Real2Sim2Real: Self-Supervised Learning of Physical Single-Step Dynamic Actions for Planar Robot Casting

Vincent Lim1,*, Huang Huang1,*, Lawrence Yunliang Chen1, Jonathan Wang1, Jeffrey Ichnowski1, Daniel Seita1, Michael Laskey2, Ken Goldberg1

Abstract—This paper introduces the task of Planar Robot Casting (PRC): where one planar motion of a robot arm holding one end of a cable causes the other end to slide across the plane toward a desired target. PRC allows the cable to reach points beyond the robot workspace and has applications for cable management in homes, warehouses, and factories. To efficiently learn a PRC policy for a given cable, we propose Real2Sim2Real, a self-supervised framework that automatically collects physical trajectory examples to tune parameters of a dynamics simulator using Differential Evolution, generates many simulated examples, and then learns a policy using a weighted combination of simulated and physical data. We evaluate Real2Sim2Real with three simulators, Isaac Gym-segmented, Isaac Gym-hybrid, and PyBullet, two function approximators, Gaussian Processes and Neural Networks (NNs), and three cables with differing stiffness, torsion, and friction. Results with 240 physical trials suggest that the PRC policies can attain median error distance (as % of cable length) ranging from 8% to 14%, outperforming baselines and policies trained on only real or only simulated examples. Code, data, and videos are available at https://tinyurl.com/robotcast.

I. INTRODUCTION

Manipulation of deformable objects using a single parameterized dynamic action can be useful for tasks such as fly fishing, lofting a blanket, and playing shuffleboard. Such tasks take as input a desired final state and output one parameterized open-loop dynamic robot action which produces a trajectory toward the final state. This is especially challenging for long-horizon trajectories with complex dynamics involving friction. Although there is substantial research on learning ballistic throwing or hitting motions for rigid objects [30, 67], there is less research into learning dynamic motions to manipulate deformable objects such as cables and fabrics. Dynamics modeling in these contexts is challenging due to uncertainty in deformability, elasticity, and friction during the object’s motion. When computing a single action in this setting (without feedback control), the complexities of state estimation and dynamics modeling are compounded by the long duration for which the system evolves after the robot action. Simulation is often used as an alternative to collecting physical examples, which can be time-consuming and hazardous. However, overcoming the simulation to reality (Sim2Real) gap is a long-standing problem in robotics [20, 32, 41, 54], and is particularly challenging for deformable objects in uncertain dynamic environments.

*Equal contribution.
1The AUTOLab at UC Berkeley (automation.berkeley.edu).
2Toyota Research Institute, CA, USA.
Correspondence to: {vincentlim, huangr}@berkeley.edu

Fig. 1: In Planar Robot Casting (PRC), a single dynamic planar motion of a robot wrist holding one end of a cable causes the other end of the cable to slide across the plane and stop near a desired target point, which may lie outside the robot workspace. (A) shows a side view of a UR5 robot with cable and planar workspace, (B) illustrates test performance on Cable 1 for 5 trials (dots) on 16 targets (stars). The gold inner sector represents the robot workspace, while the grey outer sector represents the reachable workspace of the cable. (C,D) each show five superimposed overhead views of the robot and cable with associated target points after PRC actions with the learned policy, in (C) an example with low error and in (D) an example with high error in endpoint position.

In this paper, we introduce Planar Robot Casting (PRC), where a robot arm dynamically manipulates a cable through 2D space, as illustrated in Fig. 1. We use the generic term cable to refer to any 1D deformable object with low stiffness, such as cables, ropes, and threads. We propose Real2Sim2Real, a self-supervised robot learning framework that starts by efficiently collecting a small number of physical examples, uses them to tune a simulator, and then uses a combination of physical and simulated examples to train policies for PRC. This paper makes four contributions:

1) Real2Sim2Real, a three-step robot learning framework for single-step dynamic actions executed in real environments.
2) A formulation of the Planar Robot Casting (PRC) problem, with an automated reset motion and parameterized one-step action space that can reach outside of the reachable area of the robot.
3) A physical implementation of PRC using a UR5 robot, and three simulation models of PRC: NVIDIA Isaac
Gym segmented, Isaac Gym hybrid, and PyBullet. Parameter estimation using Differential Evolution [53, 55] with associated code and datasets from 64,350 simulated experiments and 2,076 physical experiments with 3 distinct cables.

4) An application of Real2Sim2Real to PRC, with results suggesting that Real2Sim2Real can efficiently learn PRC control policies that reach target positions within median error of 14% of cable length, even when targets are outside the reachable workspace of the robot.

II. RELATED WORK

A. Deformable 1D Object Manipulation

Manipulation of 1D deformable objects such as cables has a long history in robotics; see Sanchez et al. [48] for a recent survey. Some representative applications include surgical suturing [36, 50], knot-tying [6, 17, 37, 58], untangling ropes [16, 29], deforming wires [40], inserting cables into receptacles [52, 59], and playing diabolo [12].

There is recent interest in using learning-based techniques for manipulating cables, often with pick-and-place actions and quasistatic dynamics, so that the robot deforms the cable while allowing it to settle between actions. This can be combined with learning from demonstrations [49, 66] and self-supervised learning [39, 42]. Prior work has also used quasistatic simulators [25] to train cable manipulation policies using reinforcement learning (RL) for transfer to physical robots [60, 65]. However, as we show in Sec. V, there are limits in the free-end reachability of such quasistatic procedures, motivating high-speed, dynamic motions.

B. Simulator Tuning for Sim2Real

Sim2Real transfer has emerged as an effective technique wherein a robot can learn a policy in simulation and transfer it to the real world, which can suffer from the mismatch between simulation and real [20]. One option for Sim2Real is to perform system identification [27] to tune a simulator to match real-world data, but this may be challenging with the infinite-dimensional configuration spaces of cables. An alternative is to use domain randomization over vision [47, 54] or dynamics [43] parameters in simulation. A drawback is that too much randomization can hinder policy training [35] and the randomization range needs to be hand-tuned. This has motivated work on interleaving policy training using RL in simulation and tuning simulators with real-world data [9, 13]. Recently, Chang and Padir [8] proposed a Sim2Real2Sim pipeline for cable unplugging where the Sim2Real step determines grasp points, and the Real2Sim step uses sensors to update a cable model in simulation. Similarly, we tune a simulator, but do not require repeated iterations of collecting real-world data with simulator tuning, and we train policies using supervised learning and do not need to use potentially brittle RL algorithms [1, 18]. Unlike Chang and Padir, we do not require cable sensors and we use the tuned simulator to improve the performance of the physical robot.

C. Dynamic Manipulation

In dynamic manipulation, a robot executes rapid actions to move objects to desired configurations [34]. In early work, Lynch and Mason [30] study planar dynamic manipulation primitives such as snatching, throwing, and rolling, and Ruiz-Ugalde et al. [46] build a physics-based model of pushing actions so that a robot can slide objects into a desired pose. More recent work has introduced robots that can catch items by predicting item trajectories [21], toss arbitrary objects via self-supervised learning [67], and swing items upwards using tactile feedback [56]. As with these works, we use dynamic motions, but focus specifically on cable manipulation. Furthermore, we use a tuned simulator to accelerate learning and do not require tactile sensing.

D. Dynamic Manipulation of Deformable Objects

Researchers have developed several analytic physics models to describe the dynamics of moving cables. For example, Gatti-Bonoa et al. [15] present a 2D dynamic model of fly fishing. They model the fly line as a long elastica and the fly rod as a flexible Euler-Bernoulli beam, and propose a system of differential equations to predict the movement of the fly line in space and time. In contrast to continuum models, Wang et al. [57] propose using a finite-element model to represent the fly line by a series of rigid cylinders that are connected by massless hinges. In contrast to these works, we evaluate on physical robots.

For robotic dynamic-cable manipulation, Yamakawa et al. [61–64] simplify the modeling of cable deformation by assuming each cable component follows the robot end-effector motion with constant time delay. Kim et al. [22] study a mobile robot system with a cable attached as a tail that can strike objects. They use a Rapidly-exploring Random Tree (RRT) [24] and a particle-based representation to address the uncertainty in state transition. In contrast, we aim to control cables for tasks in which we may be unable to rely on assumptions in [22, 62] for cable motion.

In closely related prior work, Zhang et al. [68] propose a self-supervised learning technique for dynamic manipulation of fixed-end cables. In contrast, we use free-end cables. Zimmermann et al. [69] study the dynamic manipulation of deformable beams and cloth in the free-end setting. They model elastic objects using the finite-element method [4] and use optimal control techniques for trajectory optimization. The simulated motions perform well in real for simple dynamical systems but performance deteriorates for complex soft bodies due to the reality gap. We develop a learned, data-driven approach for robotic manipulation of free-end cables and focus on PRC. Laezza et al. [23] develop six customizable benchmark simulation environments for 1D deformable objects, but use an commercial closed-source physics engine. We use easily-accessible physics engines to create dynamic simulation environments.

III. PROBLEM STATEMENT

In Planar Robot Casting, a robot gripper holds a cable at one endpoint and swings it along a planar surface with a
single continuous motion so that the other endpoint comes to rest at a target $s_d$. The held cable endpoint is at polar coordinate $s = (r, \theta)$, with the robot base at the origin. The objective is to find a per-cable policy $\pi$ that minimizes the expected error $E[|s_f,c - s_d,c|^2]$, where $s_f,c$ is the final state and $s_d,c$ is $s_d$ in Cartesian coordinates. We assume that the $s_d$ is reachable by the cable endpoint (see Fig. 1).

### IV. METHOD

To learn a policy that accounts for variations in cable properties such as mass, stiffness, and friction, we propose the Real2Sim2Real (R2S2R) framework (Fig. 2) and apply it to PRC. To support this framework, we define a reset procedure to bring the system into a consistent starting state and a parameterized trajectory function that generates a dynamic arm action. R2S2R first autonomously collects physical trajectories, then tunes a simulator to match the physical environment and generate simulated trajectories. Finally, R2S2R uses a weighted combination of the simulated and physical datasets to train a policy.

#### A. Reset Procedure

To automate data collection and bring the cable into a consistent starting state before each action, we define a 5-step reset procedure, in which the robot (1) lifts the cable up with the free-end touching the surface to prevent the cable from dangling; (2) continues to lift the cable such that the free-end is just above the surface; (3) hangs still for 3 seconds to stabilize; (4) swings the cable out of the plane, so that the cable straightens along the center axis of the surface and lands with its endpoint far from the robot base; and (5) slowly pulls the cable towards the robot to the reset position $(r_0,0)$. The reset motion is designed to minimize empirical uncertainty in the reset state, so we assume the start state to be consistent across trajectories. The project website illustrates the reset procedure.

#### B. Parameterized Actions

To generate quick, smooth, dynamic actions that are low dimensional, thus facilitating data generation and training, we define a parameterized action $a := (\theta_1, r_1, \theta_2, r_2, \alpha, v_{max})$ composed of two sweeping arcs (Fig. 3). The motion starts at $(r_0,0)$, arcs to $(r_1, \theta_1)$, and arcs back to $(r_2, \theta_2)$. In the motion, $\alpha$ is the wrist joint rotation about the $z$-axis during the second arc, and $v_{max}$ is the maximum velocity. This parameterization is motivated by observing human attempts at the PRC task.

We convert action parameters to a trajectory in polar coordinates using a cubic spline to smoothly interpolate the radial coefficient from $r_0$ to $r_2$, and use a maximum-velocity spline to interpolate the angular coefficient from 0 to $\theta_1$ and from $\theta_1$ to $\theta_2$, with maximum velocity $v_{max}$. The UR5 has difficulty following high-jerk trajectories, so the maximum-velocity spline uses jerk-limited bang-bang control [19]. We assume a direction change between the two arc motions, so the angular velocity at $\theta_1$ is 0. We use an analytic inverse kinematics (IK) solver to convert the splines to joint configurations. The execution for each trajectory is about 22 seconds, including 20 seconds for the reset motion and 2 seconds for the planar motion.

This parameterization enables state symmetry. For all datasets, we sample actions such that $\theta_1 > 0, \theta_2 < 0$, and $\alpha \geq 0$, to obtain targets on the left of the workspace axis of symmetry. If $(\theta_1, \theta_2, r_2, \alpha, v_{max})$ produces target $(r_d, \theta_d)$, $(-\theta_1, -\theta_2, r_2, -\alpha, v_{max})$ will produce $(r_d, -\theta_d)$. Thus, we do not evaluate targets on the right of the axis of symmetry.

#### C. Self-Supervised Physical Data Collection

For the first step of R2S2R, we autonomously collect a physical dataset $D_{phys}$ and take a subset of $D_{phys}$ to form a simulator tuning dataset $D_{tune}$ to perform system identification in simulation (Sec. IV-E). To create $D_{phys}$, we grid sample each action parameter to generate $5 \times 5 \times 5 \times 4 \times 2$
where \( t \) is the timestep, every 100 ms from the start of the robot’s trajectory. We extract the 2D waypoint location \( p_t = (x_t, y_t) \), where \( t \) is the timestep, every 100 ms from the start of the robot’s trajectory. For each trajectory \( m_j \) in \( D_{\text{phys}} \), the number of waypoints collected is \( K_j = \lceil T_j/100 \rceil \), where \( T_j \) is the duration of \( m_j \) in milliseconds.

D. Three Simulation Models for Robot Casting

R2S2R then tunes a simulator to generate data to augment policy training. We consider which simulator and model best match real from 3 options: PyBullet [11] and two versions of NVIDIA Isaac Gym [33]. We set simulated cable geometry (e.g., length and radius) to match the real cable. Fig. 4 shows a visual comparison, and the website provides further details.

PyBullet is a CPU-based deterministic rigid-body physics simulator used in prior work on deformable object manipulation [35, 49]. We model the cable as a string of capsule-shaped rigid bodies with 6-DOF spring constraints between each consecutive pair. We tune ten parameters: twist stiffness, bend stiffness, mass, lateral friction, spinning friction, rolling friction, endpoint mass, linear damping, angular damping, and dynamic friction.

NVIDIA Isaac Gym is a GPU-based robotics simulation platform that supports the FleX [31] particle-based physics simulator designed for deformable objects and rigid bodies. We test two simulation models for the cable: a segmented model and a hybrid model, both using Flex. The segmented model is a string of 18 capsule-shaped rigid bodies with consecutive pairs linked together by a ball joint. To model cable stiffness, we tune the joint friction, cable mass, endpoint mass, and planar friction to capture variation in each respective parameter in real. The hybrid model has the same rigid endpoint from the segmented model, but the rest of the cable is a soft-body rod. We tune the Young’s modulus of the soft-body rod to model both the cable stiffness and elasticity, dynamic friction of the ground plane, and rigid endpoint mass to capture variations in friction and endpoint properties. Simulation environment details are on the website.

E. Simulator Parameter Tuning for \( D_{\text{sim}} \) collection

To tune the parameters of each simulator so that simulated and real trajectories closely match, we follow system identification [27] and employ a consistent protocol that uses physical trajectories from \( D_{\text{tune}} \) (Sec. IV-C) to minimize the discrepancy between simulated and real trajectories. We define the simulation tuning objective as finding simulation parameters that minimize the average \( L^2 \) distance between the cable endpoint in simulation \( p_t \) to the target endpoint \( p_t \) over all trajectory timesteps, averaged over a batch of trajectories. We call this metric the average waypoint error.

As the simulation models may have parameters that have no physical analog, such as the joint friction in the segmented model, or have parameters that do not act similarly to their physical counterparts [28], we choose tuning algorithms that make no assumptions about the underlying dynamics model. We evaluate two derivative-free black box optimization algorithms, Bayesian Optimization (BO) and Differential Evolution (DE).

BO builds a probabilistic surrogate model and uses an acquisition function that leverages the uncertainty in the posterior to decide where to query next [14], BO has been used in policy search for RL hyperparameters [26, 51], and for tuning simulation fluid dynamics [28]. DE is an evolutionary population-based stochastic optimization algorithm widely used for global optimization of nondifferentiable, multimodal, and nonlinear objectives [53]. DE maintains a population of candidate solutions at each iteration and generates new candidates by combing each population member with another mutated population candidate member. It evaluates the fitness of the trial candidates to update the population of candidate solutions until convergence. DE has been shown to be effective at simulation parameter tuning [10].

We evaluate each algorithm via Sim2Sim tests. Using a fixed set of simulation parameters, we generate a small dataset of trajectories to tune randomly initialized simulators to test the effectiveness of DE and BO. We consider the discrepancy between ground truth and predicted simulation parameters and average waypoint error (see results in Sec. V-A). We then tune the simulator parameters for each simulation model to the physical data using the best performing tuning algorithm. We tune the simulator using the same process but with physical trajectories from \( D_{\text{tune}} \).

F. Size of Simulator Tuning Set \( D_{\text{tune}} \)

We subsample 20 trajectories from \( D_{\text{phys}} \) to generate dataset \( D_{\text{tune}} \). We evaluated whether increasing the size of \( D_{\text{tune}} \) would reduce the discrepancy between simulation and real using a test set of 30 random physical trajectories not in \( D_{\text{phys}} \). We observed negligible difference in test errors between using 20 and 60 trajectories, but the latter require \( 2 \times \) as long to tune on the hybrid and PyBullet models. Tuning with \( |D_{\text{tune}}| = 20 \) trajectories did not lead to a significant speedup in the segmented model, therefore we held this value fixed.

After tuning, R2S2R generates training data \( D_{\text{sim}} \) by grid sampling \( \theta_1, \theta_2, \gamma_2, \psi, \) and \( v_{\text{max}} \) to generate \( 15 \times 15 \times 15 \times 10 \times 2 \) values for each respective parameter. We then filter out any trajectories that violate joint limits or with collisions to generate \( |D_{\text{sim}}| = 21,450 \) (a, s) simulated trajectories.
G. Policies

1) Forward Dynamics Model: Due to the multi-modality between a target endpoint $s$ and valid actions $a$, we do not learn a policy that directly predicts $s$ given $a$. Instead, we learn a forward dynamics model $f_{\text{forw}}$ that predicts $s$ given $a$, and interpolates to select an action. We parameterize $f_{\text{forw}}$ using a fully connected neural network. Given a dataset $D$ of $(a, s_f)$ trajectories, the neural network learns to predict $s_f$ given $a$ via standard supervised regression. During evaluation, the policy $\pi$ grid samples 67,500 input actions $a$ to form a set $A$ of candidate actions. It then passes each action through the forward dynamics model $f_{\text{forw}}$, which outputs the predicted endpoint location $(x, y)$ in Cartesian space. Given a target endpoint $s_d$ (in polar coordinates), the policy selects the action that minimizes the Euclidean distance between the predicted endpoint $s_f$ and target endpoint $s_d$: $\pi(s) = \arg\min_{a \in A} \|s_{f,c} - s_{d,c}\|_2$, where $s_{f,c}$ and $s_{d,c}$ are $s_f$ and $s_d$ in Cartesian coordinates respectively. Training details can be found on the website.

2) Baseline Policies: We consider two alternative policies:

- **Cast and Pull**: Given a target cable endpoint location $s_d = (r_d, \theta_d)$, the robot rotates $\theta_d$ radians and executes a rotated reset motion as in Sec. IV-A, causing the free endpoint to land at $(r, \theta_d)$. The robot then slowly pulls the cable toward the base for a distance of $r - r_d$. The robot arm is limited to a minimum $r$ coordinate of $r_{\min} = 0.55$ to prevent the end effector from hitting the robot’s supporting table, where $r_{\min}$ is the distance from the center of the robot base to the edge of the table, limiting the reachable workspace.

- **Forward Dynamics Model using Gaussian Process**: As a learning baseline, we train a GP regressor [38, 45] using $D_{\text{phys}}$ to predict the cable endpoint location $s_f$ given input action $a$. As with the neural network forward dynamics model, we grid sample 67,500 input actions and select the trajectory that minimizes the predicted Euclidian distance to the target. We do not train a GP regressor using the simulated dataset as GP regression becomes intractable for large datasets with thousands of datapoints [5].

V. EXPERIMENTS

We evaluate the R2S2R pipeline on PRC with 3 cables (Fig. 5) using a UR5 robot (Fig. 1). The working surface is a 2.45 m wide and 1.55 m tall masonite board, painted blue and sanded to create a consistent friction coefficient across the surface. We use an overhead Logitech Brio 4K webcam recording 1920×1080 images at 60 frames per second.

We consider three forward dynamics model policies:

1) $\pi_{\text{RD}}$, trained on a small real dataset $D_{\text{phys}}$.
2) $\pi_{\text{SD}}$, trained on a large simulated dataset $D_{\text{sim}}$.
3) $\pi_{\text{R2S2R}}$, trained on the combined dataset $D_{\text{phys}} \cup D_{\text{sim}}$.

Since $|D_{\text{sim}}| \gg |D_{\text{phys}}|$, we 1) upsample $D_{\text{phys}}$ by randomly duplicating samples such that the real data make between 30% and 40% of the combined dataset to force the model to learn more from the real data, and 2) weight the loss function so that samples from $D_{\text{phys}}$ are weighted higher. We empirically choose the upsampling percentage and real sample weights to be the best performing values.

A. Comparing Real2Sim Tuning Methods

To test the effectiveness of BO and DE for tuning the Isaac Gym simulators, we generate simulated data for several hybrid models and segmented models with arbitrarily chosen parameters. We then apply BO and DE to tune the respective simulator model parameters. For BO, we use the GP-Opt [3] library, and test three acquisition functions: Expected Improvement (EI), Lower Confidence Bound (LCB), and Maximum Probability of Improvement (MPI). We use DE as implemented in SciPy [55]. With 5 tuning trajectories, EI performs better than LCB and MPI. The tuning errors using BO range from 3.39% to 17.98% for the hybrid model and 1.09% to 9.25% for the segmented model. DE consistently tunes the parameters to within 1% of the ground truth parameters for both the hybrid and segmented models, thus we use DE for all further simulator tuning results.

B. Real2Sim

Table I reports the Real2Sim simulator tuning results using DE. We observe that the segmented model consistently outperforms the PyBullet and hybrid models in minimizing the discrepancy between simulation and real. For the hybrid model, we find that the tuning algorithms reduce the value of Young’s modulus to decrease cable stiffness, but also increases the stretchiness along the length of the cable, which we did not observe in real. We attribute the PyBullet model performance to spring-system simulation instability associated with high spring forces, particularly when preventing the cable from stretching along the length of the cable.

For cable 3, we observe that the hybrid and segmented models perform similarly. Since cable 3 is stiffer, DE produces higher Young’s modulus values that both make the simulated cable stiffer and cause it to stretch less along its length, both of which are desired properties.

While the errors for cable 2 are below 10%, we observe that cables 1 and 3 are harder for each simulator to model,

<table>
<thead>
<tr>
<th></th>
<th>Cable 1</th>
<th>Cable 2</th>
<th>Cable 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sim. Model</td>
<td>Cable 1</td>
<td>Cable 2</td>
<td>Cable 3</td>
</tr>
<tr>
<td>PyBullet</td>
<td>29%</td>
<td>28%</td>
<td>28%</td>
</tr>
<tr>
<td>Isaac Gym Hybrid</td>
<td>14%</td>
<td>23%</td>
<td>11%</td>
</tr>
<tr>
<td>Isaac Gym Segm.</td>
<td>9%</td>
<td>13%</td>
<td>8%</td>
</tr>
</tbody>
</table>

TABLE I: Error for simulators tuned using DE on a set of 30 test trajectories for 3 cables (see Fig. 5). The two metrics are the median final $L^2$-distance, which is the 2D Euclidean distance between endpoint locations in simulation and reality after a trajectory terminates, and average waypoint (“Wayp.”) error. Values are expressed in percentages of cable length.
We find that policies \( \pi \) performs the best out of the policies, we proceed to evaluate the performance for 3 cables on 16 target positions using policy \( \pi_{\text{R2S2R}} \), trained for each cable separately, and repeat each action 5 times per target. Results are summarized in Table III. We quantify the aleatoric uncertainty with the 95% confidence ellipses in Fig. 6.

We attribute the epistemic uncertainty to two sources: the Sim2Real gap, where trajectories executed in simulation do not accurately reflect real, and the learning error in the forward dynamics model. The results suggest that the R2S2R pipeline can apply to other cables.

VI. CONCLUSION AND FUTURE WORK

In this paper, we present Real2Sim2Real, a self-supervised learning framework, and apply it to Planar Robot Casting. Experiments suggest that the framework, which collects physical data, tunes a simulator to match the real data, and trains policies from a weighted combination of real and simulated data, can achieve an median error between 12% and 15% of cable length on the PRC task.

As motion is restricted to a 2D plane, we do not address the 3D analog of PRC, Spatial Robot Casting (SRC), used in tasks such as fly fishing [15]. PRC is easier to visualize than SRC but can be harder to model due to inherent uncertainty about static and dynamic friction. In both PRC and SRC, the cable is infinite-dimensional and there is a long time horizon for the motion of the cable free end resulting from a motion at the controlled end.

In future work, we will explore alternate Real2Sim and Sim2Real techniques [2, 13, 43, 44] and apply Real2Sim2Real to other tasks such as grasping and fabric manipulation.

ACKNOWLEDGMENTS

This research was performed at the AUTOLAB at UC Berkeley in affiliation with the Berkeley AI Research (BAIR) Lab and the CITRIS “People and Robots” (CPAR) Initiative. The authors were supported in part by the Toyota Research Institute. We thank our colleagues Justin Kerr, Alejandro Escontrela, Ashwin Balakrishna, Ellen Novoseller, Brijen Thananjeyan, Harry Zhang, Zachary Tam, Chung Min Kim, and Ananth Rao for helpful comments.
REFERENCES


