

Learning From Demonstration Based on Environmental Constraints

Xing Li , *Graduate Student Member, IEEE*, and Oliver Brock , *Fellow, IEEE*

Abstract—We present a novel learning from demonstration approach which uses environmental constraints as the underlying representation to interpret and reproduce demonstrations. This representation based on environmental constraints separates the information that facilitates generalization from the information specific to object instances. Combined with adaptive controllers which fill in the instance-specific details during execution through explorative interaction, our approach generalizes from a single demonstration on an articulated object to different instances of the same object type. We test our approach in real-world experiments on contact-rich manipulation, using a series of mechanical locks as well as drawers and doors. The high success rate of 95% across all of these experiments provides strong evidence that environmental constraints are a powerful inductive bias for general and robust learning from demonstration.

Index Terms—Learning from demonstration, perception-action coupling.

I. INTRODUCTION

WE PRESENT an approach to kinesthetic learning from demonstration (LfD) in contact-rich manipulation tasks. The approach generalizes from a *single* demonstration of a task using an articulated object to other objects with similar but not identical kinematic structures, different sizes, and different object placements. To achieve such generalization with only a single demonstration, we must extract from the demonstration the information that facilitates generalization, ignoring information specific to the object instance. The key insight for achieving this stems from human manipulation. Humans extensively leverage contact with the environment during manipulation. We refer to this as the exploitation of environmental constraints [1]. We show that environmental constraints (ECs) are the appropriate representation for LfD, enabling generalization and robustness.

The function of many objects is encoded in their kinematic structure: doors, drawers, scissors etc. This structure constrains

Manuscript received 24 February 2022; accepted 21 June 2022. Date of publication 3 August 2022; date of current version 23 August 2022. This letter was recommended for publication by Associate Editor L. Peternel and Editor D. Kulic upon evaluation of the reviewers' comments. This work was supported in part by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through Germany's Excellence Strategy – EXC 2002/1 Science of Intelligence– Project under Grant 390523135. (*Corresponding author: Oliver Brock.*)

The authors are with the Robotics and Biology Laboratory, Technische Universität Berlin, 10587 Berlin, Germany, and also with the Science of Intelligence (SCIoI) and Cluster of Excellence, Technische Universität Berlin, 10587 Berlin, Germany (e-mail: xing.li@tu-berlin.de; oliver.brock@tu-berlin.de).

This letter has supplementary downloadable material available at <https://doi.org/10.1109/LRA.2022.3196096>, provided by the authors.

Digital Object Identifier 10.1109/LRA.2022.3196096

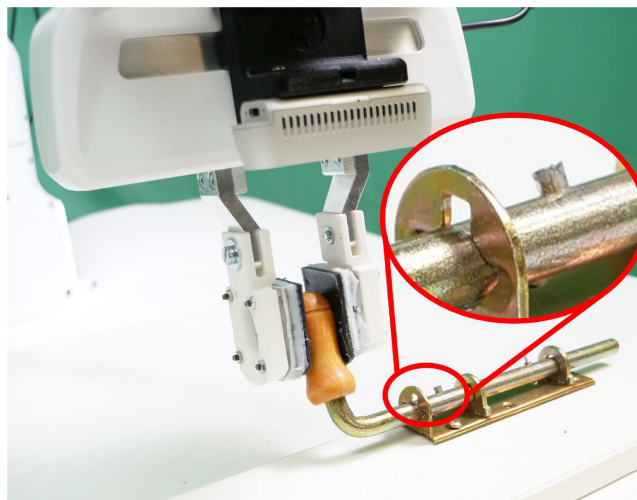


Fig. 1. Based on a single human demonstration, this robot opens a lock robustly and repeatedly. This requires passing a small pin through the narrow slot. The resulting policy robustly transfers to other locks that vary significantly from the one used in the demonstration. The policy even generalizes to placement variations of these different locks.

the relative motion of the object's parts. To the robot interacting with these objects, the kinematic model appears as an environmental constraint [1] i.e. a feature of the environment that dictates aspects of the motion the robot can perform while in contact with that object. It is an important observation that different objects of the same type can differ in the parameters of their kinematic model but the type of ECs that result from the model remain the same. This enables humans to transfer experience between objects of the same type. By using ECs as the underlying representation for the information extracted from human demonstrations, we can impart similar transfer abilities to robots.

Environmental constraints also play an important role in the robust execution of policies learned from demonstrations. When manipulating articulated objects with tight clearances, small errors during motion execution can cause the robot to be blocked. The associated contact forces lead to perceptual aliasing of contact states relevant to the policy. To address these problems, our approach leverages the ECs extracted from the demonstration as an inductive bias for motion execution.

We leverage these insights to devise an approach to LfD based on ECs. We demonstrate the learning and transfer of a policy from a single demonstration with a lock, such as the one shown in Fig. 1, to other locks of the same type but greatly differing

physical properties, achieving success in 49 out of 50 trials. We show that success remains high, even when we rotate the lock relative to the lock's orientation in the demonstration. We also demonstrate the approach on a cabinet with three compartments, each having a different opening mechanism. Again, a single demonstration suffices to open all three. Our key contribution is to present this LfD system based on ECs and demonstrate that ECs are a suitable representation to achieve transfer and robustness in LfD for contact-rich manipulation tasks.

II. RELATED WORK

Learning from demonstration is a promising way to teach a robot various manipulation skills efficiently [2]. One common paradigm for LfD is behavior cloning, which learns a policy to map raw observations into low-level actions from demonstrations. This paradigm usually uses all available information in demonstrations. We argue that separating general information from unnecessary details is important for LfD. In this work, we achieve this separation by using ECs as inductive biases.

Another important paradigm of LfD is to learn movement primitives from human demonstrations [3]. These methods assume the motion characteristics in demonstrated movements encapsulate the essential information to reproduce the demonstrated behaviors [4], [5]. However, Kalakrishnan et al. [6] found that simply playback a demonstrated trajectory on a robot is not able to accomplish contact-rich manipulation tasks. One reason is that it is difficult to track trajectories in a chaotic environment. In addition, demonstrations do not contain the information about friction, which is crucial to be compensated to reproduce the demonstrated in-contact skills [7]. We thus hypothesize that demonstrated motion characteristics are not the right representation for contact-rich manipulation tasks. Therefore, our work focuses on deriving a policy based on ECs from the demonstration.

Environmental Constraint Exploitation (ECE) is introduced by Eppener et al. [1] to describe behaviors which leverage useful features provided by the environment. This concept has been successfully applied in various robotic applications, including grasping [8], [9], motion planning [10] and in-hand manipulation [11]. Our research is inspired by [8], [11], in which manipulation tasks are decomposed into a sequence of ECEs. In this study, we leverage this insight for LfD. Two characteristics distinguish our work from them. First, we propose to extract a sequence of ECEs from a human demonstration rather than using manually designed actions or highly engineered controllers. Second, we consider the problem in which it is difficult to derive the knowledge of ECs using visual information. The robot has to reveal additional information by actively interacting with the environment [12].

Recently, Klingebeitl et al. showed the experimental result that humans tend to explicitly control and explore only a few contact states during manipulation [13]. Several works investigate extracting contact states using geometric information [14] or data-driven approaches [15]. In contrast, we tackle the problem involving extensive contact states due to the complex geometric model. Moreover, different contact states can produce similar

sensory measurements in a chaotic real-world environment. Therefore, applying these approaches to solve our problem is challenging.

One application of ECE for operating unknown mechanisms is introduced in [16]. The idea is to use an adaptive controller to find and follow the admissible motion direction restricted by ECs. This concept exploits helpful guidances provided by ECs and does not require object-specific information, which ensures generalization. However, the major problem of this concept is that the adaptive controller cannot select a specific EC to exploit. Consequently, using this concept to operate articulated objects with multi-DOFs is challenging. To overcome this problem, we propose a method for extracting the necessary parameters from a human demonstration to instantiate a sequence of adaptive controllers to exploit different ECs. In addition, we present the idea of deliberately exerting a force to maintain contact. It reveals useful information about contact-changing events, which facilitate the transitions of ECs.

Research on learning geometric constraints of ECs from demonstrations is related to our work [17], [18]. Our experiments show that it is hard to explicitly learn the geometric constraints for ECs. Furthermore, detailed geometric models are specific to the object instance. Using this information undermines the generalization ability. By contrast, our approach concentrates on extracting general information about EC, leaving unnecessary details to the control. In this way, our approach carefully balances learning and control, thus delivering substantial generalization.

III. ENVIRONMENTAL CONSTRAINTS IN LEARNING FROM DEMONSTRATION

We now present the key arguments for ECs as the appropriate representation for robust and general LfD. While this letter demonstrates this only in the context of contact-rich manipulation, we believe that our general arguments transfer to LfD as a whole.

Environmental constraints encode essential information about the intended interactions between an agent and its environment. There are three possible reasons. First, this can be the consequence of human design: The functionality of articulated objects (scissors, laptops, drawers, can openers...) is encoded and—most importantly—produced based on their articulation. Each articulation provides an environmental constraint that is operated when using the object. Second, environmental constraints are relevant to the robust execution of many skills, for example, when we slide a credit card to the edge of the table to pick it up reliably [1]. And finally, even contact-free actions are naturally expressed through environmental constraints, if one includes visual constraints, such as the servoing of an end-effector to a particular placement (which acts as a constraint). This fundamental interwovenness of environmental constraints with the interactions of embodied agents with their environment justifies ECs central role in LfD.

Interpreting demonstrations as sequences of ECs allows us to separate general information (associated with the EC) and instance-specific information. The properties of the EC (motion

and force directions and contact-changing events) will be identical across different object instances of the same type. But the exact trajectory, how far exactly the motion needs to be or how large a force needs to be, those are properties that vary across object instances. The latter is not part of the description of the EC. The use of ECs as the underlying representation for LfD thus enables the identification of motion patterns that generalize within object categories.

The general but incomplete information captured by ECs must be completed before a demonstration can be transferred successfully. Of course, a transferred demonstration must result in a fully specified trajectory. This does require the specification of information not captured by ECs, such as motion distances, force thresholds etc. To achieve this, we rely on adaptive control. We augment the motion template represented as ECs with information obtained directly from the current object instance by these adaptive controllers. This balances the responsibilities between the learning component and the execution component of LfD. Each component address one-half of the problem: The learning extracts the generalizable motion template and the executive uses sensing in the object-specific case to fill in missing details.

The key advantage offered by the inclusion of adaptive controllers with explorative behavior into the LfD pipeline. Rather than executing a given trajectory, they explore interactions with the world to obtain important information about the object instance that was not present in the EC-based representation. This contributes greatly to the robustness of LfD, as can be illustrated by the following examples. When the execution of a fixed trajectory would get “stuck” due to modeling or execution inaccuracies, the explorative movement of the adaptive controllers will fill in the right details based on sensing, allowing to get “unstuck” again. The explorative movement of the controllers also alleviates the effects of perceptual aliasing. Even though different contact states might give very similar sensor feedback in a particular instance, explorative behavior can quickly disambiguate these states.

For the reasons laid out above, ECs should be used as a representation to achieve robust and general LfD. In the remainder of the letter, we demonstrate this in detail and successfully for contact-rich manipulation.

IV. DEMONSTRATION SYSTEM

We explain here how to extract general information (i.e. motion and force directions) from a human demonstration. To acquire demonstration data, we use a Franka Emika Panda robot arm with a 1-DOF gripper as the robot platform, as shown in Fig. 1. The fingertips of the gripper are covered with silicon to allow the robot to grasp articulated objects with irregular shapes. We mount a Force-Torque sensor on the robot’s wrist to record the interaction force. The black square shape handle mounted on the wrist is used in the kinesthetic teaching phase, in which the user will hold the handle and move the robot to finish a task. Neither vision nor motion capture system is used. The demonstration is recorded in 100Hz frequency. We record end-effector position $p \in \mathbb{R}^3$ and interaction forces $f \in \mathbb{R}^3$ using the Force-Torque sensor. Both position and force measurements are in the robot’s

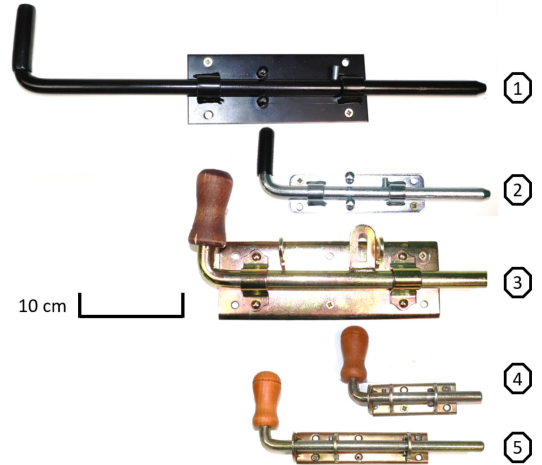


Fig. 2. The five locks used in the experiments.

base frame. Then the demonstrated trajectory Tr is decomposed into a set of vector pairs i.e. $Tr = \{(p_1, f_1), \dots, (p_n, f_n)\}$ with n timesteps.

Given a demonstration, we first decompose it into a sequence of segments. The transition between two environmental constraints is a distinct contact-changing event. For example, the robot is sliding the knob to the left. The sliding motion terminates when the pin on the bar hits the constraints, and the robot’s velocity will suddenly decrease. Based on this insight, we use such events to segment the demonstrations. Concretely, we add a segment when the velocity suddenly drops or the motion direction changes rapidly. We use zero-velocity crossings (ZVCs) [19] as the segmentation algorithm. As ZVCs often over-segmented the demonstration, we filter out the segmentation points that are very close to each other or with similar motion directions. Then we get $Tr = \{D_1, D_2, \dots, D_k\}$ with k segments. Given a segment D_h with T timesteps, we calculate the relative motion direction \hat{m}_h as a 3D unit vector using the initial p_{start} and end position p_{end} . We calculate the force direction \hat{f}_h as a 3D unit vector perpendicular to \hat{m}_h to maintain contact

$$\hat{f}_h = (1 - \hat{m}_h \cdot \hat{m}_h^T) \frac{1}{T} \sum_{i=1}^T \frac{f_i}{|f_i|} \quad (1)$$

A human demonstration is interpreted as a sequence of unit vector pairs of motion direction \hat{m} and force direction \hat{f} . We extract an ECE from each segment i.e. $ECE = (\hat{m}, \hat{f})$, as shown in Fig. 3. In the next section, we explain how to use these motion and force directions to instantiate adaptive controllers.

V. MODELING ENVIRONMENTAL CONSTRAINT EXPLOITATION

In this section, we describe how to model an ECE as an adaptive controller which incorporates the idea of Interactive Perception [12]. In addition, we explain why our ECE controller can take care of the object-specific details. Our idea comes from the observation of how humans open a door with a handle. First, they rotate the handle and let the kinematic constraint guide the motion. Simultaneously, they apply a force in the

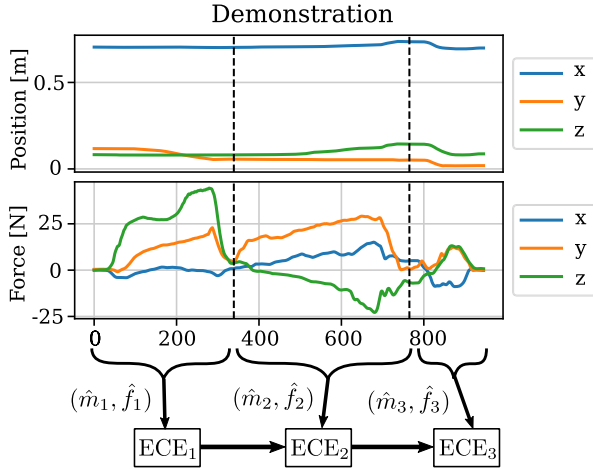


Fig. 3. This figure shows how we extract motion and force direction to instantiate EC exploitations from a human demonstration.

direction of pulling the door. This force reveals a measurable contact-changing event by maintaining contact to help people switch from rotating to pulling.

Therefore, there are two parts for an EC exploitation: 1) a controller which can follow constraints 2) deliberately exerting a force to maintain contact to reveal useful information in contact changes. Note that the force direction \hat{f} extracted from the demonstration is used to reproduce this maintaining contact behavior.

A. Constraint Following

We explain how to reproduce the constraint following behaviors. Our approach is inspired by the idea of following a motion direction with the least resistance [16]. Our adaptive controller consists of a velocity-based impedance controller and Particle Filter Optimization (PFO) [20]. The main difference between PFO and conventional particle filter is that PFO actively samples and executes an action from the estimated distribution and uses the outcome of the action as the measurement to update the distribution. Instead of estimating a particular distribution, the aim of PFO is to find the optimal p^* based on sampling. We parameterize the motion direction \hat{m} as azimuth θ and elevation ϕ in the spherical coordinate system. Each particle in PFO represents a motion direction, namely $p = (\theta, \phi)$. The optimal particle $p^* = (\theta^*, \phi^*)$ is the admissible motion direction. In addition, the optimal particle has the largest weight and vice versa.

The PFO algorithm works as follow. During one episode, PFO biases the particles' motions by adding a white noise $w \sim \mathcal{N}(0, \sigma^2)$ (prediction step). Following that, PFO samples a motion direction $x_{\text{sample}}(\theta, \phi)$ and executes it i.e. moving end-effector into this direction. Then PFO will evaluate the action's outcome and assigns a new weight to each particle (update step). The evaluation criteria is based on the observed movement. If the movement is larger than a threshold (2 mm), PFO will increase the weights of the particles surround with the observed motion direction: $w_{k+1}^i = w_k^i \cdot P(w_k^i | x_k)$, where

w_{k+1}^i denotes the weights of the i particle at the $k + 1$ iteration, motion direction x_k is derived from the observed end-effector movement Δp . The probability density function $P(\cdot | x_k)$ is defined as a normal distribution with mean x_k and predefined variance σ .

The key idea is this: if PFO finds an admissible motion direction, the particles will surround this admissible motion direction. Consequently, the subsequent sampling will be close to the admissible motion direction. Therefore, PFO will drive the robot into the admissible motion direction i.e. following constraints. On the other hand, if the sampled action does not result in motion, the particle distribution will become broad, and its range is increased due to the Gaussian noise. As a result, sampling from a large distribution will be explorative. Unlike approaches that only use estimated admissible motion direction to move the robot [16], [21], our approach introduces the explorative property by sampling action from the particle distribution. This extension allows the robot to generate probing actions to actively reveal the information of admissible motion direction, which is an example of Interactive Perception [12]. It also avoids the problem of getting stuck in the friction cone.

B. Maintaining Contact

We then introduce how to use the force direction extracted from the demonstration to maintain contact while following an EC. As the force direction depends on motion direction, as (1), we first calculate the observed motion direction \hat{m}_{obs} by averaging the observed movements

$$\hat{m}_{\text{obs}} = \text{norm} \left(\frac{1}{N} \sum_{i=1}^N \Delta p_i \right). \quad (2)$$

N is the number of observed movements during the execution of the ECE_h . The Δp_i is the i^{th} observed end-effector movements, norm is the normalization operator. We resample an action if $\hat{m}_{\text{obs}}^T \cdot x_{\text{sample}} \leq 0$. We do so to prevent the robot from moving back and forth while following an EC.

Once the \hat{m}_{obs} is known, we can calculate the corresponding force direction. Assuming that the robot is executing the ECE_h , we calculate the rotation matrix $R \in SO(3)$ which can map extracted motion direction \hat{m}_h to observed \hat{m}_{obs} along the sphere geodesic (shortest path) under the condition of $\hat{m}_h \neq \hat{m}_{\text{obs}}$ [22]:

$$R = I + \hat{m}_h \cdot \hat{m}_{\text{obs}}^T - \hat{m}_{\text{obs}} \cdot \hat{m}_h^T + \frac{1}{\langle \hat{m}_{\text{obs}}, \hat{m}_h \rangle} (\hat{m}_h \cdot \hat{m}_{\text{obs}}^T - \hat{m}_{\text{obs}} \cdot \hat{m}_h^T)^2. \quad (3)$$

We apply this rotation matrix to align the force direction \hat{f}_h in the current motion direction $\hat{f} = R \cdot \hat{f}_h$, where \hat{f} denotes the corresponded force direction to the observed motion direction \hat{m}_{obs} , as shown in Fig. 4.

C. Executing an Environmental Constraint Exploitation

When the observed motion direction \hat{m}_{obs} and the corresponding force direction \hat{f} are available, we can calculate a virtual

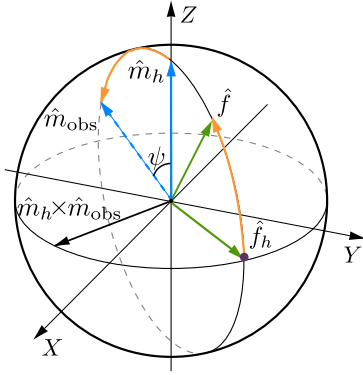


Fig. 4. We use the $\hat{m}_h \times \hat{m}_{\text{obs}}$ as rotation axis and the ψ as rotation angle to calculate rotation matrix to align the \hat{m}_h to \hat{m}_{obs} as well as \hat{f}_h to \hat{f} .

force $F_v \in \mathbb{R}^3$ to move the robot's end-effector

$$F_v = K_m \cdot \hat{m}_{\text{obs}} \cdot v_d + K_f \cdot \hat{f} \cdot v_d \quad (4)$$

where $K_m \in \mathbb{R}^{3 \times 3}$ and $K_f \in \mathbb{R}^{3 \times 3}$ are positive symmetric controller gain matrices for motion and force respectively, $v_d \in \mathbb{R}^3$ is the desired velocity magnitude. We do not control the orientation part. Therefore, we set $\vec{0}$ to angular terms of the desired external force $F_{\text{ext}} \in \mathbb{R}^6$ namely $F_{\text{ext}} = [F_v^T, 0^T]^T$. We map the F_{ext} to the motor torques $\tau_c \in \mathbb{R}^7$ of the robot arm under the operational space schema [23]

$$\tau_c = J^T \cdot F_{\text{ext}} + \tau_{\text{dyn}}(q, \dot{q}, \ddot{q}) + (I - J^T J^{T+}) \cdot \tau_{\text{null}}, \quad (5)$$

where $J \in \mathbb{R}^{6 \times 7}$ is the Jacobian, $J^{T+} \in \mathbb{R}^{6 \times 7}$ is the pseudo-inverse of the Jacobian transpose, $\tau_{\text{dyn}}(q, \dot{q}, \ddot{q})$ represents the to be compensated dynamical forces such as gravity and coriolis force of the arm, and $\tau_{\text{null}} \in \mathbb{R}^7$ represents the null-space torques.

VI. CHAINING A SEQUENCE OF ECES

Given a sequence of ECES, we construct a hybrid automaton [24] to reproduce the demonstration. Each state in the automaton represents an ECE. The edges are transition events between ECES. These transitions are based on a contact-changing event that involves two parts: 1) The robot observes movement (≥ 10 mm) in the force direction. 2) The robot hits a constraint i.e. it fails to move after 30 sampling trials.

Once the ECE_{k-1} terminates and the subsequent ECE_k is to be instantiated, the observed motion direction \hat{m}_{obs} can be used to predict the motion direction \hat{m}_k for the ECE_k i.e. choosing the next EC to exploit. Similar to the calculation of \hat{f} as (1), we calculate the rotation matrix R which maps the motion direction \hat{m}_{k-1} of ECE_{k-1} to the \hat{m}_{obs} . We then get the subsequent motion direction by $\hat{m}_k = R \cdot \hat{m}_{\text{obs}}$, which is used to initialize the particles for the PFO.

Maintaining contact plays an important role in executing chains of ECES. First, having to maintain contact reduces the set of admissible motion directions. This facilitates a single ECE. Second, maintaining contact leverages the idea of Interactive Perception [12] to exploit the correlation between forceful interaction and contact-changing events, which indicates the

TABLE I
COMPARISON OF SUCCESS RATES ON ALL LOCKS

Method	Lock 1	Lock 2	Lock 3	Lock 4	Lock 5
DMP w/o Force	0%	0%	0%	0%	0%
DMP with Force	40%	0%	0%	0%	0%
Our Approach	100%	100%	90%	100%	100%

transition between two ECES with higher certainty. Both factors contribute to the robustness of executing policies.

VII. EXPERIMENTS AND RESULTS

We evaluate our approach on locks and cabinet opening tasks. In all experiments, we assume the robot knows how to grasp the knobs or handles. The quantitative results demonstrate generalization and robustness achieved by exploiting environmental constraints in LfD in contact-rich manipulation tasks. We also compare with approaches which uses geometric parameters as the representation.

A. ECs Enable Generalization in LfD

We examine generalization in the context of a lock-opening task. Based on a single demonstration with a single lock (see Fig. 5), we test the transfer of the extracted policy to other locks. We use five locks that vary in appearance and size but share the same locking principle (see Fig. 2). To unlock, we must first slide the knob to the left while pushing down. This motion ends when the pin on the bar hits the constraint (see inset in Fig. 1). We then must lift the knob while pushing against the constraint. Finally, the pin passes the small slot, and the knob is put down to the open position. The video supplement illustrates this functionality.

After recording a single demonstration on lock 1, our LfD approach segments the recorded data into three parts as described in Section IV. Each segment extracts an ECE, consisting of a motion direction and a force direction. To execute our approach in transfer experiments, the desired velocity v_d is set to 0.13 m/s. The control gains are $K_m = \text{diag}(150)$ and $K_f = \text{diag}(75)$. Each transfer experiment starts with a known grasp location. Everything else is determined by the extracted policy. We define a transfer trial to be successful if the lock ends up in the opening position.

We use two versions of Dynamic Movement Primitives (DMP) as a baseline, one based on position control (DMP without force) and one based on position/force control (DMP with force) [25]. All three methods use the same demonstration data. We run 10 trials per lock for each method, for a total of 50 transfers on all locks, based on a single demonstration. The results are given in Table I.

Our approach succeeds 49 out of 50 times.¹ The single failure occurs with lock 3 and we will discuss it in detail in Section VII-C. This demonstrates that our method generalizes very well, overcoming the three challenges we outlined in III. We interpret this as confirmation that exploiting environmental constraints are indeed the proper representation to achieve generalization in LfD. Fig. 6 illustrates the working of the proposed

¹[Online]. Available: <https://www.youtube.com/watch?v=LIT5n7sG74o>

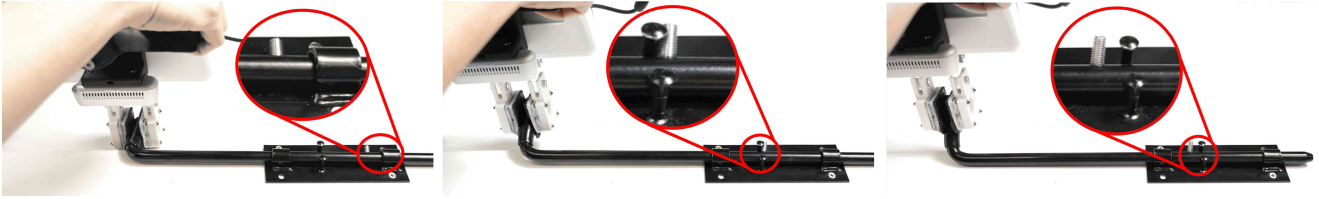


Fig. 5. This figure shows a kinesthetic demonstration on lock 1. We first slide to the left while pushing down (left). We then lift the knob while pushing against the constraint (middle). Finally, we put the pin down while pushing to the left (right). By providing such a demonstration, we can extract useful (force) information to exploit environmental constraints.

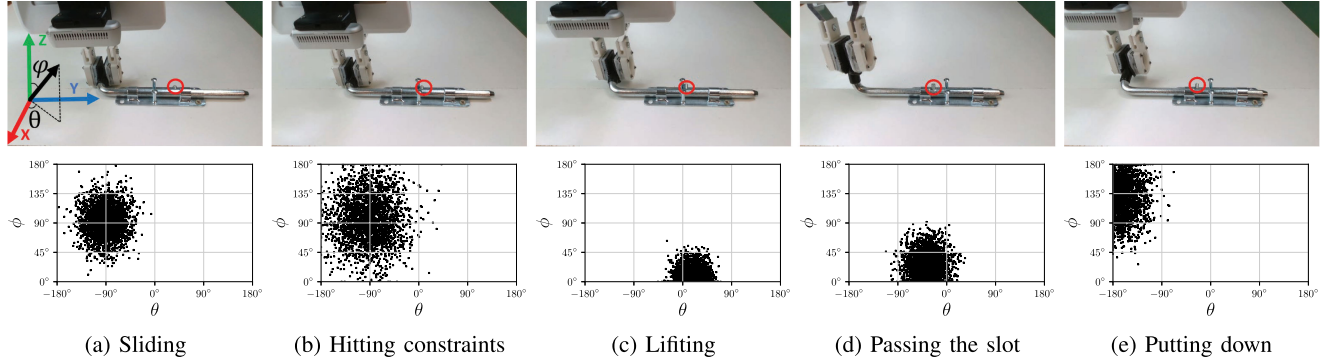


Fig. 6. This figure shows five steps in one opening trial on lock 2. The pin is labeled with a red circle. The second row is the particle distribution. From (a) to (b), the robot hits constraints after sliding. It then tries to sample actions near the sliding direction but fails to observe movements. Therefore, the distribution gradually increases.

method for a transfer experiment. The second row of that figure shows the distribution of particles for sampling actions. We can see that sampling actions from these distributions involve exploration property, which allows the robot to actively probe the environment and regulate the modeling and actuation errors.

The baseline approaches based on DMPs do not transfer well (see Table I). We attribute this to the specific nature of contact-rich manipulation tasks, in which retracing a position-based demonstration is insufficient for generalization. Even worse, the two baselines do not even successfully open the same lock on which the demonstration is provided. The failure modes will be discussed in detail in Section VII-C. This confirms the results obtained in prior works [6], [26].

B. Execution Based on ECs is Robust

We test the robustness of transfer with respect to variations in lock placement. As our method relies on *relative* motion directions, we only need to test variations in lock orientation; generalization to positional placements is automatic. Also here, we claim that robustness is a consequence of exploiting environmental constraints. We place each lock at varying orientations (20° , 40° , and 60°) relative to the orientation during the demonstration. For each of the five locks, we run 10 trials per orientation. The setup is identical to the experiments described in Section VII-A, and we use the same demonstration. Table II contains the results.

Our LfD method achieves high success rates even in the presence of orientation errors up to 40° .² Specifically, the method

TABLE II
SUCCESS RATES OF OUR APPROACH ON OPENING LOCKS WITH INCREASING ORIENTATION ERRORS

Orientation Error (deg)	Lock 1	Lock 2	Lock 3	Lock 4	Lock 5	Avg.
20	90%	100%	100%	80%	100%	94%
40	100%	100%	80%	80%	70%	86%
60	50%	80%	30%	40%	90%	58%

opens the lock successfully in 43 out of 50 trials. In 5 out of 7 unsuccessful trials, the method opened the lock by passing the pin through the slot but failed to achieve the final opening position (see Fig. 1). Even when the locks are rotated by 60° , our approach robustly manipulates lock 2 and lock 5. These results show that the exploitation of ECs enables our approach to compensate for perception and actuation errors, leading to robust transfer of a single demonstration.

C. Failure Analysis

All observed failure cases result from a misestimation of the motion direction associated with the active EC. This false estimation results in a realignment of the expected contact changing event associated with the EC. The motion is thus terminated prematurely and the remainder of the execution is equally misaligned.

There are two reasons for the false estimation of the motion direction. First, significant play between the moved part and the environment can lead to the detection of false motion directions. This happens in the single failure case of our approach in Table I. This type of failure can be avoided by adjusting the motion

²[Online]. Available: <https://www.youtube.com/watch?v=j6sF-RUbrzI>

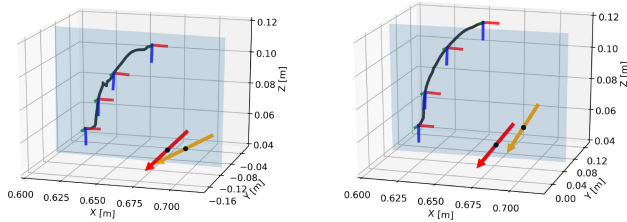


Fig. 7. Extraction of geometric parameters for lock 1 (left) and lock 4 (right), given a demonstration of a lifting action: end-effector trajectory (black line); orientation of the end-effector (coordinate frames); estimated lock axis (orange arrow); ground truth lock axis (red arrow); note the significant modeling inaccuracies (different between orange and red arrows).

TABLE III

ESTIMATION ERRORS FOR LEARNED GEOMETRIC PARAMETERS OF THE LOCK AXIS: POSITION ERRORS ARE CALCULATED BASED ON THE ESTIMATED AND GROUND-TRUTH VALUE OF THE CENTER OF ROTATION

Locks	Geometric Model Errors				
	x (cm)	y (cm)	z (cm)	Orientation (deg)	Rotation Angle (deg)
Lock 1	1.29	0.09	0.11	5.11	9
Lock 4	1.87	0.09	1.23	1.58	16

The orientation errors is computed as the relative angle between estimated and ground-truth orientations. The estimation errors of the axis of the lock are the deviation of the rotation angles.

threshold for detecting a motion direction (Δp) to the amount of play measured by the controller. The second reason for the false estimation of the motion direction is variability in the environment. When the expected motion direction based on the demonstration differs significantly from the actual motion direction, our controller has a difficult time finding the correct direction. This effect explains the data in Table II. The average success rate decreases with increasing orientation error of the lock placement. This type of failure emphasizes the need for visual information to compensate for the inaccuracy [27].

D. Comparison to Explicit Geometric Constraints

ECs are derived from geometric information but discard some object-specific geometric details in order to generalize better. In this section, we present experiments to show that an attempt to estimate the motion geometry of the lock does not lead to a successful generalization of the demonstration. We use least-squares regression [18], [28] to estimate the position and orientation of the lock and the rotation angle of the knob. This estimation relies on a single demonstration (so as to be comparable to our method) with locks 1 and 4. The estimation error compared to hand-measured ground truth is given in Fig. 7 and Table III.

Given the estimated geometric parameters, We attempt to open the locks by generating trajectories for sliding and lifting actions. We track the trajectories using the same impedance controllers and settings as with our method. However, this method fails to open the lock 1, in which the robot gets stuck while lifting the knob, as the protrusion of the black pin constrains the motion of the silver-colored pin (see middle inset in Fig. 5). Furthermore, the method also fails to open lock 4. Here, the estimated rotation angle is not precise enough to be able to pass

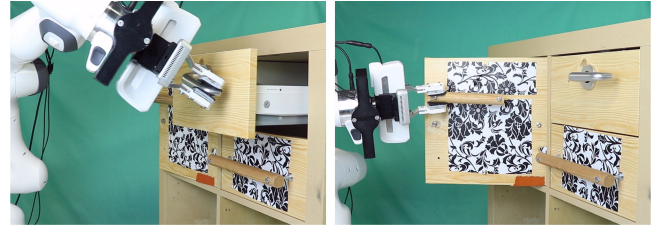


Fig. 8. The robot learns to open the compartments of a cabinet from a single demonstration on a drawer with a door handle.

through the small slot. These experimental results illustrate how difficult it is to extract geometric parameters of sufficient precision to enable generalization in learning from demonstration, even when generalization should happen on the same object.

In addition to this empirical evidence, there are also fundamental challenges to the estimation of exact geometric information from demonstrations. Visual information is subject to occlusions, in particular when manipulating smaller objects. And proprioception/tactile information is often obfuscated by slippage between the robot and the manipulated object. By relying on EC as a representation for demonstrations, we are able to represent the geometric information that supports generalization while relying on controllers to fill in the object-specific details during execution.

E. Other Manipulation Tasks

We also conducted experiments with opening the compartments of a cabinet, each with a different mechanism.³ We have a prismatic mechanism (drawer), a revolute mechanism (door with a handle), and a combined mechanism (drawer that must be opened with a door handle), as shown in Fig. 8. We only record a single demonstration for the combined mechanism. Our approach derives the policy from this demonstration which successively exploits two ECs provided by a revolute joint of the door handle and a prismatic joint. To compute the success rate, we run 15 trials per component with three different grasp locations. Our approach succeeds 45/45 times. This superior success rate stems from that these compartments are fully constrained in 1 DOF, different from locks with multi-DOFs. In such cases, there is no ambiguity about the admissible motion direction. Therefore, estimation of the motion direction becomes easier and our approach circumvents the problem of misalignment, as introduced in Section VII-C. Overall, these experiments provide additional evidence that our approach effectively leverages the information obtained in human demonstrations to produce policies that transfer robustly to similar mechanisms.

VIII. LIMITATIONS

The application domain of our approach is restricted due to two crucial assumptions. First, our approach leverages the guidance provided by the environmental constraints, which assumes the manipulated objects are fully constrained and demonstrations can be segmented using contact-changing events.

³[Online]. Available: <https://www.youtube.com/watch?v=-iju4owhLZ4>

Therefore, our approach is restricted to the task of manipulation of articulated objects, such as doors, drawers or locks. However, we could still apply our central idea of using ECs to a wide range of LfD tasks. For example, learning in-hand manipulation skills from human demonstrations should exploit ECs provided by the fingers to funnel the action outcome [11], and learning grasping strategies from humans should not ignore the insight that humans extensively exploit the constraints present in the environment [1]. The second crucial assumption is that we assume the demonstration is informative i.e. has enough information about ECs. However, if humans are familiar with the manipulated objects, they might give demonstrations which contain much less interaction with the environment [29]. For this reason, it is essential to not only rely on the information from human demonstration, but also actively interact with the environment to fill in the missing information [12], or to online adapt the imperfect information to the environment uncertainty [7], [30].

IX. CONCLUSION

We demonstrate the efficacy of environmental constraints as an underlying representation for general and robust learning from demonstration. Environmental constraints enable the extraction of generalizable information from human demonstrations, separating out instance-specific information that hinders generalization. Using this approach, a single demonstration suffices to achieve generalization to a variety of different instances of the same object type. The proposed method then augments the information by using adaptive controllers. These controllers acquire information about the novel instance through explorative interaction. This combination of general, transferable motion templates and filling in the instance-specific details during execution based on interactive sensing leads to highly robust generalization. We demonstrate this in the context of contact-rich manipulation tasks with articulated objects. Our results validate environmental constraints as a key ingredient for general and robust learning from demonstration.

REFERENCES

- [1] C. Eppner, R. Deimel, J. Álvarez-Ruiz, M. Maertens, and O. Brock, "Exploitation of environmental constraints in human and robotic grasping," *Int. J. Robot. Res.*, vol. 34, no. 7, pp. 1021–1038, 2015.
- [2] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, "Survey: Robot programming by demonstration," in *Springer Handbook of Robotics*, B. Siciliano and O. Khatib, Eds., Berlin, Germany: Springer, 2008.
- [3] S. Schaal, "Is imitation learning the route to humanoid robots?," *Trends Cogn. Sci.*, vol. 3, no. 6, pp. 233–242, 1999.
- [4] P. Pastor, H. Hoffmann, T. Asfour, and S. Schaal, "Learning and generalization of motor skills by learning from demonstration," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2009, pp. 763–768.
- [5] A. Paraschos, C. Daniel, J. R. Peters, and G. Neumann, "Probabilistic movement primitives," in *Proc. Adv. Neural Inf. Process. Syst.*, 2013, vol. 26. [Online]. Available: <https://proceedings.neurips.cc/paper/2013/hash/e53a0a2978c28872a4505bdb51db06dc-Abstract.html>
- [6] M. Kalakrishnan, L. Righetti, P. Pastor, and S. Schaal, "Learning force control policies for compliant manipulation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, 2011, pp. 4639–4644.
- [7] A. Conkey and T. Hermans, "Learning task constraints from demonstration for hybrid force/position control," in *Proc. IEEE-RAS 19th Int. Conf. Humanoid Robots*, 2019, pp. 162–169.
- [8] C. Eppner and O. Brock, "Planning grasp strategies that exploit environmental constraints," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2015, pp. 4947–4952.
- [9] E. Páll and O. Brock, "Analysis of open-loop grasping from piles," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2021, pp. 2591–2597.
- [10] A. Sieverling, C. Eppner, F. Wolff, and O. Brock, "Interleaving motion in contact and in free space for planning under uncertainty," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, 2017, pp. 4011–4017.
- [11] A. Bhatt, A. Sieler, S. Puhlmann, and O. Brock, "Surprisingly robust in-hand manipulation: An empirical study," in *Proc. Robot.: Sci. Syst.*, 2021. [Online]. Available: <http://www.roboticsproceedings.org/rss17/p089.html>
- [12] J. Bohg et al., "Interactive perception: Leveraging action in perception and perception in action," *IEEE Trans. Robot.*, vol. 33, no. 6, pp. 1273–1291, Dec. 2017.
- [13] E. Klingbeil, S. Menon, and O. Khatib, "Experimental analysis of human control strategies in contact manipulation tasks," in *Proc. Int. Symp. Exp. Robot.*, 2016, pp. 275–286.
- [14] J. Xiao and X. Ji, "Automatic generation of high-level contact state space," *Int. J. Robot. Res.*, vol. 20, no. 7, pp. 584–606, 2001.
- [15] S. Scherzinger, A. Roennau, and R. Dillmann, "Contact skill imitation learning for robot-independent assembly programming," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, 2019, pp. 4309–4316.
- [16] G. Niemeyer and J.-J. Slotine, "A simple strategy for opening an unknown door," *Proc. Int. Conf. Robot. Automat.*, 1997, vol. 2, pp. 1448–1453.
- [17] C. Pérez-D'Arpino and J. A. Shah, "C-Learn: Learning geometric constraints from demonstrations for multi-step manipulation in shared autonomy," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2017, pp. 4058–4065.
- [18] G. Subramani, M. Zinn, and M. Gleicher, "Inferring geometric constraints in human demonstrations," in *Proc. Conf. Robot Learn.*, 2018, pp. 223–236.
- [19] J. R. Flanagan, M. C. Bowman, and R. S. Johansson, "Control strategies in object manipulation tasks," *Curr. Opin. Neurobiol.*, vol. 16, no. 6, pp. 650–659, 2006.
- [20] B. Liu, S. Cheng, and Y. Shi, "Particle filter optimization: A brief introduction," in *Proc. Int. Conf. Swarm Intell.*, 2016, pp. 95–104.
- [21] E. Lutscher, M. Lawitzky, G. Cheng, and S. Hirche, "A control strategy for operating unknown constrained mechanisms," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2010, pp. 819–824.
- [22] F. A. F. Tojo, "A constructive approach towards the method of solution-regions," *J. Math. Anal. Appl.*, vol. 472, no. 2, pp. 1803–1819, 2019.
- [23] O. Khatib, "A unified approach for motion and force control of robot manipulators: The operational space formulation," *IEEE J. Robot. Automat.*, vol. 3, no. 1, pp. 43–53, Feb. 1987.
- [24] M. Egerstedt, "Behavior based robotics using hybrid automata," in *Hybrid Systems: Computation and Control*, vol. 1790, Berlin/Heidelberg, Germany: Springer, 2000, pp. 103–116.
- [25] P. Pastor, L. Righetti, M. Kalakrishnan, and S. Schaal, "Online movement adaptation based on previous sensor experiences," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, 2011, pp. 365–371.
- [26] M. Racca, J. Pajarinen, A. Montebelli, and V. Kyrki, "Learning in-contact control strategies from demonstration," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, 2016, pp. 688–695.
- [27] S. Hutchinson, "Exploiting visual constraints in robot motion planning," in *Proc. IEEE Int. Conf. Robot. Automat.*, 1991, pp. 1722–1723.
- [28] R. Martín-Martín, C. Eppner, and O. Brock, "The RBO dataset of articulated objects and interactions," *Int. J. Robot. Res.*, vol. 38, no. 9, pp. 1013–1019, 2019.
- [29] M. Suomalainen, Y. Karayiannidis, and V. Kyrki, "A survey of robot manipulation in contact," *Robot. Auton. Syst.*, 2022, Art. no. 104224, doi: [10.1016/j.robot.2022.104224](https://doi.org/10.1016/j.robot.2022.104224).
- [30] R. Martín-Martín and O. Brock, "Coupled recursive estimation for online interactive perception of articulated objects," *Int. J. Robot. Res.*, 2019, Art. no. 0278364919848850. [Online]. Available: <https://journals.sagepub.com/doi/full/10.1177/0278364919848850>