Active Perception for Deformable Linear Objects Stiffness Estimation

Chengxiao Dong, Alessio Caporali, Hongyu Lan and Gianluca Palli

***Abstract*— Estimating the stiffness of Deformable Linear Objects (DLOs) is crucial for robust manipulation. Inferring this hidden property depends heavily on the physical interaction strategy. Through a 1D CNN-based analysis of predefined probing modes, we first demonstrate that boundary constraints and grasp locations drastically alter stiffness identifiability. While fixed-end setups yield highly informative responses, they are rarely practical in unconstrained tasks. Consequently, we move beyond manual heuristics and reframe DLO parameter identification as an active perception problem. We propose a Reinforcement Learning (RL) framework that autonomously learns informative interaction strategies for free cables. By coupling a Proximal Policy Optimization (PPO) agent with a trajectory-aware estimator, the system dynamically excites the DLO to extract stiffness from diverse, stochastic manipulation sequences. Achieving a Mean Absolute Error (MAE) of 0.0192, our approach provides a robust, active paradigm that overcomes the limitations of static probing in unconstrained environments.**

***Index Terms*— Deformable Linear Objects, Active Perception, Stiffness Estimation, Reinforcement Learning.**

I. INTRODUCTION

Manipulating Deformable Linear Objects (DLOs) poses distinct challenges [1], [2] as their infinite degrees of freedom make shape control inherently underactuated [3]. Predictable task execution thus relies on understanding their intrinsic physical parameters. Properties like bending stiffness govern shape evolution and contact mechanics [4], yet remain hidden state variables requiring dynamic excitation to be measured [5].

The limiting factor in parameter estimation is often generating informative interaction data, rather than solely model capacity. Traditional methods rely on environmental anchors or static fixtures [4], [5] to produce predictable force-deformation mappings [6]. While recent work explores online estimation during dynamic manipulation to relax these assumptions [4], [5], [7], manually designed heuristics in unconstrained tabletop tasks [8] often reduce to passive data collection. Sub-optimal probing motions fail to sufficiently excite the DLO, yielding data that lacks the informative deformations needed to isolate stiffness.

Framing parameter identification as an active perception problem offers a robust alternative to static heuristics [9], [10]. After quantifying the information disparity across predefined boundary conditions, we introduce an active perception framework for unconstrained cables. By coupling a reinforcement learning agent [11], [12] with a trajectory-aware neural network [13], the robot learns to adaptively perturb unanchored cables to elicit informative feedback [14]. Instead of relying on fixed routines, the system autonomously discovers interaction strategies that maximize parameter identifiability [15].

The authors are with DEI - Department of Electrical, Electronic and Information Engineering, University of Bologna, 40136 Bologna, Italy.

Corresponding author: chengxiao.dong2@unibo.it

This work was supported by the China Scholarship Council (CSC).

The authors acknowledge Algoryx Simulation AB for providing an academic license of AGX Dynamics.

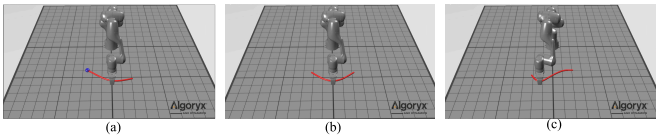

Fig. 1. Variations of DLO grasping location and boundary conditions. (a) `Mode 1`: Pinned-free ends, midpoint grasp. (b) `Mode 2`: Free-free ends, symmetric grasp. (c) `Mode 3`: Free-free ends, asymmetric grasp.

II. METHODOLOGY

A. Task Formulation and Simulation Environment

We simulate a Franka Panda robot manipulating a cable in AGX Dynamics [16]. The cable (radius 0.0075 m, discretized into 48 segments) has a density of 1000 kg/m^3, a twisting stiffness of 1.0×10^6, and a stretching stiffness of 1.0×10^8. Friction coefficients for the cable-table and cable-finger contacts are 0.3 and 0.5, respectively.

Our target property, bending stiffness E_{bend}, is sampled from $[10^4, 10^6]$. To stabilize network training, we logarithmically map E_{bend} to a normalized scalar $\phi \in [0,1]$:

$$\phi = \frac{\log_{10} E_{\text{bend}} - \log_{10} E_{\min}}{\log_{10} E_{\max} - \log_{10} E_{\min}} \tag{1}$$

where $E_{\min} = 10^4$ and $E_{\max} = 10^6$.

We evaluate three predefined setups (Fig. 1): `Mode 1` (midpoint grasp, one pinned end), `Mode 2` (midpoint grasp, free ends), and `Mode 3` (quarter-point grasp, free ends). In each mode, the robot executes a 6 s sequence (grasping and a 0.15 m horizontal drag) followed by 2 s of static recording. Sampled at 10 Hz (inclusive of the $t = 0$ initial state), this 8 s interaction produces an 81-step temporal sequence. By tracking the 3D relative displacements of 25 cable nodes, each episode yields a 75-dimensional trajectory. Specifically, this feature vector concatenates the $\Delta x, \Delta y$, and Δz positional offsets of all 25 nodes relative to their initial resting state, strictly isolating the cable's pure structural deformation from global spatial translation. We generated 10,000 episodes per mode for offline analysis.

B. Supervised Stiffness Estimator

We train a supervised estimator to evaluate the informativeness of each predefined mode, randomly splitting the dataset into 80/10/10% for training, validation, and testing.

We process the multivariate deformation time series using a 3-layer 1D CNN that progressively extracts temporal features across the 75 spatial dimensions. The network channels expand as $75 \to 64 \to 128 \to 256$ (using kernel sizes 7, 5, and 3, respectively), with each layer followed by BatchNorm1d and ReLU. After aggregating the feature maps via temporal adaptive average pooling, an MLP head ($256 \to 64 \to 1$) with a sigmoid activation regresses the normalized stiffness ϕ.

The model is optimized using Adam for up to 400 epochs, incorporating early stopping (patience: 40) on the validation loss to prevent overfitting.

C. Active Interaction Learning Framework

We formulate unconstrained cable interaction as an active perception RL problem. At 10 Hz, each 81-step episode (including the $t = 0$ reference) produces a 97D observation. This comprises a 22D proprioceptive state (14D kinematics, 3D end-effector and midpoint coordinates, 1D gripper state, and a 1D temporal progress ratio) and 75D cable deformation features (3D relative displacements of 25 nodes). The 4D continuous action defines the gripper state and Cartesian displacement $(\Delta x, \Delta y, \Delta z)$.

To handle these heterogeneous inputs, a dual-stream Proximal Policy Optimization (PPO) [17] agent employs an MLP to encode the 22D state and a 1D CNN for the 75D spatial deformation. These representations are fused to parameterize a Gaussian policy. Concurrently, the stiffness estimator processes the full 101D trajectory (97D observation + 4D action) across all 81 steps, using temporal convolutions and Multi-Head Attention to capture delayed action-response dependencies.

Training proceeds in three phases: (1) Behavior Cloning (50 episodes) from Mode 2 to bootstrap exploration; (2) estimator pretraining (200 episodes); and (3) joint PPO training using a 20,000-capacity replay buffer. To fairly evaluate autonomous optimization, empirical statistics (e.g., Fig. 3) strictly reflect Phase 3. Here, PPO updates every 32 episodes (4 epochs, batch size 64) with policy LR 3×10^{-4}, value LR 1×10^{-3}, $\gamma = 0.99$, and GAE $\lambda = 0.95$.

The episode-level reward $R^{(e)}$ is decomposed into step-level signals for PPO:

$$R^{(e)} = \beta(e) \sum_{t=1}^{T_e} r_{\text{shape},t} + w_{\text{est}} r_{\text{est}}^{(e)} g^{(e)} - r_{\text{idle}}^{(e)} \tag{2}$$

The primary objective, $r_{\text{est}}^{(e)} = \exp(-5|\phi_{\text{pred}}^{(e)} - \phi_{\text{gt}}^{(e)}|)$, drives estimation accuracy. To prevent the agent from exploiting the estimator without physically exciting the cable, a multiplicative gate $g^{(e)} \in [0,1]$ scales with midpoint displacement, zeroing $r_{\text{est}}^{(e)}$ if deformation is negligible. An idle penalty $r_{\text{idle}}^{(e)}$ applies if $g^{(e)} \approx 0$. Dense shaping terms $r_{\text{shape},t}$ (for approaching and grasping) guide early exploration but linearly anneal to zero ($\beta(e) \to 0$) between episodes 2000 and 5000, enforcing pure parameter identification. During buffer insertion, shaping rewards apply per-step, while the gated estimation reward is uniformly distributed across the final 10% of steps to account for the delayed nature of temporal inference.

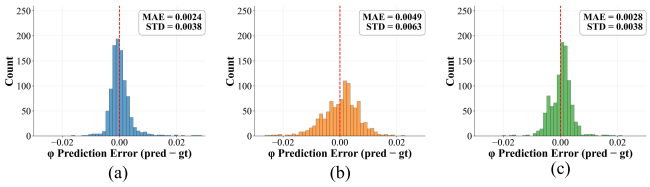

Fig. 2. Comparison of stiffness estimation error distributions. (a) `Mode 1`. (b) `Mode 2`. (c) `Mode 3`. The fixed-end setup (a) yields the lowest estimation error, while the free cable grasp (c) provides significantly more informative deformations than the symmetric midpoint grasp (b).

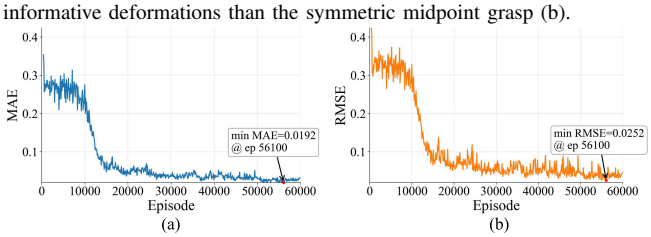

Fig. 3. Training curves for the active perception policy (moving averages over 60,000 episodes). The agent achieves stable convergence for both (a) MAE and (b) RMSE under unconstrained conditions.

III. EXPERIMENTS AND RESULTS

A. Supervised Probing Mode Comparison

Offline results (Fig. 2) demonstrate that interaction strategies dictate prediction quality. The fixed-end setup (`Mode 1`) achieved the lowest Mean Absolute Error (MAE) of 0.0024 for ϕ. The symmetric free-cable setup (`Mode 2`) performed worst (MAE 0.0049). Shifting the grasp to the quarter point (`Mode 3`) broke this mechanical symmetry, improving the MAE to 0.0028. This gap demonstrates how boundary constraints and grasp locations govern data informativeness.

B. Active RL Interaction Performance

To overcome `Mode 2`'s limitations, our RL framework learns adaptive manipulation strategies for unconstrained cables. The policy was trained for 60,000 episodes (Fig. 3), achieving a minimum MAE of 0.0192 and an RMSE of 0.0252, averaged over 100 episodes.

The agent dynamically drags the cable to actively maximize stiffness excitations. Although the RL-driven MAE is numerically higher than the static baseline ($\sim$0.003), this trade-off is physically justified. Fixed heuristics force estimators to overfit a single predictable trajectory. Conversely, the RL policy generates diverse, stochastic interactions. This active exploration introduces estimation variance but critically prevents overfitting, yielding a robust parameter identification paradigm for complex, unconstrained environments.

IV. CONCLUSION

We reframe DLO stiffness estimation from passive regression to active perception. Our initial analysis establishes that boundary constraints and grasp locations significantly influence stiffness identifiability. Consequently, we propose a Reinforcement Learning framework that learns adaptive manipulation strategies, dynamically coupling physical interaction with robust parameter estimation. Future work will explore latent representations for broader physical properties and sim-to-real hardware deployment.

REFERENCES

[1] J. Sanchez, J.-A. Corrales, B.-C. Bouzgarrou, and Y. Mezouar, "Robotic manipulation and sensing of deformable objects in domestic and industrial applications: a survey," *The International Journal of Robotics Research*, vol. 37, no. 7, pp. 688–716, 2018.

[2] J. Zhu, A. Cherubini, C. Dune, D. Navarro-Alarcon, F. Alambeigi, D. Berenson, F. Ficuciello, K. Harada, J. Kober, X. Li, *et al.*, "Challenges and outlook in robotic manipulation of deformable objects," *IEEE Robotics & Automation Magazine*, vol. 29, no. 3, pp. 67–77, 2022.

[3] N. Lv, J. Liu, and Y. Jia, "Dynamic modeling and control of deformable linear objects for single-arm and dual-arm robot manipulations," *IEEE Transactions on Robotics*, vol. 38, no. 4, pp. 2341–2353, 2022.

[4] C. Wang, Y. Zhang, X. Zhang, Z. Wu, X. Zhu, S. Jin, T. Tang, and M. Tomizuka, "Offline-online learning of deformation model for cable manipulation with graph neural networks," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 5544–5551, 2022.

[5] A. Caporali, P. Kicki, K. Galassi, R. Zanella, K. Walas, and G. Palli, "Deformable linear objects manipulation with online model parameters estimation," *IEEE Robotics and Automation Letters*, vol. 9, no. 3, pp. 2598–2605, 2024.

[6] T. Bretl and Z. McCarthy, "Quasi-static manipulation of a kirchhoff elastic rod based on a geometric analysis of equilibrium configurations," *The International Journal of Robotics Research*, vol. 33, no. 1, pp. 48–68, 2014.

[7] N. U. Shinde, X. Liang, F. Liu, Y. Zhang, F. Richter, S. L. Herbert, and M. C. Yip, "Jiggle: An active sensing framework for boundary parameters estimation in deformable surgical environments," in *2024 Robotics: Science and Systems (RSS)*, 2024, pp. 1–20.

[8] K. Chen, Z. Bing, Y. Wu, F. Wu, L. Zhang, S. Haddadin, and A. Knoll, "Real-time contact state estimation in shape control of deformable linear objects under small environmental constraints," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 13 833–13 839.

[9] R. Bajcsy, Y. Aloimonos, and J. K. Tsotsos, "Revisiting active perception," *Autonomous Robots*, vol. 42, no. 2, pp. 177–196, 2018.

[10] K. Zhang, B. Li, K. Hauser, and Y. Li, "Adaptigraph: Material-adaptive graph-based neural dynamics for robotic manipulation," *arXiv: 2407.07889*, 2024.

[11] K. Lv, M. Yu, Y. Pu, X. Jiang, G. Huang, and X. Li, "Learning to estimate 3-d states of deformable linear objects from single-frame occluded point clouds," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 7119–7125.

[12] M. Li, H. Yu, and C. Choi, "Routing manipulation of deformable linear object using reinforcement learning and diffusion policy," in *2025 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2025, pp. 1–7.

[13] J. Lyu, Z. Li, X. Shi, C. Xu, Y. Wang, and H. Wang, "Dywa: Dynamics-adaptive world action model for generalizable non-prehensile manipulation," in *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*. IEEE, 2025, pp. 11 058–11 068.

[14] J. Bohg, K. Hausman, B. Sankaran, O. Brock, D. Kragic, S. Schaal, and G. S. Sukhatme, "Interactive perception: Leveraging action in perception and perception in action," *IEEE Transactions on Robotics*, vol. 33, no. 6, pp. 1273–1291, 2017.

[15] T. Schneider, C. de Farias, R. Calandra, L. Chen, and J. Peters, "Apple: Toward general active perception via reinforcement learning," *arXiv preprint arXiv:2505.06182*, 2025.

[16] Algoryx Simulation, "AGX Dynamics," 2026. [Online]. Available: https://www.algoryx.se/documentation/complete/agx/tags/latest/doc/html/index.html

[17] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.