robot accumulates insights into the object's kinematics, aiding in the development of effective priors for future manipulations. These priors facilitate quicker subsequent manipulations, yet the absence of closed-loop control could lead to inefficiencies, particularly when initial conditions vary between attempts.

Looking ahead, enhancing Tac-Man to include closed-loop control elements could significantly improve its time efficiency. Such a development would enable quicker and more responsive adaptations during initial exploration, potentially reducing the overall time required for manipula-

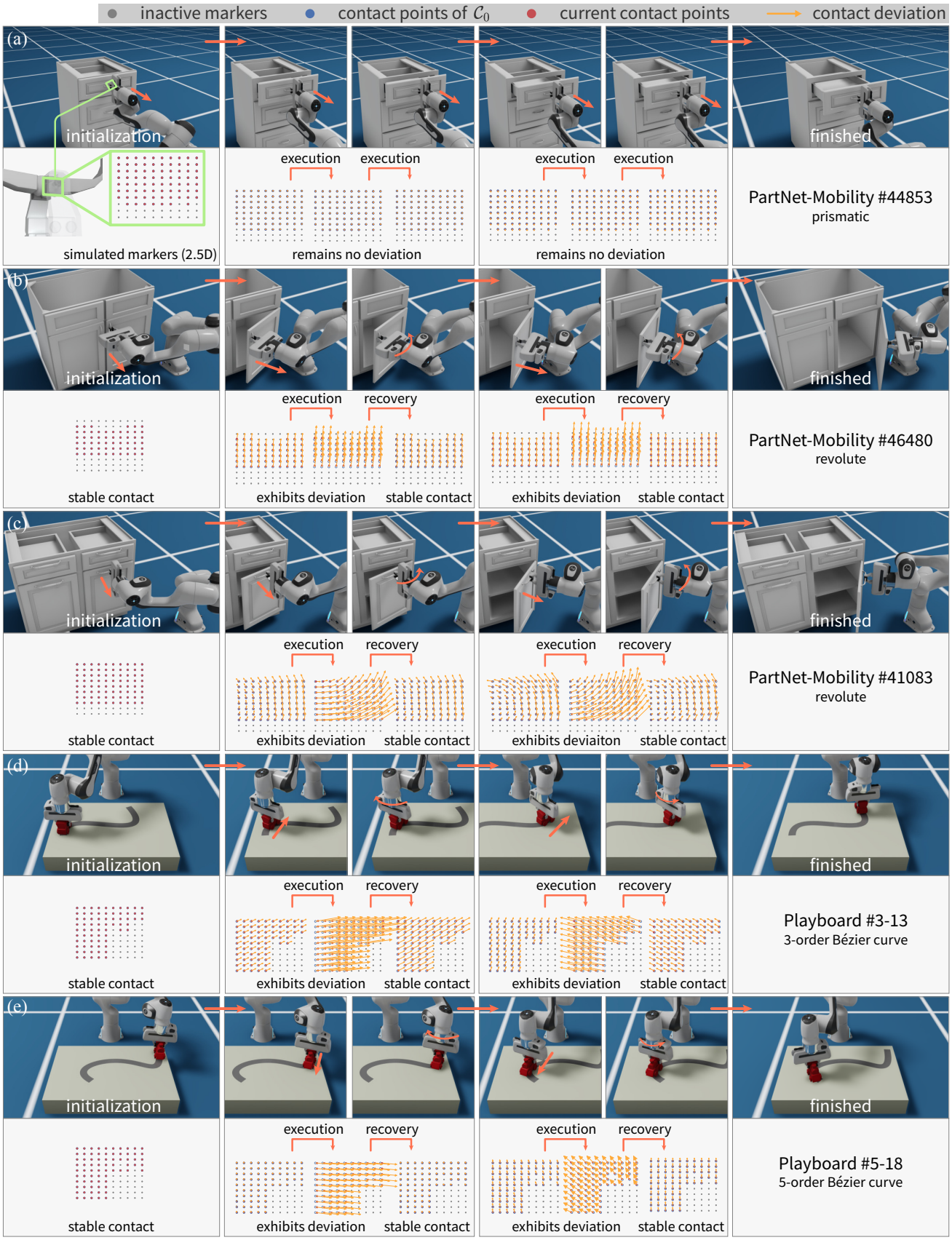

Fig. 12: **Examples of simulated manipulation.** Each row captures a sequence of manipulating an articulated object, showcasing scenarios with basic joint types (a-c) and intricate trajectories (d-e). For each scenario, the depiction includes the initial and final states, the *execution-recovery* iterations, and the tactile patterns observed. This visual representation underscores the effectiveness of our simulation setup in demonstrating Tac-Man's capability to adapt and succeed across varying articulation challenges. For video demonstrations, refer to Supplementary Video S5.

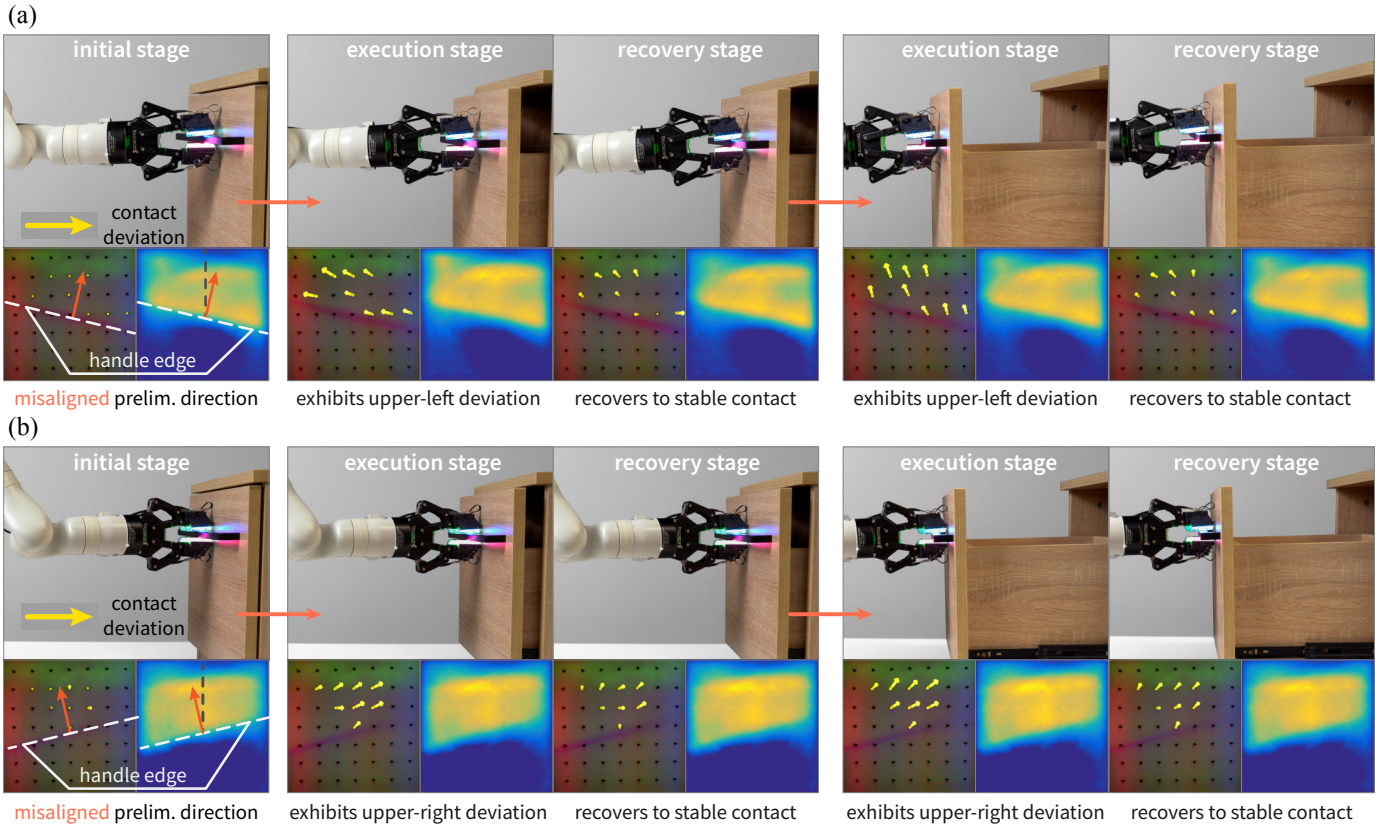

Fig. 13: **Manipulation under misaligned preliminary direction.** Even given a misaligned preliminary direction, Tac-Man is still able to use it to successfully open the drawer.

TABLE III: **Time for completing various manipulation tasks.**

Case	Fig. 7(e)	Fig. 7(f)	Fig. 7(g)	Fig. 7(h)	Fig. 8(b)	Fig. 9(a)	Fig. 9(b)
Time (s)	≈ 480	≈ 600	≈ 480	≈ 180	≈ 750	≈ 280	≈ 900

tions. It would lead to a more balanced approach, achieving manipulation robustness and efficiency on a par with human performance, especially in dynamic environments as discussed in Sec. IV-E.

VII. Conclusion

This paper introduces a novel prior-free approach to the manipulation of articulated objects, characterized by its reliance on tactile feedback rather than predefined object models. Central to this approach is the regulation of stable contact throughout the manipulation process, enabling dynamic adjustments of the robot's grip to ensure successful interaction with various objects.

The efficacy of Tac-Man is demonstrated through a series of experiments and simulations, highlighting its robustness in diverse scenarios. This includes its ability to handle ambiguous perceptions, adapt to imperfect parameter estimates, and respond effectively to unforeseen perturbations. The versatility of Tac-Man is further evidenced by its ability to manage objects with intricate trajectories, showcasing its practical utility in real-world applications.

In summary, this research makes a timely contribution to the field of robotic manipulation. It introduces a flexible and intuitive solution for interactions with articulated objects in environments marked by complexity and uncertainty. The adoption of a prior-free methodology marks a step forward in developing robotic systems that are more adaptable and capable of performing sophisticated manipulation tasks, paving the way for enhanced automation and efficiency in various applications.

Acknowledgment

Our sincere thanks go to Miss Yujie Fan and Miss Zhen Chen (BIGAI) for their expertise and creativity in figure design. We appreciate the technical support provided by Hongjie Li (PKU) in setting up the simulations. Special acknowledgment is due to Dr. Chi Zhang (BIGAI) and Muzhi Han (UCLA) for their participation in insightful discussions that enriched this work. We are grateful to Yida Niu (PKU) for developing the initial version of the Kinova driver, an essential component of our research setup. Lastly, we thank Mr. Mish Toszeghi (QMUL) for his meticulous proofreading of our manuscript.

Appendix

A. Material List

For detailed demonstrations related to this study, a comprehensive list of supplementary video URLs is provided. These videos visualize experiments and simulations discussed in this paper. Please refer to Tab. A1 for the complete set of links to these supplementary materials.

TABLE A1: **Complete URLs of supplementary materials.**

Material	Description and URL
Supp. Video S0	Manipulation trajectories in Fig. 1 https://vimeo.com/907222643/9793e630b0
Supp. Video S1	Manipulation under ambiguous priors https://vimeo.com/907224197/f38b1dea62
Supp. Video S2	Manipulation under imperfect priors https://vimeo.com/907224440/b059314753
Supp. Video S3	Manipulation under unknown priors https://vimeo.com/907224671/cb9aa51fcb
Supp. Video S4	Manipulation under obsolescent priors https://vimeo.com/907224990/3b4181a4e0
Supp. Video S5	Simulation studies https://vimeo.com/907225198/b641a5a84d

B. Details in Simulation Setup

Tac-Man leverages the compliance of the silica gel material on the gripper to track contact. Simulating such compliance within a rigid body simulator presents unique challenges, for which we have implemented a series of specialized designs. These designs are crucial to accurately compute the contact dynamics and effectively mimic the compliance of the gripper's material. The simulation setup aims to replicate the tactile feedback mechanism as closely as possible to real-world conditions. Key hyper-parameters play a significant role in this process, and their values are determined based on empirical observations and testing. The specific values used in our simulations, which are instrumental in achieving realistic contact dynamics, are enumerated in Tab. A2. This table provides a detailed overview of the hyper-parameters, ensuring transparency and reproducibility of the simulation results.

Establishing contact: In the simulated environment, a point cloud grid of resolution $n_{\text{res}} \times n_{\text{res}}$ is sampled on each fingertip of the gripper, resulting in $C_0^{(\text{sim})} = \{\mathbf{p}_i^0 | i = 1, \ldots, m\}$. The notation $((\cdot)^{(\text{sim})})$ indicates simulated quantities. To mimic real-world sensor noise, Gaussian noise $\epsilon \sim \mathcal{N}(0, \zeta^2)$ is added to the point cloud. Once the gripper grasps the handle, we define the initial contact set $C_1^{(\text{sim})}$ as:

$$C_1^{(\text{sim})} = \{\mathbf{p}_i | \mathcal{D}(\mathbf{p}_i^0) < \epsilon^{(\text{sim})}\}, \tag{A1}$$

where $\mathcal{D}(\mathbf{p})$ denotes the distance from point $\mathbf{p}$ to the gripper's surface.

Simulating contact: With $C_1^{(\text{sim})}$ established, this contact set is then affixed to the handle's frame. As the manipulation progresses, $C_1^{(\text{sim})}$ is transformed back to the gripper's frame T_i to update the contact $C_i^{(\text{sim})}$. The iterative process between the *execution* and *recovery* stages in the simulation adheres to the methodology detailed in Sec. III and utilizes the same hyper-parameters as in the real-world experiments (Tab. I).

Simulating gripper compliance: To realistically simulate gripper compliance within a rigid body simulation framework, we introduce a mechanism that fixes the joint on the object during the *recovery stage*. This approach is designed to approximate the flexibility and adaptive nature of an actual gripper's material, ensuring that the simulated interaction dynamics closely resemble those observed in real-world scenarios.

C. Preprocess PartNet-Mobility

The PartNet-Mobility dataset [36] comprises 1,045 articulated objects, and GAPartNet [38] enriches it with detailed part-level annotations. In this dataset, each object includes a base link and one or more moving links, connected via prismatic or revolute joints. Objects with multiple joints often necessitate sequential manipulations to fully access all their links (for instance, opening an oven door before pulling out the grill inside). For the sake of simplicity in our study, we focus on objects with a single accessible handle. We select objects that necessitate manipulating just one joint by grasping the linked handle. From this refined subset, we randomly chose 50 objects with revolute joints and another 50 with prismatic joints, adhering to these specified criteria.

D. Generate Random 1-DoF Playboards

Our simulation setup, as depicted in Fig. 12(b), incorporates playboards consisting of a baseboard (colored white) and a toy train (colored red). The train's movement is restricted to a predetermined trajectory (colored grey). To achieve this, we constructed a groove on a flat cube board and set up a distance joint between the board and the toy train, limiting its movement to within the plane of the board. Two cylindrical locating pins beneath the train's collision model are inserted into the groove, further guiding the train along the set trajectory.

The trajectories are generated as Bézier curves with n_{ctrl} control points, $B(s), s \in [0, 1]$, resulting in $(n_{\text{ctrl}} - 1)$-order curves. The control points are uniformly sampled and scaled to fit within a $H\text{cm} \times W\text{cm}$ area. This scaling ensures that the trajectories, confined within the convex hull of the control points, remain within the defined area. Any self-intersecting curves, taking into account a padding of w, are discarded. The groove, based on these curves, is carved into a $(H + 2w)\text{cm} \times (W + 2w)\text{cm}$ board. The initial position of the toy train is set at $B(\eta)$, where η is a small positive value to prevent initial pin-border collisions. Fig. 11(b) illustrates an example curve and corresponding play-board, with control points marked as red 'x's. Additional examples in Fig. 11(c) demonstrate that higher-order curves generally exhibit greater complexity, posing increased manipulation challenges. The hyper-parameters for this setup are listed in Tab. A2.

TABLE A2: **Hyper-parameters defined in Appxs. B and C.**

Hyper-parameter	n_{res}	$\epsilon^{(\text{sim})}$	ζ_n	n_{ctrl}	H	W	w	r_{pin}	η
Value	10	0.00025 m	2 mm	3, 4, 5, 6	40 cm	60 cm	4 cm	1.5 cm	0.02

References

[1] E. Krotkov, D. Hackett, L. Jackel, M. Perschbacher, J. Pippine, J. Strauss, G. Pratt, and C. Orlowski, "The darpa robotics challenge finals: Results and perspectives," *Journal of Field Robotics*, vol. 34, no. 2, pp. 229–240, 2017.

[2] A. Billard and D. Kragic, "Trends and challenges in robot manipulation," *Science*, vol. 364, no. 6446, p. eaat8414, 2019.

[3] S. Chitta, B. Cohen, and M. Likhachev, "Planning for autonomous door opening with a mobile manipulator," in *International Conference on Robotics and Automation (ICRA)*, 2010.

[4] F. Burget, A. Hornung, and M. Bennewitz, "Whole-body motion planning for manipulation of articulated objects," in *International Conference on Robotics and Automation (ICRA)*, 2013.

[5] Z. Jiao, Z. Zhang, X. Jiang, D. Han, S.-C. Zhu, Y. Zhu, and H. Liu, "Consolidating kinematic models to promote coordinated mobile manipulations," in *International Conference on Intelligent Robots and Systems (IROS)*, 2021.

[6] Y. Karayiannidis, C. Smith, F. E. Vina, P. Ogren, and D. Kragic, ""open sesame!" adaptive force/velocity control for opening unknown doors," in *International Conference on Intelligent Robots and Systems (IROS)*, 2012.

[7] K. Hausman, S. Niekum, S. Osentoski, and G. S. Sukhatme, "Active articulation model estimation through interactive perception," in *International Conference on Robotics and Automation (ICRA)*, 2015.

[8] 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.

[9] R. Hu, W. Li, O. Van Kaick, A. Shamir, H. Zhang, and H. Huang, "Learning to predict part mobility from a single static snapshot," *ACM Transactions on Graphics (TOG)*, vol. 36, no. 6, pp. 1–13, 2017.

[10] B. Abbatematteo, S. Tellex, and G. Konidaris, "Learning to generalize kinematic models to novel objects," in *Conference on Robot Learning (CoRL)*, 2019.

[11] X. Li, H. Wang, L. Yi, L. J. Guibas, A. L. Abbott, and S. Song, "Category-level articulated object pose estimation," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

[12] C. Moses, M. Noseworthy, L. P. Kaelbling, T. Lozano-Pérez, and N. Roy, "Visual prediction of priors for articulated object interaction," in *International Conference on Robotics and Automation (ICRA)*, 2020.

[13] V. Zeng, T. E. Lee, J. Liang, and O. Kroemer, "Visual identification of articulated object parts," in *International Conference on Intelligent Robots and Systems (IROS)*, 2021.

[14] M. Mittal, D. Hoeller, F. Farshidian, M. Hutter, and A. Garg, "Articulated object interaction in unknown scenes with whole-body mobile manipulation," in *International Conference on Intelligent Robots and Systems (IROS)*, 2022.

[15] M. Han, Z. Zhang, Z. Jiao, X. Xie, Y. Zhu, S.-C. Zhu, and H. Liu, "Scene reconstruction with functional objects for robot autonomy," *International Journal of Computer Vision (IJCV)*, vol. 130, no. 12, pp. 2940–2961, 2022.

[16] J. Lv, Q. Yu, L. Shao, W. Liu, W. Xu, and C. Lu, "Sagci-system: Towards sample-efficient, generalizable, compositional, and incremental robot learning," in *International Conference on Robotics and Automation (ICRA)*, 2022.

[17] Z. Jiang, C.-C. Hsu, and Y. Zhu, "Ditto: Building digital twins of articulated objects from interaction," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.

[18] B. Eisner, H. Zhang, and D. Held, "Flowbot3d: Learning 3D articulation flow to manipulate articulated objects," in *Robotics: Science and Systems (RSS)*, 2022.

[19] R. Martín-Martín and O. Brock, "Coupled recursive estimation for online interactive perception of articulated objects," *International Journal of Robotics Research (IJRR)*, vol. 41, no. 8, pp. 741–777, 2022.

[20] Z. Zhang, L. Zhang, Z. Wang, Z. Jiao, M. Han, Y. Zhu, S.-C. Zhu, and H. Liu, "Part-level scene reconstruction affords robot interaction," in *International Conference on Intelligent Robots and Systems (IROS)*, 2023.

[21] P. Pastor, H. Hoffmann, T. Asfour, and S. Schaal, "Learning and generalization of motor skills by learning from demonstration," in *International Conference on Robotics and Automation (ICRA)*, 2009.

[22] T. Welschehold, C. Dornhege, and W. Burgard, "Learning mobile manipulation actions from human demonstrations," in *International Conference on Intelligent Robots and Systems (IROS)*, 2017.

[23] D.-A. Huang, S. Nair, D. Xu, Y. Zhu, A. Garg, L. Fei-Fei, S. Savarese, and J. C. Niebles, "Neural task graphs: Generalizing to unseen tasks from a single video demonstration," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[24] T. Zhang, Z. McCarthy, O. Jow, D. Lee, X. Chen, K. Goldberg, and P. Abbeel, "Deep imitation learning for complex manipulation tasks from virtual reality teleoperation," in *International Conference on Robotics and Automation (ICRA)*, 2018.

[25] C. Lynch, M. Khansari, T. Xiao, V. Kumar, J. Tompson, S. Levine, and P. Sermanet, "Learning latent plans from play," in *Conference on Robot Learning (CoRL)*, 2020.

[26] H. Xiong, Q. Li, Y.-C. Chen, H. Bharadhwaj, S. Sinha, and A. Garg, "Learning by watching: Physical imitation of manipulation skills from human videos," in *International Conference on Intelligent Robots and Systems (IROS)*, 2021.

[27] Y. Qin, Y.-H. Wu, S. Liu, H. Jiang, R. Yang, Y. Fu, and X. Wang, "Dexmv: Imitation learning for dexterous manipulation from human videos," in *European Conference on Computer Vision (ECCV)*, 2022.

[28] J. Wong, A. Tung, A. Kurenkov, A. Mandlekar, L. Fei-Fei, S. Savarese, and R. Martín-Martín, "Error-aware imitation learning from teleoperation data for mobile manipulation," in *Conference on Robot Learning (CoRL)*, 2022.

[29] R. Gong, J. Huang, Y. Zhao, H. Geng, X. Gao, Q. Wu, W. Ai, Z. Zhou, D. Terzopoulos, and S.-C. Zhu, "Arnold: A benchmark for language-grounded task learning with continuous states in realistic 3d scenes," in *International Conference on Computer Vision (ICCV)*, 2023.

[30] J. Ye, J. Wang, B. Huang, Y. Qin, and X. Wang, "Learning continuous grasping function with a dexterous hand from human demonstrations," *IEEE Robotics and Automation Letters (RA-L)*, vol. 8, no. 5, pp. 2882–2889, 2023.

[31] Y. Urakami, A. Hodgkinson, C. Carlin, R. Leu, L. Rigazio, and P. Abbeel, "Doorgym: A scalable door opening environment and baseline agent," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2019.

[32] Z. Xu, Z. He, and S. Song, "Universal manipulation policy network for articulated objects," *IEEE Robotics and Automation Letters (RA-L)*, vol. 7, no. 2, pp. 2447–2454, 2022.

[33] Y. Chen, T. Wu, S. Wang, X. Feng, J. Jiang, Z. Lu, S. McAleer, H. Dong, S.-C. Zhu, and Y. Yang, "Towards human-level bimanual dexterous manipulation with reinforcement learning," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.

[34] H. Geng, Z. Li, Y. Geng, J. Chen, H. Dong, and H. Wang, "Partmanip: Learning cross-category generalizable part manipulation policy from point cloud observations," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.

[35] K. Mo, S. Zhu, A. X. Chang, L. Yi, S. Tripathi, L. J. Guibas, and H. Su, "Partnet: A large-scale benchmark for fine-grained and hierarchical part-level 3D object understanding," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[36] F. Xiang, Y. Qin, K. Mo, Y. Xia, H. Zhu, F. Liu, M. Liu, H. Jiang, Y. Yuan, H. Wang, *et al.*, "Sapien: A simulated part-based interactive environment," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

[37] L. Liu, W. Xu, H. Fu, S. Qian, Q. Yu, Y. Han, and C. Lu, "AKB-48: a real-world articulated object knowledge base," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.

[38] H. Geng, H. Xu, C. Zhao, C. Xu, L. Yi, S. Huang, and H. Wang, "Gapartnet: Cross-category domain-generalizable object perception and manipulation via generalizable and actionable parts," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.

[39] Y. Zhu, T. Gao, L. Fan, S. Huang, M. Edmonds, H. Liu, F. Gao, C. Zhang, S. Qi, Y. N. Wu, *et al.*, "Dark, beyond deep: A paradigm shift to cognitive ai with humanlike common sense," *Engineering*, vol. 6, no. 3, pp. 310–345, 2020.

[40] M. K. Johnson and E. H. Adelson, "Retrographic sensing for the measurement of surface texture and shape," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.
[41] W. Yuan, S. Dong, and E. H. Adelson, "Gelsight: High-resolution robot tactile sensors for estimating geometry and force," *Sensors*, vol. 17, no. 12, p. 2762, 2017.
[42] Y. Qin, W. Yang, B. Huang, K. Van Wyk, H. Su, X. Wang, Y.-W. Chao, and D. Fox, "Anyteleop: A general vision-based dexterous robot arm-hand teleoperation system," in *Robotics: Science and Systems (RSS)*, 2023.
[43] B. Zheng, S. Verma, J. Zhou, I. W. Tsang, and F. Chen, "Imitation learning: Progress, taxonomies and challenges," *IEEE Transactions on Neural Networks and Learning Systems*, no. 99, pp. 1–16, 2022.
[44] Y. Qin, H. Su, and X. Wang, "From one hand to multiple hands: Imitation learning for dexterous manipulation from single-camera teleoperation," *IEEE Robotics and Automation Letters (RA-L)*, vol. 7, no. 4, pp. 10873–10881, 2022.
[45] Z. Huang, J. Xu, S. Dai, K. Xu, H. Zhang, H. Huang, and R. Hu, "Nift: Neural interaction field and template for object manipulation," in *International Conference on Robotics and Automation (ICRA)*, 2023.
[46] A. Tekden, M. P. Deisenroth, and Y. Bekiroglu, "Grasp transfer based on self-aligning implicit representations of local surfaces," *IEEE Robotics and Automation Letters (RA-L)*, 2023.
[47] P. Li, T. Liu, Y. Li, Y. Geng, Y. Zhu, Y. Yang, and S. Huang, "Gendexgrasp: Generalizable dexterous grasping," in *International Conference on Robotics and Automation (ICRA)*, 2023.
[48] N. Likar, B. Nemec, and L. Žlajpah, "Virtual mechanism approach for dual-arm manipulation," *Robotica*, vol. 32, no. 6, p. E3, 2014.
[49] A. J. Schmid, N. Gorges, D. Goger, and H. Worn, "Opening a door with a humanoid robot using multi-sensory tactile feedback," in *International Conference on Robotics and Automation (ICRA)*, 2008.
[50] H. Yousef, M. Boukallel, and K. Althoefer, "Tactile sensing for dexterous in-hand manipulation in robotics—a review," *Sensors and Actuators A: physical*, vol. 167, no. 2, pp. 171–187, 2011.
[51] J. Lloyd and N. F. Lepora, "Pose and shear-based tactile servoing," *International Journal of Robotics Research (IJRR)*, vol. 0, no. 0, 2024.
[52] J. A. Fishel and G. E. Loeb, "Sensing tactile microvibrations with the biotac—comparison with human sensitivity," in *International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, 2012.
[53] H. Liu, X. Xie, M. Millar, M. Edmonds, F. Gao, Y. Zhu, V. J. Santos, B. Rothrock, and S.-C. Zhu, "A glove-based system for studying hand-object manipulation via joint pose and force sensing," in *International Conference on Intelligent Robots and Systems (IROS)*, 2017.
[54] Z. Lu and H. Yu, "Gtac-hand: A robotic hand with integrated tactile sensing and extrinsic contact sensing capabilities," *Transactions on Mechatronics (TMECH)*, 2023.
[55] J. Kim, M. Lee, H. J. Shim, R. Ghaffari, H. R. Cho, D. Son, Y. H. Jung, M. Soh, C. Choi, S. Jung, *et al.*, "Stretchable silicon nanoribbon electronics for skin prosthesis," *Nature Communications*, vol. 5, no. 1, pp. 1–11, 2014.
[56] Y. Wu, Y. Liu, Y. Zhou, Q. Man, C. Hu, W. Asghar, F. Li, Z. Yu, J. Shang, G. Liu, *et al.*, "A skin-inspired tactile sensor for smart prosthetics," *Science Robotics*, vol. 3, no. 22, p. eaat0429, 2018.
[57] C. M. Boutry, M. Negre, M. Jorda, O. Vardoulis, A. Chortos, O. Khatib, and Z. Bao, "A hierarchically patterned, bioinspired e-skin able to detect the direction of applied pressure for robotics," *Science Robotics*, vol. 3, no. 24, p. eaau6914, 2018.
[58] W. W. Lee, Y. J. Tan, H. Yao, S. Li, H. H. See, M. Hon, K. A. Ng, B. Xiong, J. S. Ho, and B. C. Tee, "A neuro-inspired artificial peripheral nervous system for scalable electronic skins," *Science Robotics*, vol. 4, no. 32, p. eaax2198, 2019.
[59] G. Li, S. Liu, L. Wang, and R. Zhu, "Skin-inspired quadruple tactile sensors integrated on a robot hand enable object recognition," *Science Robotics*, vol. 5, no. 49, p. eabc8134, 2020.
[60] Y. Yu, J. Li, S. A. Solomon, J. Min, J. Tu, W. Guo, C. Xu, Y. Song, and W. Gao, "All-printed soft human-machine interface for robotic physicochemical sensing," *Science Robotics*, vol. 7, no. 67, p. eabn0495, 2022.
[61] R. Li, R. Platt, W. Yuan, A. Ten Pas, N. Roscup, M. A. Srinivasan, and E. Adelson, "Localization and manipulation of small parts using gelsight tactile sensing," in *International Conference on Intelligent Robots and Systems (IROS)*, 2014.
[62] A. C. Abad and A. Ranasinghe, "Visuotactile sensors with emphasis on gelsight sensor: A review," *IEEE Sensors Journal*, vol. 20, no. 14, pp. 7628–7638, 2020.
[63] A. C. Abad and A. Ranasinghe, "Low-cost gelsight with uv markings: Feature extraction of objects using alexnet and optical flow without 3d image reconstruction," in *International Conference on Robotics and Automation (ICRA)*, 2020.
[64] M. Lambeta, P.-W. Chou, S. Tian, B. Yang, B. Maloon, V. R. Most, D. Stroud, R. Santos, A. Byagowi, G. Kammerer, *et al.*, "Digit: A novel design for a low-cost compact high-resolution tactile sensor with application to in-hand manipulation," *IEEE Robotics and Automation Letters (RA-L)*, vol. 5, no. 3, pp. 3838–3845, 2020.
[65] I. H. Taylor, S. Dong, and A. Rodriguez, "Gelslim 3.0: High-resolution measurement of shape, force and slip in a compact tactile-sensing finger," in *International Conference on Robotics and Automation (ICRA)*, pp. 10781–10787, 2022.
[66] W. Li, M. Wang, J. Li, Y. Su, D. K. Jha, X. Qian, K. Althoefer, and H. Liu, "L^3 f-touch: A wireless gelsight with decoupled tactile and three-axis force sensing," *IEEE Robotics and Automation Letters (RA-L)*, 2023.
[67] J. Zhao and E. H. Adelson, "Gelsight svelte: A human finger-shaped single-camera tactile robot finger with large sensing coverage and proprioceptive sensing," in *International Conference on Intelligent Robots and Systems (IROS)*, 2023.
[68] W. Li, A. Alomainy, I. Vitanov, Y. Noh, P. Qi, and K. Althoefer, "F-touch sensor: Concurrent geometry perception and multi-axis force measurement," *IEEE Sensors Journal*, vol. 21, no. 4, pp. 4300–4309, 2020.
[69] S. Wang, Y. She, B. Romero, and E. Adelson, "Gelsight wedge: Measuring high-resolution 3D contact geometry with a compact robot finger," in *International Conference on Robotics and Automation (ICRA)*, 2021.
[70] P. J. Besl and N. D. McKay, "Method for registration of 3-d shapes," in *Sensor fusion IV: control paradigms and data structures*, vol. 1611, pp. 586–606, Spie, 1992.
[71] K. Mo, L. J. Guibas, M. Mukadam, A. Gupta, and S. Tulsiani, "Where2act: From pixels to actions for articulated 3D objects," in *International Conference on Computer Vision (ICCV)*, 2021.