

Supplementary Materials: Surrogate-aided Learning of Active Tether-Net Maneuver to Capture Rotating Space Debris:

Feng Liu¹, Achira Boonrath¹, Eleonora M. Botta², Souma Chowdhury³

This document serves as supplementary material to the following paper [1].

F. Liu, A. Boonrath, E. M. Botta, S. Chowdhury, “Surrogate-aided Learning of Active Tether-Net Maneuver to Capture Rotating Space Debris”, *Submitted to IEEE Transactions on Aerospace and Electronic Systems*, under review.

I. MODELS FOR THE TETHER-NET DEPLOYMENT & CAPTURE PHASES

A. Surrogate Models to Determine Maneuverable Units (MUs)’ Trajectories and Capture Success

MUs Deployment Phase Surrogate Model: The Long Short-Term Memory (LSTM) neural network is selected for MUs trajectory prediction due; the structure of the LSTM model is shown in the upper half of Table I.

Capture Phase Surrogate Model: The purpose of the capture phase surrogate model is to act as a binary classifier, based on the definitions of success stated in the associated paper [1]; the structure of the Convolutional Neural Network (CNN) used for the model is shown in the lower half of Table I.

TABLE I: Neural Network Architectures

Component	Layer Type	Parameters	Output Shape
LSTM (Deployment phase)			
Input	Input Layer	Size: 18	(ξ , 18)
Dense	Linear	in_feature=18, out_feature=1024	(ξ , 1024)
Activation	Tanh		(ξ , 1024)
Repeat	Repeat Vector	Times: 100	(ξ , 100, 1024)
LSTM	LSTM	input_size=1024, hidden_size=1024, num_layers=1	(ξ , 100, 1024)
Dense	Linear	in_features=1024, out_features=12	(ξ , 100, 12)
CNN (Capture phase)			
Input	Input Layer	Size: 33	(ξ , 33)
Convolution	Conv1D	in_channels=1, out_channels=8, kernel_size=3	(ξ , 8, 31)
Activation	ReLU		(ξ , 8, 31)
Pooling	MaxPool1D	kernel_size=2	(ξ , 8, 15)
Flatten	Flatten		(ξ , 120)
Dense	Linear	in_features=120, out_features=512	(ξ , 512)
Activation	ReLU		(ξ , 512)
Regularization	Dropout	p=0.5	(ξ , 512)
Dense	Linear	in_features=512, out_features=1	(ξ , 1)
Activation	Sigmoid		(ξ , 1)

* ξ is the batch size.

The training (loss function) histories for both surrogate models are shown in Fig. 1, with notable reductions observed over time and convergence of the validation loss by the end of the training processes.

II. RESULTS: TESTING PERFORMANCE OF THE LEARNED MANEUVER POLICIES

A. Surrogate vs. High-Fidelity Trajectories

Figure 2 provides a comparison of the predicted (surrogate-based) and ground-truth (higher-fidelity sim-based) trajectories of the 4 MUs for an example scenario, demonstrating position errors to be within 0.5 m.

¹ Ph.D. Student, Mechanical & Aerospace Engineering, IEEE Student Member

² Assistant Professor, Mechanical & Aerospace Engineering, IEEE Member

³ Associate Professor, Mechanical & Aerospace Engineering, IEEE Senior Member, Corresponding author

All authors are with University at Buffalo, Buffalo, NY 14260, USA

{fliu23, achirabo, ebotta, soumacho} @buffalo.edu

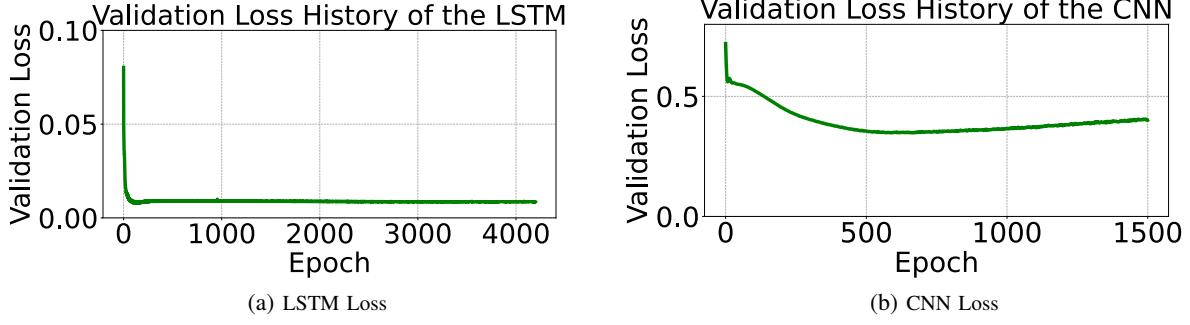


Fig. 1: Surrogate Models Validation Loss History

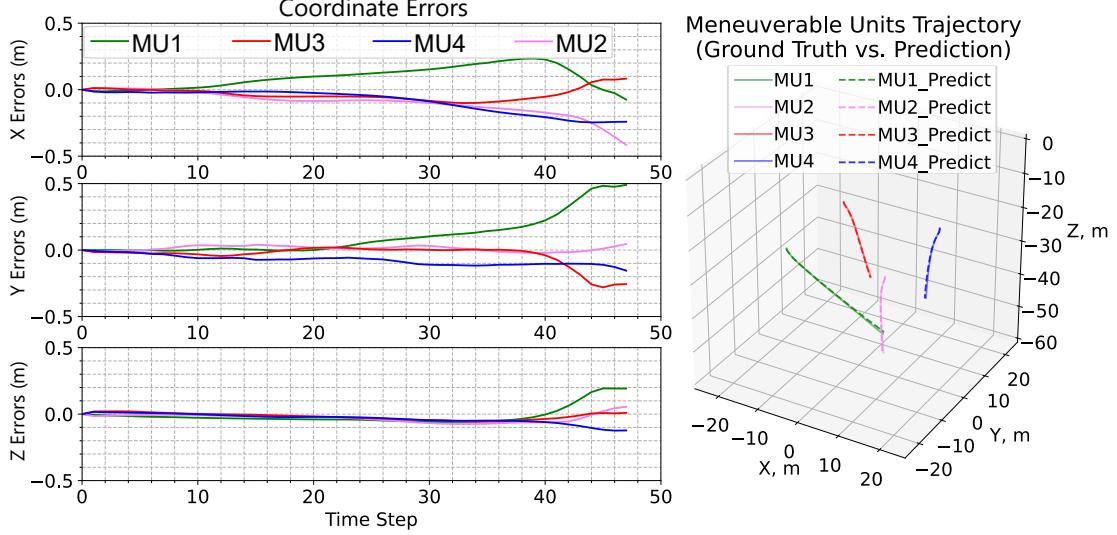


Fig. 2: Example Scenario: Surrogate Predicted Deployment Phase Trajectories vs. that given by the Higher-fidelity Simulation

B. Example Capture Scenario Visualization

Example visualizations of different simulated scenarios are shown in Fig. 5 and 6, utilizing the baseline control and the RL_{PS} -derived policy, respectively. The initial conditions of the scenarios and the control parameters are given in Table II. It can be observed from both visualization examples that in the RL_{PS} derived control, one of the MU acts like an anchor, pulling back the net, while the other three MU generate rotational motion. As a result, the entire tether-net system behaves like a flail, swinging and propelling toward the target. This is partly because the allowed thrust angle range restrains the MUs from flying directly towards the target in RL_{PS} control, unlike the baseline control. It should also be noted that the baseline control resulted in a failed debris capture in this scenario, unlike when the RL_{PS} control is applied. It should be noted that this is a first-of-its-kind example demonstration on how active rotation of the net (based on learned control in this case) relative to the debris provide capture benefits.

TABLE II: Target States and MUs Control Parameters in the Example Scenario Shown in Figs. 5 and 6

	Variables	Unit	Value
Target States	X_D, Y_D, Z_D	m	-6.10, 7.10, -56.69]
	$O_{D_{x0}}, O_{D_{y0}}, O_{D_{z0}}$	deg	-43.77, -58.41, -82.02
MU Thrust Controls (Baseline)	$\omega_{D_{x0}}, \omega_{D_{y0}}, \omega_{D_{z0}}$	deg/s	7.79, 8.23, 16.80
	$\psi_{T,1}, \psi_{T,2}, \psi_{T,3}, \psi_{T,4}$	deg	-33.43, -137.61, 43.13, 127.67
	$\theta_{T,1}, \theta_{T,2}, \theta_{T,3}, \theta_{T,4}$	deg	13.98, 11.50, 15.89, 13.81
	F_T	N	8.5
MU Thrust Controls (RL_{PS})	$\psi_{T,1}, \psi_{T,2}, \psi_{T,3}, \psi_{T,4}$	deg	50.60, -45, 34.82, 109.11
	$\theta_{T,1}, \theta_{T,2}, \theta_{T,3}, \theta_{T,4}$	deg	55.0, 35.0, 50.0, 35.0
	F_T	N	5

C. Robustness Analysis

Table III shows the comparison of the success rate and total fuel cost of the RL methods and the Baseline methods with uncertainty in the observations. The values in the lower-half table represent the results when the observations of the target state

TABLE III: Comparison of the Performances of RL Policies and the Baseline Method on Unseen Test Scenarios with Noisy Observations – extended results

Noisy Observation Tests (100 scenarios, 20 MC sample for each)					
Method	Success Rate (%)	Total Fuel Cost (kg)			
		Median	Mean	Std	
Baseline	69.0	0.185	0.186	0.0187	
RL-PS	75.0	0.158	0.157	0.0171	
RL-MFU	69.9	0.160	0.160	0.0162	
RL-SWS	78.3	0.169	0.170	0.0191	

are noisy. This comparison aims to analyze the robustness of the control methods when potential uncertainties are attributed to imperfect sensing and/or state estimation.

To evaluate the robustness of the control strategy under observation noise, we first generated 100 test scenarios using Latin Hypercube Sampling (LHS). For each of these 100 scenarios, 20 noisy variants were created by adding Gaussian noise to the target state. Note that the variant target states are only for the inputs to the control methods, while in the simulation, the target states are still defined by the 100 test scenarios. Each noisy variant includes perturbations applied to all 9 components of the target state vector. The Gaussian noise has zero mean and a standard deviation equal to 10% of the respective target state component's range, as defined in the paper. This setup reflects the uncertainty introduced by sensing and/or state estimation errors in practical applications. Note that for the baseline method, only target positions are influential to the control, while for the RL methods, all 9 components (position, initial orientation, and angular velocity in X, Y, and Z axes) are influential. The number of common successful capture scenarios is 769 out of 2,000 scenarios.

As Table III shows, with the noisy observations, the mean and median fuel costs of all the RL methods are still significantly lower than that of the baseline method, and the success rates of all the RL methods are higher than that of the baseline method. This shows that the RL generated policies provide some implicit robustness, even when they are not explicitly trained on sample scenarios with noisy observation. Another finding is that the standard deviations of all the methods are not affected much by the uncertainty in the observation.

D. Analyzing Differences in Control Actions

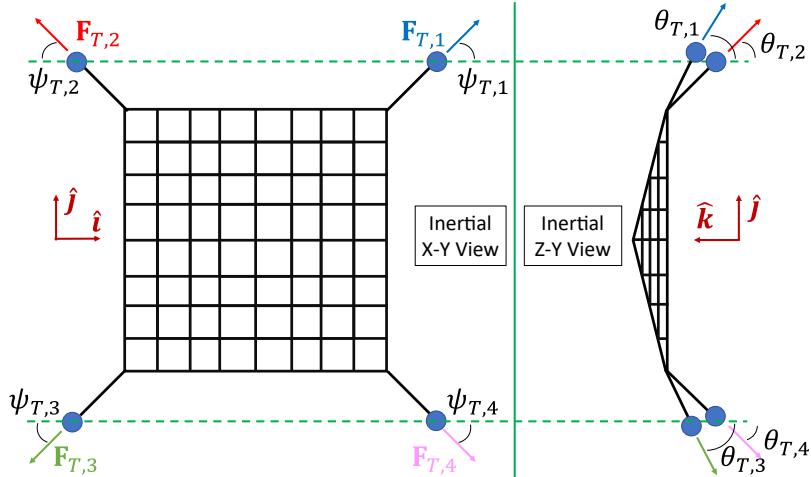


Fig. 3: Thrust angles $\psi_{T,i}$ and $\theta_{T,i}$ for the MUS.

Here we analyze how the control actions (thrusts) vary across the RL-derived policies, and differ from those given by the baseline. The thrust magnitude F_T is found to converge to a value of 5.0 N in each of the RL methods – notably smaller than the 8.5 N prescribed in the baseline – and stays mostly invariant across scenarios. This value corresponds to the lower bound set for this action parameter, with the following likely explanation. This lowest thrust magnitude still allows suitably large net mouth opening at t_{open} , while keeping fuel costs as low as possible. However, the lower net speed resulting from the low thrust magnitude should affect how the net wraps around the target, where this effect is also impacted by the target's rotation rates; this phenomenon requires further analysis to fully understand the tendency of the RL policies to converge to the lowest thrust magnitude setting.

Figure 4 shows the distribution of the difference between the thrust angles computed by each of the RL derived policy and the baseline method, i.e., ($|\psi_{T,i} - \psi_{T,\text{base},i}|$) and ($|\theta_{T,i} - \theta_{T,\text{base},i}|$). Since the baseline method computes the thrust angles to

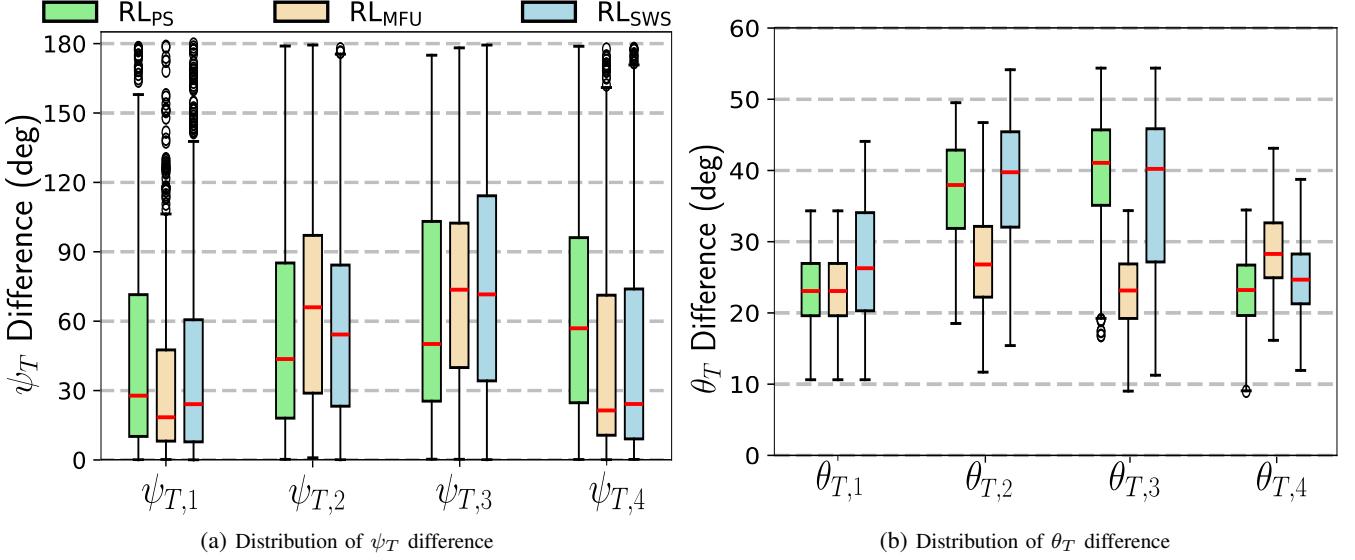


Fig. 4: Distribution of the difference of RL and baseline thrust angle control in the 500 successful scenarios. $\psi_{T,i}$ and $\theta_{T,i}$ are the thrust angle on the $\hat{\mathbf{i}}\cdot\hat{\mathbf{j}}$ plane and $\hat{\mathbf{j}}\cdot\hat{\mathbf{k}}$ plane, respectively (shown in Fig. 3).

use from the aiming point vectors, this analysis provides an understanding of how the RL-derived policy deviates or adjusts these angles, so as to improve capture success and fuel costs.

It can be seen from Fig. 4(a) that the RL-generated thrust angles ψ_T , defined on the X-Y inertial frame, varied notably from that of the baseline. This relatively large difference, ranging from 0° to 180° , can be partly attributed to the broadly defined range of the RL-controlled ψ_T (i.e., -45° to 135°). The difference also indicates that the RL has effectively explored and identified thrust angles significantly distinct from the baseline, leading to successful captures. The differences between the RL and baseline given θ_T , i.e., $\theta_{T,i} - \theta_{T,\text{base},i}$, are comparatively smaller, as shown in Fig. 4(b). This is due to the narrower range (i.e., 35° to 55°) allowed for θ_T in the RL-trained policies. The median differences for θ_T is also found to be consistently positive, suggesting that the RL policy favors wider net openings. This is to be expected as the reward term R_{open} maximizes the net's mouth area at t_{open} .

REFERENCES

- [1] F. Liu, A. Boonrath, E. M. Botta, and S. Chowdhury, "Surrogate-aided learning of active tether-net maneuver to capture rotating space debris," *IEEE Transactions on Aerospace and Electronic Systems (Under Review)*, 2025.

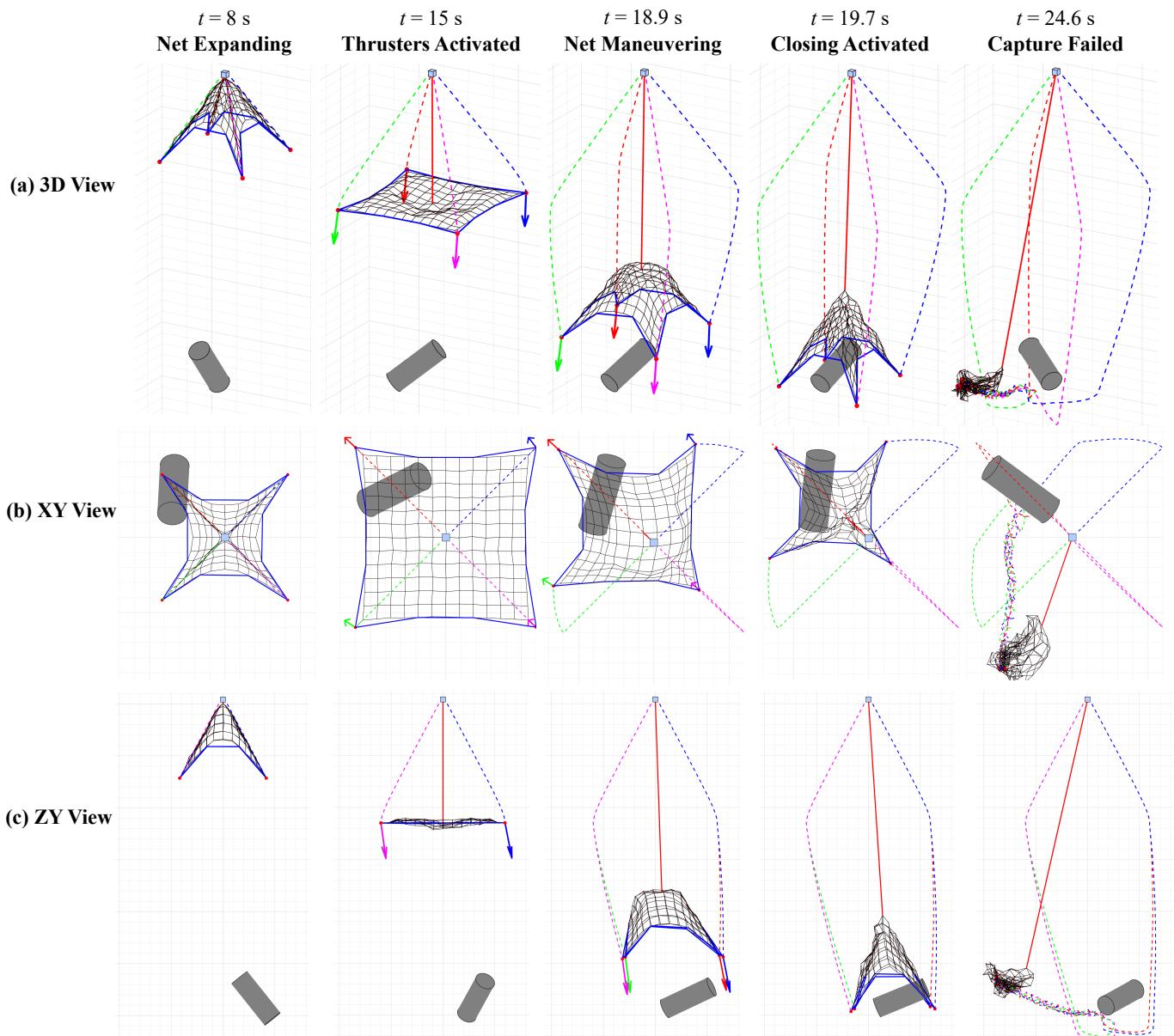


Fig. 5: Example Visualization with the Baseline Aiming Point Control

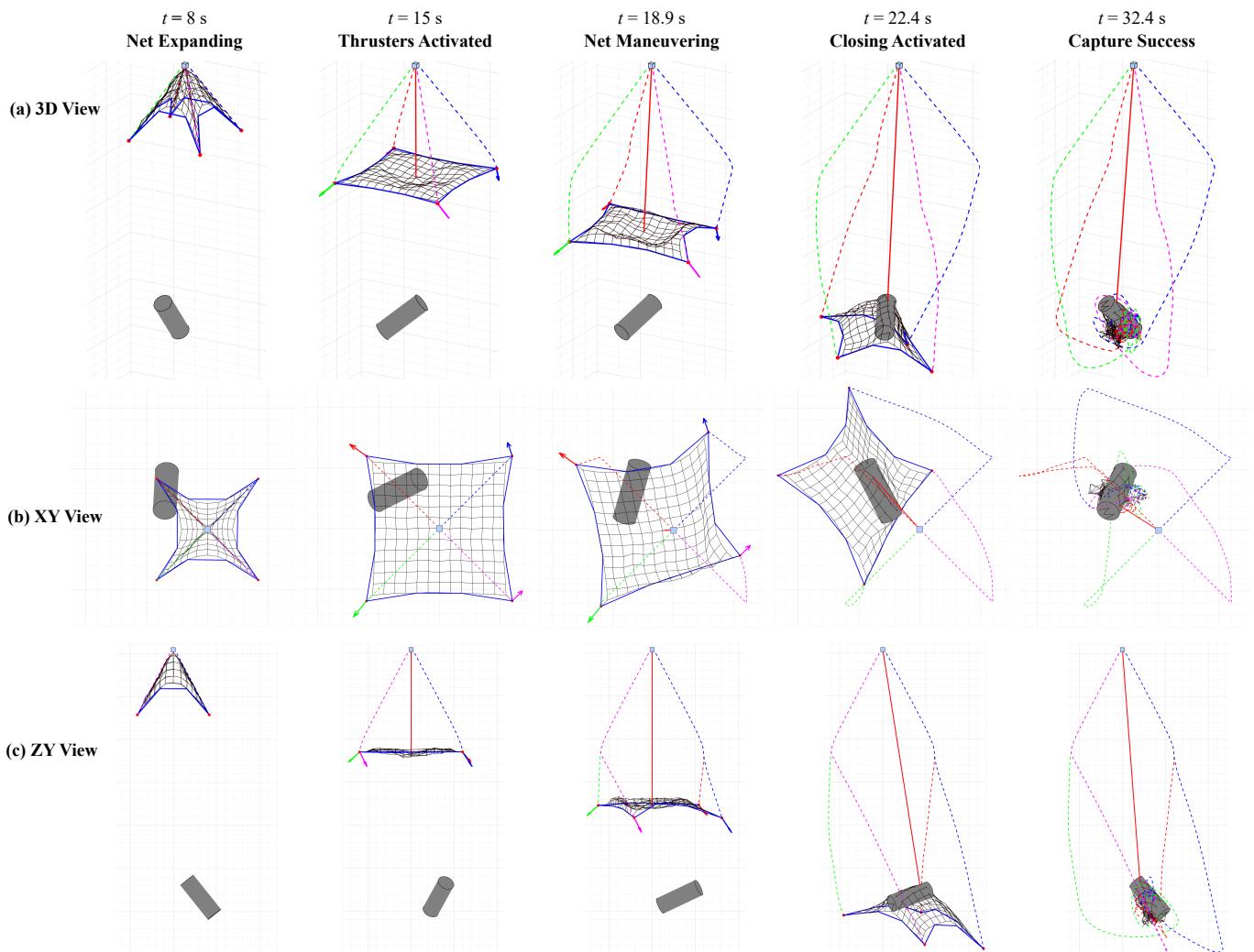


Fig. 6: Example Visualization with RL_{PS} -derived Control Policy