

Learning Task Constraints from Demonstration for Hybrid Force/Position Control

Adam Conkey¹ and Tucker Hermans^{1,2}

Abstract—We present a novel method for learning hybrid force/position control from demonstration. We learn a **dynamic constraint frame** aligned to the direction of desired force using Cartesian Dynamic Movement Primitives. In contrast to approaches that utilize a fixed constraint frame, **our approach easily accommodates tasks with rapidly changing task constraints over time**. We activate only one degree of freedom for force control at any given time, ensuring motion is always possible orthogonal to the direction of desired force. Since we utilize demonstrated forces to learn the constraint frame, we are able to compensate for forces not detected by methods that learn only from the demonstrated kinematic motion, such as frictional forces between the end-effector and the contact surface. We additionally propose novel extensions to the Dynamic Movement Primitive (DMP) framework that encourage robust transition from free-space motion to in-contact motion in spite of environment uncertainty. We incorporate force feedback and a dynamically shifting goal to reduce forces applied to the environment and retain stable contact while enabling force control. Our methods exhibit low impact forces on contact and low steady-state tracking error.

I. INTRODUCTION

Many tasks, such as wiping a window, scrubbing a floor, and mixing in a bowl require motion along a surface while maintaining a desired force. In order to automate such constrained-motion tasks, robots must be able to control force and position simultaneously. Though forces can be applied to an object using only position control, it is generally unsafe to do so without force feedback since excessively large forces can be imposed on the object (and the robot) in the presence of estimation errors. Controlling forces relative to desired motion is essential for ensuring completion of constrained-motion tasks without risking damage to the environment or the robot.

Hybrid force/position control is a popular control scheme for constrained-motion tasks [1], [2] since position and force control objectives can be tracked simultaneously without conflict. Control is performed with respect to a (possibly time-varying) Cartesian coordinate system $\mathcal{C}_t \in SO(3)$, denoted the *constraint frame*, that may be arbitrarily located in space. Common choices for the constraint frame include the world frame, the robot’s tool frame, or a frame attached to an object of interest [3]. *Task constraints* determine which dimensions of the constraint frame are controlled for position and which dimensions are controlled for force. They are typically defined by a diagonal binary selection matrix $\mathbf{S}_t \in \mathbb{R}^{6 \times 6}$ where $\mathbf{S}_t(i, i) = 1$ activates position control for Cartesian dimension $i \in \{0, \dots, 5\}$ at time t and $\mathbf{S}_t(i, i) = 0$ enables force control.

Specifying an appropriate constraint frame and task constraints is difficult and prone to error for complex tasks. Improper constraint specification is especially problematic for transitioning from free-space motion to **being in contact with a surface**, as large forces may be applied to the surface if constraints are enabled too soon or too late [4], [5]. Even

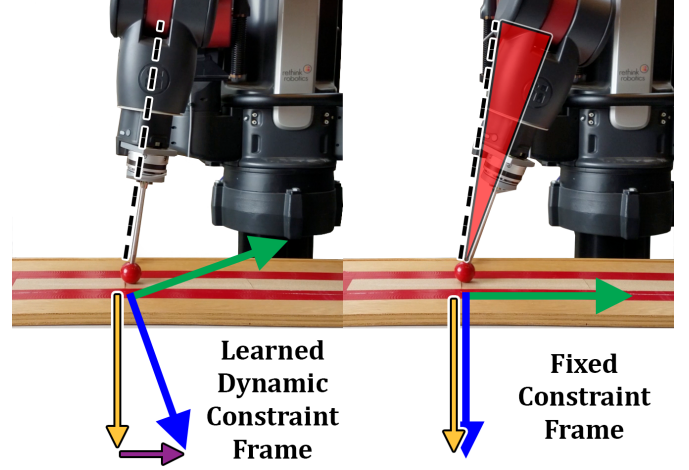


Fig. 1: Illustration of a learned dynamic constraint frame (our method) versus a fixed constraint frame while sliding on a high-friction surface. Green and blue arrows indicate the y and z axes of the constraint frame, respectively, where force control is activated for the z -dimension. The fixed-frame approach tracks the desired force applied to the surface (yellow arrow), but incurs large pose error (red triangle) due to inhibited motion from frictional forces between the end-effector and the contact surface. Our method actively compensates for these frictional forces (purple arrow), thereby tracking both desired force (yellow arrow) and desired pose (dotted black line).

when constraints are properly specified, small perturbations in the environment configuration or **perceptual estimation errors can interfere with the timing of the constraints**. It is therefore desirable for a robot to learn the constraints of a task and to adapt them to environment uncertainty online.

While learning from demonstration has proven successful for **learning task constraints** [6]–[9], existing approaches focus on learning **axis-aligned constraints** with respect to a chosen fixed frame. This is a limitation when desired forces are time-varying and span multiple dimensions of the constraint frame, as the robot loses a degree of freedom for motion with each degree of freedom devoted to force control. **Time-varying task constraints** have been learned from kinematic motion for the purpose of **generalizing null space policies** [10]. However, learning task constraints only from kinematic motion leaves out valuable information contained in **forces** observed during demonstration, such as **frictional forces** between the robot’s tool and the **contact surface** that need to be compensated for.

In this paper, we present a novel approach to learning task constraints for hybrid force/position control from demonstration. We learn a **dynamic constraint frame** using Cartesian Dynamic Movement Primitives (CDMPs) such that **a principal axis of the constraint frame is always aligned to the direction of desired force**. Our approach has the following benefits over existing methods that utilize fixed constraint frames and that learn only from kinematic motion:

- 1) **We accommodate tasks with constraints that change rapidly over time.** We show experimentally on the task

¹School of Computing; Robotics Center; University of Utah, USA.

²NVIDIA, USA. Email: adam.conkey@utah.edu, thermans@cs.utah.edu

of mixing in a bowl, in which forces vary in a non-trivial manner across all three dimensions of commonly used fixed frames, that we are able to **track both motion and force objectives**. The task was not achievable controlling with respect to a fixed constraint frame.

- 2) **We activate only one degree of freedom for force control at any given time.** Our method ensures motion is always possible orthogonal to the direction of desired force, whereas fixed-frame methods can render orthogonal motion impossible if all degrees of freedom need to be activated for force control.
- 3) **We compensate for frictional forces between the end-effector and the contact surface while sliding.** Methods that **learn task constraints only from kinematic motion** neglect crucial information contained in observed forces, such as the frictional forces between the end-effector and the contact surface. We show experimentally on the task of sliding on a high-friction surface that frictional forces induce dramatic pose error when controlling with respect to a fixed constraint frame, while our method actively compensates for these forces and tracks desired force with little pose error (see Figure 1).

We additionally extend the Dynamic Movement Primitive (DMP) framework to encourage robust **transition from free-space motion to constrained motion**. Our extensions incorporate **force feedback and contact awareness to reduce contact forces** and **gradually transition into tracking desired forces**. We also define a dynamically changing goal that transitions as a function of the robot’s contact with the environment. These modifications account for cases in which the surface to be contacted is not exactly at the anticipated position, e.g. due to **perceptual error** or **perturbation of the environment**.

We structure the remainder of the paper as follows. We review related work in the areas of learning force/position control and Dynamic Movement Primitives in Section II. In Section III we present the base methods from the prior art we utilize in our framework. The details of our novel contributions are provided in Section IV. We describe our experimental setup in Section V and present the associated results in Section VI. Section VII concludes with a brief discussion of our methods and directions for future work.

II. RELATED WORK

We review two general areas of related research. The first covers learning simultaneous control of force and position. The second area includes methods that incorporate force feedback into Dynamic Movement Primitives.

A. Learning Force/Position Control

The literature in learning from demonstration for simultaneous control of position and force has focused on 1) learning which dimensions of the constraint frame should be selected for position or force control [6], [9], [11], [12] and, to a lesser extent, 2) learning the best constraint frame to control with respect to [7], [9]. A key insight that has motivated constraint selection methods is that dimensions of the constraint frame that consistently exhibit high variance over time in force and low variance over time in position should favor force control, and position control otherwise [9]. In [9], a criterion based on trajectory variance is defined that modulates a stiffness

parameter of a Cartesian impedance controller, allowing force tracking when stiffness is low. Impedance stiffness is set to zero in [12] for compliant dimensions orthogonal to the dimension of highest variance in motion. A series of boolean checks in [6] over variance in force and position variables determines which axes of the robot’s tool frame is enabled for PI force control or Cartesian impedance control. In [11], binary constraint selection for a hybrid force/position controller is made by enabling position control when the computed position variance is found to be greater than the force variance.

Constraint frames are often chosen manually based on the requirements of the task [1]. Common choices include the world frame [12], [13], surface normals [11], the robot’s tool frame [6], [8], [10], [14], and frames attached to objects of interest in the environment [9]. **The robot selects an appropriate constraint frame from a collection of pre-defined candidate frames in [9] based on the observed trajectory variance over multiple demonstrations.** In [7], candidate frames include the start and end frames of a human-robot collaboration task, and an appropriate frame is chosen as over the motion progresses. However, methods that use a fixed constraint frame cannot be used for a task in which desired forces span all three dimensions of the constraint frame, since it would require all dimensions to be enabled for force control, thereby preventing simultaneous motion. A careful choice of constraint frame can mitigate this problem, but for tasks in which desired forces vary in a complex manner, fixed frame selection is not feasible.

Estimating task constraints and null space projections thereof can be used for generalizing a task to different environment configurations [10], [15]. The task constraint matrix and null space projection are estimated from motion data in [15] and incorporated into an operational space controller. [15] does not, however, consider task constraints for the purpose of force control. In [10], the robot estimates task constraints to command a policy learned from demonstration in the task null space. While [10] uses a force/torque sensor to align the robot end-effector to the normal of a curved surface for generalizing a learned planar task, it does not consider explicit task constraints for force control and assumes the robot is already in contact with the surface before initiating the task.

B. Force Feedback for Dynamic Movement Primitives

Dynamic Movement Primitives (DMPs) are a widely used policy representation for learning robot motion that afford real-time obstacle avoidance [16], dynamic goal changing [17], and can be learned from demonstration using standard regression techniques [18]. Various features of DMPs have been used to augment motion trajectories with force information. Kormushev et al. [19] synchronize position trajectories and force profiles (specified separately with a haptic input device) using the DMP phase variable. In [20], force error is incorporated into the phase variable to aid in assembly tasks learned from demonstration. Temporal coupling terms in [8] provide pose disturbance detection when executing tasks that repeatedly make and break contact with a surface. Compliant Movement Primitives [21] **encode both motion and joint torques** as a DMP to reduce contact forces during unexpected collisions. Velocity in periodic DMPs is modulated based on a passivity criterion in [22] to efficiently perform wiping tasks in a stable manner.

Having both motion trajectories and force profiles encoded as DMPs allows standard reinforcement learning methods such as PI^2 to be readily applied in order to learn the optimal forces needed for completing a task [13], [23].

Kober et al. [24] learn DMPs for individual segments of a multi-phase task and achieve force and position tracking with a hybrid force/position controller. However, [24] selects a fixed constraint frame based on convergence metrics of the DMPs, whereas our method uses a dynamic constraint frame learned from forces observed during demonstration. Several complementary works to ours use force information to guide transitions between primitives [25]–[27], but they do not address the problem of robustly transitioning between **free-space motion and in-contact task phases**. Steinmetz et al. [8] handle the case of ensuring contact when an expected contact is not satisfied, but require switching between multiple controllers, which is known to suffer stability issues [28]. Additionally, [8] cannot adapt to contacts made sooner than expected. Our extensions to the DMP framework enable robust transitions from free-space motion to constrained motion using a single unified controller.

III. BACKGROUND

In this section, we present the base methods we employ in our framework. We first define the hybrid force/position control law we use in Section III-A, and then present a standard formulation of DMPs in Section III-B. Our novel contributions will be presented in Section IV.

A. Hybrid Force/Position Controller

We utilize the operational space hybrid force/position controller defined in [29] which we present here for clarity. The controller we use has the form

$$\boldsymbol{\tau} = \boldsymbol{\tau}_f + \boldsymbol{\tau}_x + \left[\mathbf{I}_n - (\mathbf{J}^\# \mathbf{J})^T \right] \boldsymbol{\tau}_0 + \mathbf{g} \quad (1)$$

where $\boldsymbol{\tau}_x$ and $\boldsymbol{\tau}_f$ are joint torques corresponding to position and force control laws, respectively, $\boldsymbol{\tau}_0$ is an arbitrary joint space control law to be commanded in the null space of hybrid force/position control, and \mathbf{g} is gravity compensation in joint space. $\mathbf{I}_n \in \mathbb{R}^{n \times n}$ is the n -dimensional identity matrix where n is the number of robot joints, $\mathbf{J} \in \mathbb{R}^{6 \times n}$ is the analytic Jacobian, and $\mathbf{J}^\#$ is the generalized Jacobian pseudo-inverse derived in [29] as

$$\mathbf{J}^\# = \mathbf{M}^{-1} \mathbf{J}^T \boldsymbol{\Lambda} \in \mathbb{R}^{n \times 6} \quad (2)$$

where \mathbf{M} is the joint space inertia matrix and $\boldsymbol{\Lambda}$ is the inertia matrix reflected into task space defined by

$$\boldsymbol{\Lambda} = [\mathbf{J} \mathbf{M}^{-1} \mathbf{J}^T]^{-1} \in \mathbb{R}^{6 \times 6} \quad (3)$$

We use the null space projection to command a low-gain PD controller $\boldsymbol{\tau}_0$ in joint space that tracks a desired posture keeping the robot away from joint limits when possible.

We use a Cartesian inverse dynamics controller defined as

$$\boldsymbol{\tau}_x = \mathbf{J}^T \boldsymbol{\Lambda} \boldsymbol{\Omega} \left(\mathbf{K}_p(\mathbf{x}_d - \mathbf{x}) + \mathbf{K}_d(\dot{\mathbf{x}}_d - \dot{\mathbf{x}}) + \ddot{\mathbf{x}}_d - \dot{\mathbf{J}}\dot{\mathbf{q}} \right) \quad (4)$$

where $\mathbf{x}_d, \dot{\mathbf{x}}_d, \ddot{\mathbf{x}}_d$ are desired Cartesian poses, velocities, and accelerations, $\mathbf{x}, \dot{\mathbf{x}}$ are actual poses and velocities, $\dot{\mathbf{q}}$ are joint

velocities, and $\mathbf{K}_p, \mathbf{K}_d \in \mathbb{R}^{6 \times 6}$ are positive semi-definite gain matrices. $\boldsymbol{\Omega} = \boldsymbol{\Omega}(\mathbf{S})$ is a block tensor transformation that performs selection for position control in the constraint frame:

$$\boldsymbol{\Omega}(\mathbf{S}) = \begin{bmatrix} \mathbf{R}^T \mathbf{S} \mathbf{R} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}^T \mathbf{S} \mathbf{R} \end{bmatrix} \in \mathbb{R}^{6 \times 6} \quad (5)$$

for $\mathbf{R} = {}^c\mathbf{R}_0$ the rotation matrix from the base frame to the constraint frame and \mathbf{S} the selection matrix defined in Section I. Note that \mathbf{S} and \mathbf{R} generally vary with time but we drop the subscripts here for consistency with the controller definition.

We control forces using the following PI control law

$$\boldsymbol{\tau}_f = \mathbf{J}^T \tilde{\boldsymbol{\Omega}} \left(\mathbf{K}_f(\mathbf{F}_d - \mathbf{F}) + \mathbf{K}_I \sum_{t-\Delta_t}^t (\mathbf{F}_d - \mathbf{F}) \right) \quad (6)$$

where $\mathbf{K}_f, \mathbf{K}_I \in \mathbb{R}^{6 \times 6}$ are positive semi-definite gain matrices, $\mathbf{F}_d, \mathbf{F} \in \mathbb{R}^6$ are desired and actual forces, and Δ_t is the window of error accumulation. $\tilde{\boldsymbol{\Omega}} = \boldsymbol{\Omega}(\tilde{\mathbf{S}})$ is the force control selection matrix where $\tilde{\mathbf{S}} = \mathbf{I}_6 - \mathbf{S}$. Force tracking occurs for each dimension of the constraint frame that has $\mathbf{S}(i, i) = 0$. We can achieve pure free-space motion by setting $\mathbf{S} = \mathbf{I}_6$.

In Section VI-C.1 we experimentally compare against PI force control with Integral Error Scaling (IES) presented in [5]. This technique attenuates the integral error when it opposes the desired direction of force in order to mitigate the chance of the end-effector breaking contact with the surface. For IES, the integral error term in Equation (6) $\Delta \mathbf{F} = \mathbf{F}_d - \mathbf{F}$ switches to $\Delta \mathbf{F} = \beta(\mathbf{F}_d - \mathbf{F})$ for $\beta \in [0, 1)$ when $\mathbf{F}_d - \mathbf{F} < 0$.

B. Dynamic Movement Primitives

We learn DMPs for position trajectories and force profiles following the formulation of [30] characterized by:

$$\tau \dot{v} = \alpha_v(g - y) - \beta_v v - \alpha_v(g - y_0)s + \alpha_v f(s) \quad (7)$$

$$\tau \dot{y} = v \quad (8)$$

$$\tau \dot{s} = -\alpha_s s \quad (9)$$

$$f(s) = \frac{\sum_i w_i \Psi_i(s)}{\sum_i \Psi_i(s)} s \quad (10)$$

$$\Psi_i(s) = \exp(-h_i(s - c_i)^2) \quad (11)$$

Equations (7) and (8) define a first order critically damped dynamical system for an appropriate choice of $\alpha_v, \beta_v \in \mathbb{R}$ where y is the state variable being tracked, y_0 is the initial state, g is the goal, and $f(s)$ a forcing function. Equation (9) specifies the evolution of a phase variable that decouples the system from explicit time. Equation (10) defines the forcing function as a normalized linear combination of basis functions. We use Gaussian basis functions as is common in the literature [30], where Equation (11) parameterizes them with centers c_i and widths h_i . Each degree of freedom receives its own DMP which are synchronized by the common phase variable s .

For orientation trajectories, we learn Cartesian Dynamic Movement Primitives (CDMPs) similar to [30] but with the full quaternion error as suggested in [31]:

$$\tau \dot{\boldsymbol{\omega}} = \alpha_\omega \delta(\mathbf{q}, \mathbf{q}_d) - \beta_\omega \boldsymbol{\omega} - \alpha_\omega \delta(\mathbf{q}_0, \mathbf{q}_d)s + \alpha_\omega \mathbf{f}(s) \quad (12)$$

$$\tau \dot{\mathbf{q}} = \frac{1}{2} \boldsymbol{\omega} * \mathbf{q} \quad (13)$$

Equations 12 and 13 are analogous to Equations 7 and 8 where we define the difference function for quaternions $\mathbf{q}_i = (v_i, \mathbf{u}_i)$ as $\delta(\mathbf{q}_1, \mathbf{q}_2) = 2\log(\mathbf{q}_2 * \bar{\mathbf{q}}_1)$. Here $\bar{\mathbf{q}}$ denotes quaternion conjugation and the quaternion product is

$$\mathbf{q}_1 * \mathbf{q}_2 = (v_1 v_2 - \mathbf{u}_1^T \mathbf{u}_2) + (v_1 \mathbf{u}_2 + v_2 \mathbf{u}_1 + \mathbf{u}_1 \times \mathbf{u}_2)$$

IV. METHODS

We now present the details of our novel contributions. We describe our novel approach to learning task constraints with a dynamic constraint frame in Section IV-A, and our novel extensions to the DMP framework in Section IV-B that allow for robust transition from free-space to in-contact motion.

A. Learning Time-Varying Task Constraints

Instead of learning the selection matrix \mathbf{S}_t for a fixed constraint frame (as in, e.g. [6], [9]–[12]), we propose learning a dynamic constraint frame \mathcal{C}_t for which \mathbf{S}_t can be specified in a canonical way. Our key insight is that, at each timestep, we can align a principal axis of the constraint frame to the direction of desired force $\mathbf{F}_{d,t}$, thereby requiring only one degree of freedom for force control. We set the z -axis¹ to be axis-aligned to desired forces observed during demonstration, resulting in selection matrix values of $\mathbf{S}_t(2,2) = 0$ and $\mathbf{S}_t(i,i) = 1$ otherwise. This corresponds to force control along the z -axis of the constraint frame and position control on all other axes.

We create the input to the learning procedure from the forces observed during demonstration by defining the z -axis at each time step to be the observed force vector normalized to unit length. We construct the other axes by selecting the end effector y -axis as a candidate orthogonal axis and use cross products to create a valid right-handed coordinate system. We learn a single CDMP (described in Section III-B) from the constructed input data using ridge regression. The output is a smoothly varying trajectory for \mathcal{C} with a z -axis that tracks the direction of desired force. We obtain a smoothly varying estimate of the magnitude of desired forces $\|\hat{\mathbf{F}}_d\|$ to be applied along the z -axis of \mathcal{C} by learning a DMP from $\|\mathbf{F}_d\|$. Our method inherits the generalization benefits of DMPs well known in the literature [18]. Thus, any modulations applied to the robot’s motion (e.g. temporal modulation) can also be applied to the learned constraint frame and desired forces, ensuring motion and force objectives remain in sync.

We show in our experiments in Section VI-B that controlling with respect to our learned constraint frame allows desired forces to be tracked using one degree of freedom for force control, even when desired forces span multiple dimensions of fixed reference frames such as the world or tool frames. We also show in Section VI-A.2 that we achieve compensation of frictional forces while sliding without explicitly modeling frictional properties of the robot or the environment. This improves upon the typical hybrid force/position control paradigm that makes the simplifying assumption of frictionless contact [3]. Previous approaches for learning hybrid force/position control from demonstration (e.g. [8], [11], [24]) do not discover these forces and rely on low-friction environments to demonstrate their methods.

¹The choice of z is arbitrary.

B. Extended DMPs for Making Stable Contact

We extend the DMP framework for the purpose of encouraging robust transition from position control to force control when making contact with a surface.

1) *Halt DMP at Surface Contact*: To bring the system to a halt when the robot detects contact, we modify Equation (8):

$$\tau \dot{y} = \frac{v}{1 + \alpha_f \sigma(f) |f|} \quad (14)$$

where f is the sensed force in the same task space dimension as y and $\alpha_f \in \mathbb{R}$ determines how sensitive the system is to contact forces. We define the contact classifier $\sigma(f)$ as

$$\sigma(f) = \begin{cases} 1 & \mu_w > \mu_0 \\ 0 & \text{otherwise} \end{cases} \quad (15)$$

where μ_w is the mean value of $|f|$ over a sliding window of size n and μ_0 is the mean of the noise inherent to the sensor. We show in Section VI-C.1 that our method lowers impact forces when contacting a surface earlier than anticipated.

The right-hand side of Equation (14) has a similar form to a term proposed in [31] for halting a DMP system when pose error accumulates and in [20] when force error accumulates. However, in [20] and [31] the terms are applied to the phase variable and not the transformation system. We apply our term directly to the transformation system velocity as it allows us to selectively decouple the halting behavior of different dimensions. We show in our experiments (Section VI-C.2) that the robot can halt motion in a dimension with an expected contact, while the remaining unconstrained dimensions continue to converge to their desired goal states. This could not be achieved if the term were used in the shared phase variable, as it synchronizes control across all dimensions.

2) *Change in Goal Based on Contact Conditions*: If we assume the robot made the intended contact, but at an earlier time (see Section VII for a further discussion of this assumption), then the modification in Equation (14) alone does not suffice for completing the task. The DMP will remain in a halted state until the force disappears, which will not happen when the force is due to an intended contact and not a transient disturbance. We instead desire the free-space DMP system to gracefully terminate its execution and transition into the in-contact phase of the task. We achieve this by allowing the goal to dynamically change determined by

$$\dot{g} = \alpha_c \sigma(f)(y - g) + \alpha_{nc}(1 - \sigma(f))(g_0 - g) \quad (16)$$

where g_0 is the original goal, g is the current goal, y the current DMP state, and $\sigma(f)$ the contact classifier in Equation (15).

Equation (16) smoothly moves the current goal to coincide with the robot’s current state when the robot detects stable contact. Once the goal and state coincide, the robot ends the free-space task phase and transitions to the in-contact phase. If a disturbance caused the sensed force and it disappears before the transition occurs, Equation (16) affords a smooth transition back to the original goal and the phase proceeds from that point as it would if no contact had been made. Parameters $\alpha_c, \alpha_{nc} \in \mathbb{R}$ control the rate of goal transition.

When the surface is farther than expected, the pose DMPs will converge to their respective goals before making contact

with the surface. At goal convergence, each term in Equation (7) approaches zero, but we can still incite movement toward the desired contact by moving the goal g in the direction of the desired contact by a small amount ϵ . This moves the end-effector at a constant velocity towards the desired contact, achieving similar behavior to [8] and [28]. Our method is advantageous over these methods as we do not require controller switching [8], and we only require a single demonstration as opposed to hundreds of real-robot trials [28].

3) *Incremental Force Control on Contact*: When the robot first makes contact with the surface, an initial impact force will be applied to the surface that depends on the velocity at impact; a higher approach velocity results in a higher impact force. Though we mitigate impact forces with the DMP feedback in Equation (14), we still desire to enable force control when in contact in order to avoid sustained application of high impact forces and to gracefully transition into the constrained motion phase of the task. However, when the force error is large, enabling force control instantaneously can make retaining stable contact with the surface difficult, particularly for a stiff environment [4], [5].

We propose to overcome this difficulty by incrementally enabling force control for the desired dimension by leveraging the gradual goal transition of Equation (16). Instead of a strictly binary selection matrix \mathbf{S}_t for the hybrid force/position controller, we allow the Cartesian dimension i transitioning to force control to continuously vary from 1 to 0 determined by

$$\mathbf{S}_t(i, i) = \exp\left(1 - \frac{|y_c - g_c|}{|y - g|}\right) \quad (17)$$

where y_c is the system state at the time of contact, y is the current system state, g_c is the DMP goal at the time of contact, and g is the current DMP goal. This expression initializes to 1 when the robot initially makes contact, and converges to 0 as the goal g converges to the current system state y . This allows the controller to smoothly transition from position control to force control as $\mathbf{S}_t(i, i)$ runs through convex combinations of the two control laws. We show in Section VI-C.1 that this technique affords stable contact and steady-state tracking when making contact at different approach velocities.

V. EXPERIMENTAL SETUP

We validate our methods on a Baxter robot equipped with a 6-axis Optoforce HEX-E-200N force-torque sensor at the wrist. Both the robot state and the force-torque sensor state² are sampled at 1000Hz. Robot controllers operate at a rate of 1000Hz. The end-effector is a hard plastic sphere threaded to the tip of a steel shaft which affords a point contact that can vary easily over the course of the trajectory. Experiments were performed using an Intel Core i7-4770 CPU @ 3.40GHz x 8 computer with 8GB of RAM running Ubuntu 14.04 and ROS Indigo. All software and data associated with this paper will be made available upon publication.

²We employ an online low-pass filter on the force sensor with a cutoff frequency of 1.5Hz. This is a lower cutoff frequency than typically used, but the added noise reduction is beneficial for learning from the sensor readings. The filter adds a small time delay on the order of 20 milliseconds, which is insignificant given the 1000Hz sampling rate of the sensor.

We provide kinesthetic demonstrations by manually moving the robot arm in gravity-compensation mode. Once recorded, the system autonomously segments the demonstrations using the contact classifier in Equation (15) into three phase types: making-contact, in-contact, and breaking-contact. Desired goal forces for making contact are equal to the initial desired forces for the sliding phase. A DMP is learned for each DOF in each task phase as described in Sections III-B and IV-A. DMP parameters were set according to guidance in the prior art [18]. All DMP and controller parameters are kept the same in all experiments unless otherwise stated in the text. We now lay out an overview of our experimental protocol; we present the associated results in Section VI.

A. Sliding on a flat surface.

We show that our learned dynamic constraint frame (Section IV-A) actively compensates for frictional forces between the end-effector and the contact surface while sliding. The demonstrator is instructed to slide the end-effector along the surface while applying a small force and keeping the end-effector perpendicular to the table surface. We conduct experiments across both low and high-friction surfaces and compare using our learned dynamic constraint frame and the fixed world frame. Our results show that controlling with respect to our learned constraint frame affords accurate tracking of both the desired force profile and pose trajectory, while controlling with respect to a fixed frame results in considerable pose error due to frictional forces inhibiting the end-effector's motion.

B. Sliding on a curved surface.

We demonstrate the ability of our learned constraint frame to easily accommodate tasks where the constraints vary rapidly over time with the task of performing a mixing motion on the interior curved surface of a bowl. Desired forces for the task vary across all three dimensions of commonly used frames such as world and tool frames, preventing simultaneous motion when fixed-frame control approaches are utilized. Our method, on the other hand, tracks the desired force profile and pose trajectory using only one degree of freedom for force control.

C. Making contact with a surface

We demonstrate the benefits of our methods presented in Section IV-B for making robust contact with a surface. We perform tests for making stable contact at varying table heights and approach velocities, when given a single demonstration for moving the end-effector straight down from an initial position above the table to a desired contact point on the table at a nominal height. We compare against using the standard DMP formulation with no force feedback ("open-loop" below) and against using PI force control with and without Integral Error Scaling. We also consider the case where the end-effector moves at an angled approach from above the table to a desired contact point on the table. These final experiments illustrate the advantage of putting our DMP feedback on the transformation system instead of the canonical system.

VI. RESULTS

We now present our experimental results following the same structure outlined in Section V.

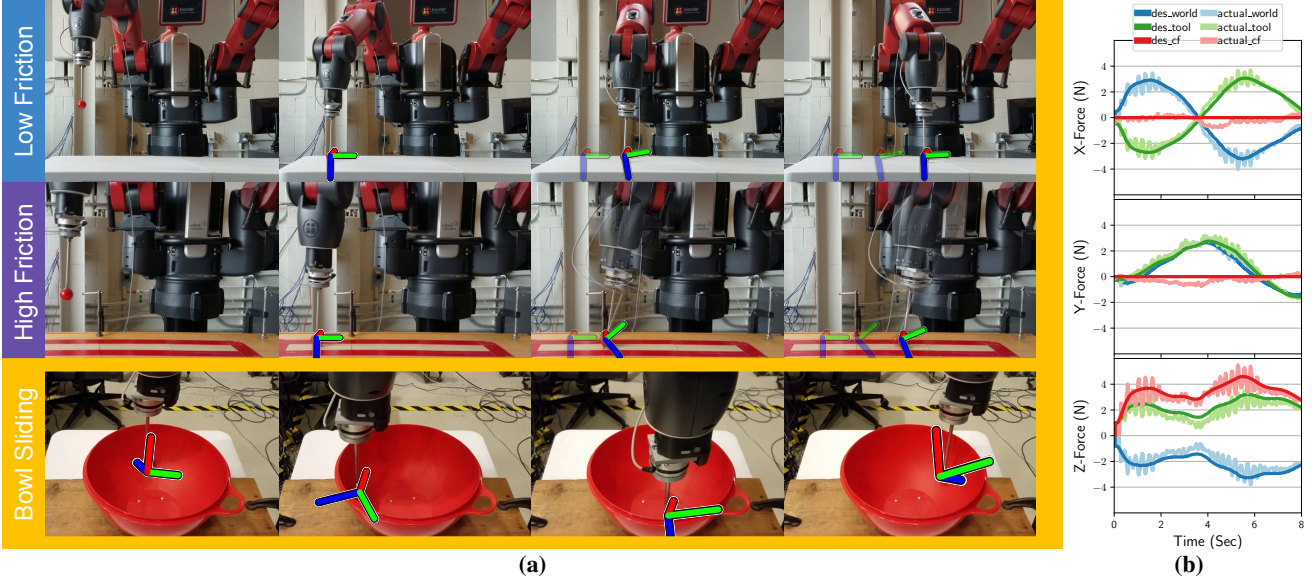


Fig. 2: (a): Example scenarios performing hybrid force/position control with respect to a learned constraint frame. We test sliding on low and high friction surfaces as well as a curved (bowl) surface. Coordinate frames show the pose of the learned constraint frame over time (red = x , green = y , blue = z). Row 2 shows the difference in pose deviation for controlling with respect to the learned constraint frame (primary image) and the world frame (semi-transparent overlay). **(b):** Compares the desired and actual forces observed in the world frame, tool frame, and learned constraint frame while controlling with respect to the learned constraint frame for the bowl scenario.

A. Sliding on a flat surface.

We perform the sliding task on two surfaces with drastically different friction properties: a smooth plastic table and a piece of wood covered with sand paper. In each case, the demonstrator slides the end-effector along the surface while applying a small force and attempts to keep the end-effector oriented perpendicular to the surface.

1) *Low-friction surface:* We learned a DMP trajectory and force profile from the provided demonstration and executed it on the robot, as visualized in Figure 2a. We compare our method of controlling with respect to a learned dynamic constraint frame described in Section IV-A against controlling desired forces in the z -axis of the fixed world frame, which in our setup is orthogonal to the table surface. Figure 3 compares the resulting pose error and force profiles. The force profiles for both methods are similar and adhere closely to desired forces. However, the L2-norm of the pose error is noticeably higher for controlling with respect to the world frame. This is because unmodeled friction between the end-effector and the table drags down the sliding motion. Our learned constraint frame is less influenced by this effect since it is aligned to the forces observed in demonstration, including compensation forces due to friction. See Figure 1 for a diagrammatic visualization of the differences observed in controlling with respect to learned and fixed constraint frames.

2) *High-friction surface:* We performed the same sliding experiment on a wooden board covered in 150 grit sand paper. As seen in Figure 2a, the z -axis of the learned constraint frame points primarily into the table where desired forces dominate, but it also points slightly in the direction of motion. This is due to the learned constraint frame aligning not only to the forces explicitly imposed by the user during demonstration, but also the compensation forces the user was implicitly applying to overcome friction and maintain the desired end-effector orientation while sliding.

The pose error for controlling with respect to the world frame is exacerbated further on the high-friction surface, as seen in Figures 1, 2a, and 3 (and our supplementary video). The pose error for controlling with respect to our learned constraint frame remains low, and is in fact lower than for the smooth table due to the overall lower forces being applied to the sand paper surface. Figure 3 (right side) shows the force profiles observed in each case. Both methods show good tracking in the z dimension. Our method also exhibits good tracking of the compensation forces for friction in the x and y dimensions. We highlight that we achieve this without

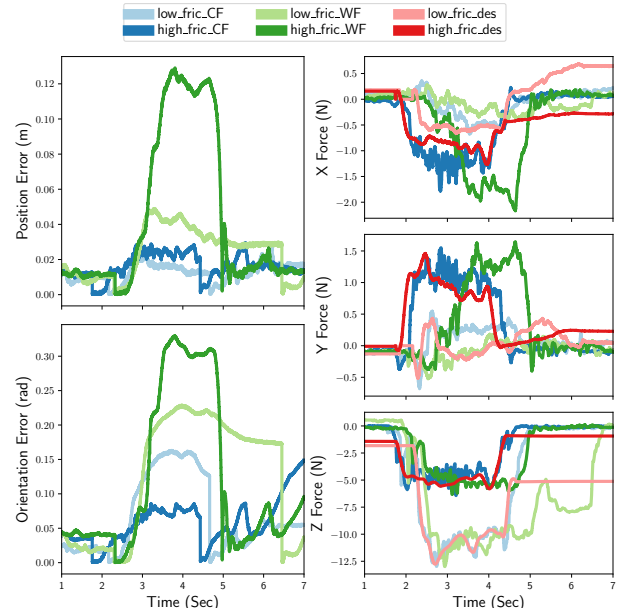


Fig. 3: Results for sliding on low and high friction surfaces comparing control with respect to the world frame (WF) and our learned constraint frame (CF). The left subplots show the L2 norm of the position and orientation error of the end-effector over time. The right subplots show the force profiles in the x , y , and z dimensions.

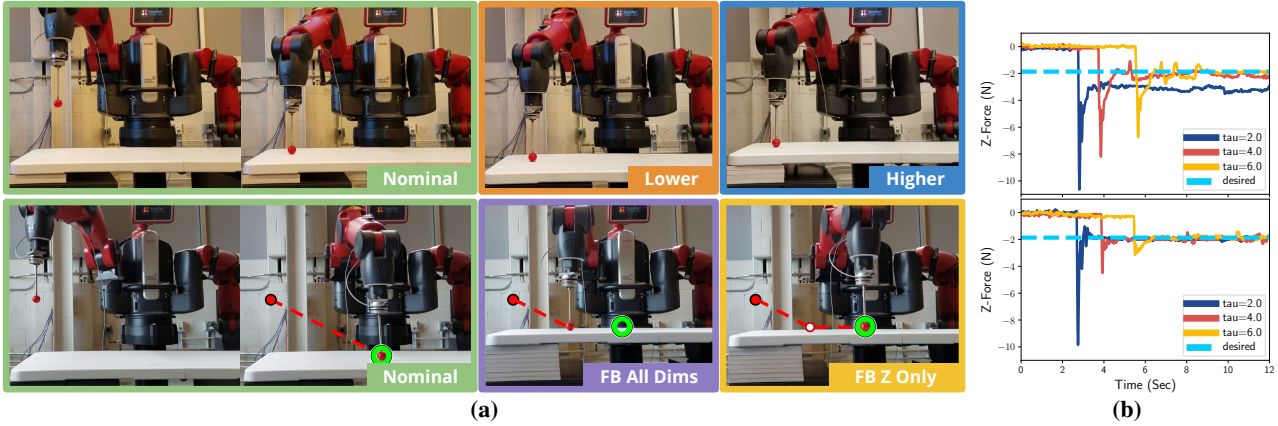


Fig. 4: (a): Scenarios we consider for robustly making contact with a surface. The top row illustrates making contact with a surface when positioned at a different location than observed in demonstration. The bottom row illustrates the case of making a slanted approach to a desired contact point on the surface and the effect of our DMP extensions when contact occurs sooner than anticipated. The dotted red line is the trajectory taken in each case. The green circle highlights the desired goal position with respect to the surface. (b): The results for making contact sooner than expected at different approach velocities. The DMP temporal scaling parameter τ governs trajectory velocity, where lower values correspond to lower velocities. The dotted line shows the desired force. Top: IES force control. Bottom: our method.

modeling friction, and by using only one dimension of the constraint frame for force control. Interestingly, the x and y forces for controlling with respect to the world frame reach a similar magnitude, but at a delayed time. We suggest this is because the frictional forces are being passively reacted to, as opposed to being actively commanded as our method does.

B. Sliding on a curved surface.

A more complex force profile is achieved by sliding the end-effector along the inside of a mixing bowl, as pictured in Figure 2a. We learn a DMP trajectory and force profile and control the execution with respect to the dynamic constraint frame learned from observed forces. We record the forces observed during execution and transform them to the world and tool frames (both commonly chosen constraint frames [3]) for the sake of comparison. The recorded force profiles from one execution are visualized in Figure 2b. The figure illustrates that the desired forces simultaneously vary in a non-trivial manner across all three dimensions of the fixed world and tool frames. This implies that in order to track the forces with these frames, all three dimensions would need to be activated for force control, prohibiting simultaneous position tracking. Our method, on the other hand, requires activation of only one degree of freedom for force control to track the desired forces, thereby ensuring motion orthogonal to the direction of desired force is always possible.

We attempted to compare against controlling with respect to a fixed constraint frame. However, we were unable to perform the experiments safely. Enabling force control for only one dimension, for example x in the world frame, would work as long as motion was primarily orthogonal to that direction. However, as soon as the end-effector started moving along that axis, control became unpredictable and had to be terminated due to fears of damaging the robot and force sensor. Using a fixed frame for this task requires very precise timing of constraint specification. We will in future work seek a reasoned criterion for determining this specification.

C. Making contact with a surface

In our final experiments, we demonstrate the efficacy of our methods for making robust contact presented in Section IV-B

with the task of making contact with a table for which the height may be higher or lower than anticipated.

1) *Straight down approach:* We initialized the robot end-effector to hover 20cm in the world z -axis above a table at a nominal height of 77cm measured from the ground to the table surface. We recorded a demonstration that moved the end effector from its initial position to a desired contact point on the table. The start and end poses of the trajectory can be seen in Figure 4a. We then varied the height of the table to 74cm and 80cm. These heights were chosen to be large enough to clearly illustrate the benefits of our methods while still allowing for open-loop position trajectories to be executed for reference without applying unsafe forces.

For the lower height of 74cm, open-loop position control leaves the end-effector hovering approximately 3cm above the desired contact point. We use our DMP extension described in Section IV-B.2 to slowly change the goal in the direction of the desired contact. We chose a value of $\epsilon = 0.0005$ to move the goal, as this value generates a slow enough speed to easily make stable contact. Once the contact classifier detects contact, the robot enables force control and tracks the desired initial sliding-phase force of approximately 2N.

For making contact at the higher height of 80cm, we compare our method of DMP force feedback with incremental force control against PI force control with and without Integral Error Scaling (IES) described in Section III-A. We test three different execution speeds by varying the DMP temporal scaling parameter τ which approximately corresponds to trajectory duration in seconds. We chose values for $\tau \in [2, 4, 6]$. For each method we use the same control gains $\mathbf{K}_f = \text{diag}(0.2)$, $\mathbf{K}_I = \text{diag}(70.0)$ which were empirically found to exhibit good steady-state tracking once already in contact.

We found that PI force control alone could not make stable contact at any speed using these control gains; control immediately went unstable and had to be terminated for safety. By introducing IES with a value of $\beta = 0.001$, stable contact was retained at each speed. However, as seen in the top of Figure 4b, there is steady-state tracking error of approximately 1.5N for the case of $\tau = 2.0$. The results for our method are

shown in the bottom of Figure 4b. We achieve stable contact, steady-state tracking, and reduce impact forces in all cases.

2) *Angled approach*: Results for this case are pictured in Figure 4a. The end-effector was initialized to be approximately 25cm above a table height of 74cm. A demonstration was recorded that moved the end-effector at an angled approach to the table along a straight line trajectory to a desired contact point on the table. We compare two different behaviors possible with our DMP feedback term defined in Equation (14) on a table height of 86cm to make the difference apparent. When Equation (14) is activated for all task space dimensions (equivalent to applying the change on the canonical system as previously proposed) using $\alpha_f = 10.0$, the end-effector halts as soon as contact is detected and moves no further. The end-effector reached the goal in the z -direction³, but cannot reach the (x, y) goal even though those directions are unconstrained. To achieve full goal convergence, we activate Equation (14) only for z , the dimension in which contact is expected. In this case the end-effector makes contact, halts in the z -direction but continues to converge to the position goal in other directions.

VII. DISCUSSION AND FUTURE WORK

We presented a novel solution to learning hybrid force/position control from demonstration. Our experimental results demonstrate that using a **dynamic constraint frame** aligned to the direction of desired force allows three-dimensional forces to be controlled accurately using only one degree of freedom in the constraint frame. We additionally found that controlling with respect to our learned constraint frame compensates for frictional forces without any explicit modeling of friction, thereby reducing pose deviation over controlling with respect to a fixed frame. An interesting avenue for future work is to learn to adapt to surfaces with higher or lower friction than was observed in demonstration. Reinforcement learning may be one promising approach to achieve this sort of generalization [13].

Our novel extensions to the DMP framework were shown to provide robust transition from free-space motion to surface-constrained motion in spite of environment uncertainty. Our method affords reduced impact forces and better steady-state tracking on higher velocity impacts than other comparable methods. As indicated in Section IV-B.2, we assume an early contact is the intended contact, as opposed to an undesired collision. We make this assumption since the robot only uses a wrist force/torque sensor to classify contacts. In most cases the robot **could avoid observed obstacles using collision avoidance techniques for DMPs** [16]. When unintended contact cannot be avoided, other perceptual modalities such as visual and tactile feedback can allow for more robust classification of intended and unintended contacts. We leave multi-sensory, robust contact classification as a direction for future work.

REFERENCES

- [1] M. H. Raibert and J. J. Craig, "Hybrid position/force control of manipulators," *Journal of Dynamic Systems, Measurement, and Control*, vol. 102, no. 127, pp. 126–133, 1981.
- [2] V. Ortenzi, R. Stolkin, J. Kuo, and M. Mistry, "Hybrid motion/force control: a review," *Advanced Robotics*, vol. 31, pp. 1102–1113, 2017.
- [3] B. Siciliano, L. Sciacivco, L. Villani, and G. Oriolo, *Robotics: modelling, planning and control*. Springer, 2009.
- [4] N. Mandal and S. Payandeh, "Experimental evaluation of the importance of compliance for robotic impact control," in *Control Applications, Second IEEE Conference on*, 1993, pp. 511–516.
- [5] L. S. Wilfinger, "A comparison of force control algorithms for robots in contact with flexible environments," 1992.
- [6] L. Peternel, L. Roza, D. Caldwell, and A. Ajoudani, "A Method for Derivation of Robot Task-Frame Control Authority from Repeated Sensory Observations," *RAL*, vol. 2, pp. 719–726, 2017.
- [7] L. Roza, S. Calinon, and D. G. Caldwell, "Learning force and position constraints in human-robot cooperative transportation," in *RO-MAN*, 2014, pp. 619–624.
- [8] F. Steinmetz, A. Montebelli, and V. Kyrki, "Simultaneous kinesthetic teaching of positional and force requirements for sequential in-contact tasks," in *Humanoids*, 2015, pp. 202–209.
- [9] A. L. P. Ureche, K. Umezawa, Y. Nakamura, and A. Billard, "Task parameterization using continuous constraints extracted from human demonstrations," *Transactions on Robotics*, vol. 31, 2015.
- [10] L. Armesto, J. Moura, V. Ivan, M. S. Erden, A. Sala, and S. Vijayakumar, "Constraint-aware learning of policies by demonstration," *IJRR*, vol. 37, no. 13-14, pp. 1673–1689, 2018.
- [11] Z. Deng, J. Mi, Z. Chen, L. Einig, C. Zou, and J. Zhang, "Learning human compliant behavior from demonstration for force-based robot manipulation," in *ROBIO*, 2016, pp. 319–324.
- [12] M. Suomalainen and V. Kyrki, "Learning compliant assembly motions from demonstration," in *IROS*, 2016, pp. 871–876.
- [13] M. Hazara and V. Kyrki, "Reinforcement learning for improving imitated in-contact skills," in *Humanoids*, 2016, pp. 194–201.
- [14] M. Racca, J. Pajarinen, A. Montebelli, and V. Kyrki, "Learning in-contact control strategies from demonstration," in *IROS*, 2016.
- [15] H.-C. Lin, P. Ray, and M. Howard, "Learning task constraints in operational space formulation," in *ICRA*, 2017, pp. 309–315.
- [16] D.-H. Park, H. Hoffmann, P. Pastor, and S. Schaal, "Movement reproduction and obstacle avoidance with dynamic movement primitives and potential fields," in *Humanoids*, 2008, pp. 91–98.
- [17] P. Pastor, H. Hoffmann, T. Asfour, and S. Schaal, "Learning and generalization of motor skills by learning from demonstration," in *ICRA*, 2009, pp. 763–768.
- [18] A. J. Ijspeert, J. Nakanishi, H. Hoffmann, P. Pastor, and S. Schaal, "Dynamical movement primitives: learning attractor models for motor behaviors," *Neural Computation*, vol. 25, no. 2, pp. 328–373, 2013.
- [19] P. Kormushev, S. Calinon, and D. G. Caldwell, "Imitation learning of positional and force skills demonstrated via kinesthetic teaching and haptic input," *Advanced Robotics*, vol. 25, no. 5, pp. 581–603, 2011.
- [20] F. J. Abu-Dakka, B. Nemec, J. A. Jørgensen, T. R. Savarimuthu, N. Krüger, and A. Ude, "Adaptation of manipulation skills in physical contact with the environment to reference force profiles," *Autonomous Robots*, vol. 39, no. 2, pp. 199–217, 2015.
- [21] M. Deniša, T. Petrić, A. Gams, and A. Ude, "A Review of Compliant Movement Primitives," in *Robot Control*, 2016.
- [22] E. Shahriari, A. Kramberger, A. Gams, A. Ude, and S. Haddadin, "Adapting to contacts: Energy tanks and task energy for passivity-based dynamic movement primitives," in *Humanoids*, 2017, pp. 136–142.
- [23] M. Kalakrishnan, L. Righetti, P. Pastor, and S. Schaal, "Learning force control policies for compliant manipulation," in *IROS*, 2011.
- [24] J. Kober, M. Gienger, and J. J. Steil, "Learning movement primitives for force interaction tasks," in *ICRA*, 2015, pp. 3192–3199.
- [25] D. Kappler, P. Pastor, M. Kalakrishnan, M. Wuthrich, and S. Schaal, "Data-Driven Online Decision Making for Autonomous Manipulation," in *RSS*, 2015.
- [26] O. Kroemer, C. Daniel, G. Neumann, H. Van Hoof, and J. Peters, "Towards learning hierarchical skills for multi-phase manipulation tasks," in *ICRA*, 2015, pp. 1503–1510.
- [27] P. Pastor, M. Kalakrishnan, L. Righetti, and S. Schaal, "Towards associative skill memories," in *Humanoids*, 2012, pp. 309–315.
- [28] D. Driess, P. Englert, and M. Toussaint, "Constrained bayesian optimization of combined interaction force/task space controllers for manipulations," in *ICRA*, 2017, pp. 902–907.
- [29] O. Khatib, "A unified approach for motion and force control of robot manipulators: The operational space formulation," *IEEE Journal on Robotics and Automation*, vol. 3, no. 1, pp. 43–53, 1987.
- [30] P. Pastor, L. Righetti, M. Kalakrishnan, and S. Schaal, "Online movement adaptation based on previous sensor experiences," in *IROS*, 2011, pp. 365–371.
- [31] A. Ude, B. Nemec, T. Petrić, and J. Morimoto, "Orientation in cartesian space dynamic movement primitives," in *ICRA*, 2014, pp. 2997–3004.

³Based on our change in goal technique, the goal converges to the current pose in the z -direction when stable contact has been retained sufficiently long.