



# From Space to Time: Enabling Adaptive Safety with Learned Value Functions via Disturbance Recasting



Website

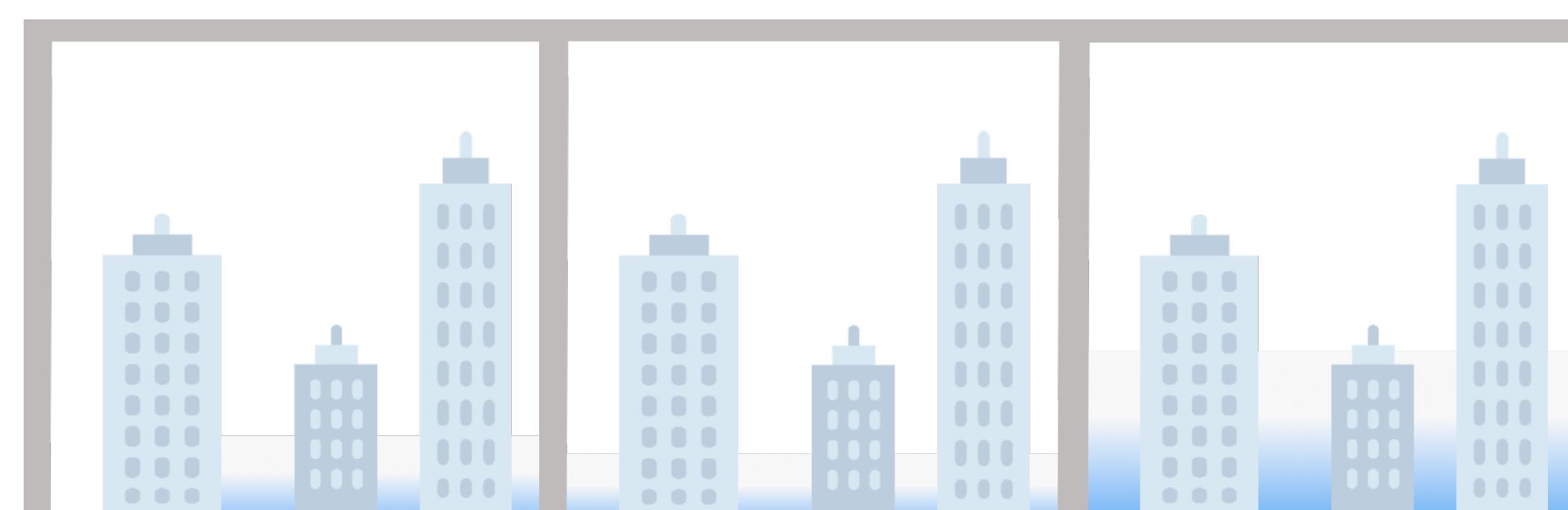
Sander Tonkens\*, Nikhil Uday Shinde\*, Azra Begzadić\*, Michael C. Yip, Jorge Cortés, Sylvia Herbert

Paper

## Motivation and Background:

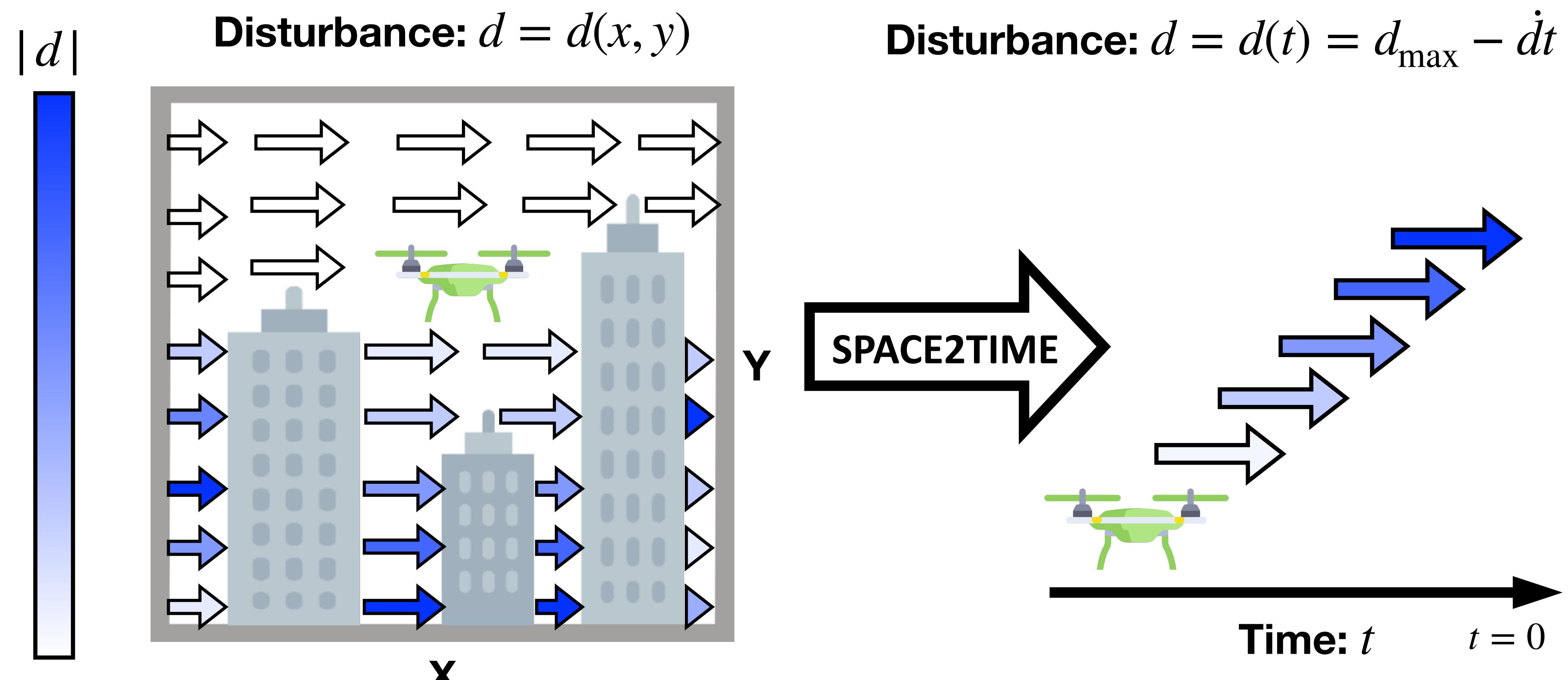


Not accounting for spatial variation in wind can lead to catastrophic failures during drone flight



Learned value functions are not amenable to online fine tuning, and learning a separate value function for every possible spatial variation is infeasible

**Key Idea:** Reparameterize spatial variations in disturbance as temporal variations.

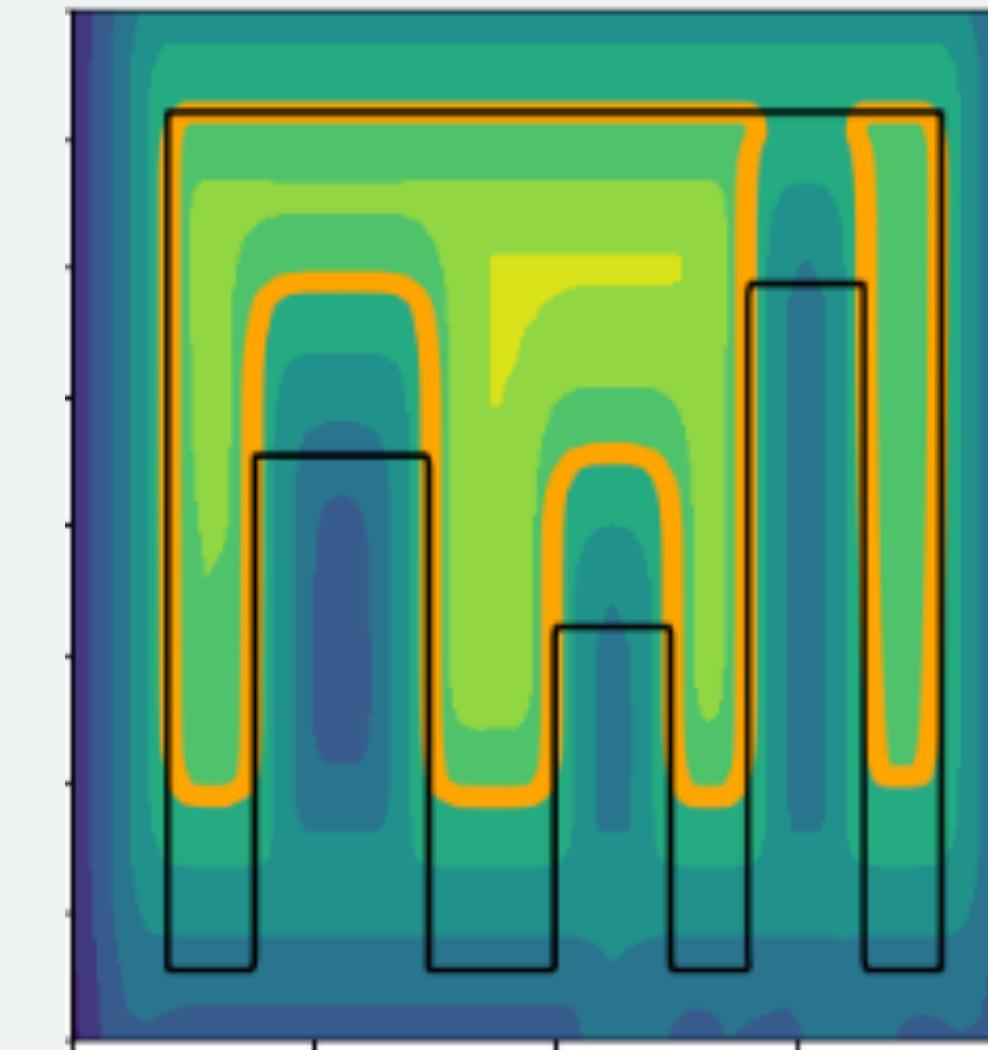
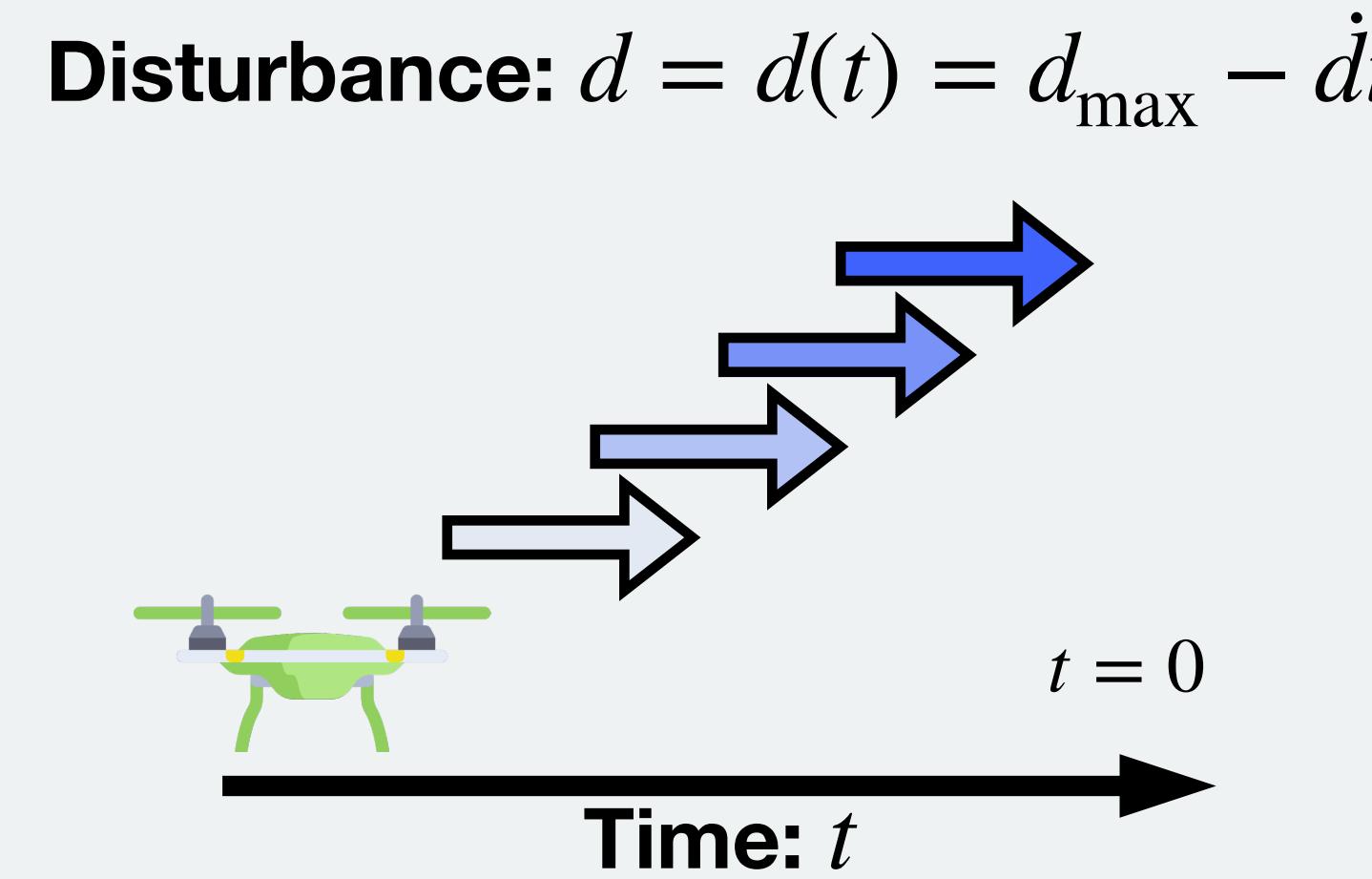


Enables **learning safety value functions offline** and using them in spatial varying environments *online*. SPACE2TIME ensures the system is **realistic about the present** (based on current disturbance) and **pessimistic about the future** (assuming worst measured rate of disturbance change to persist)

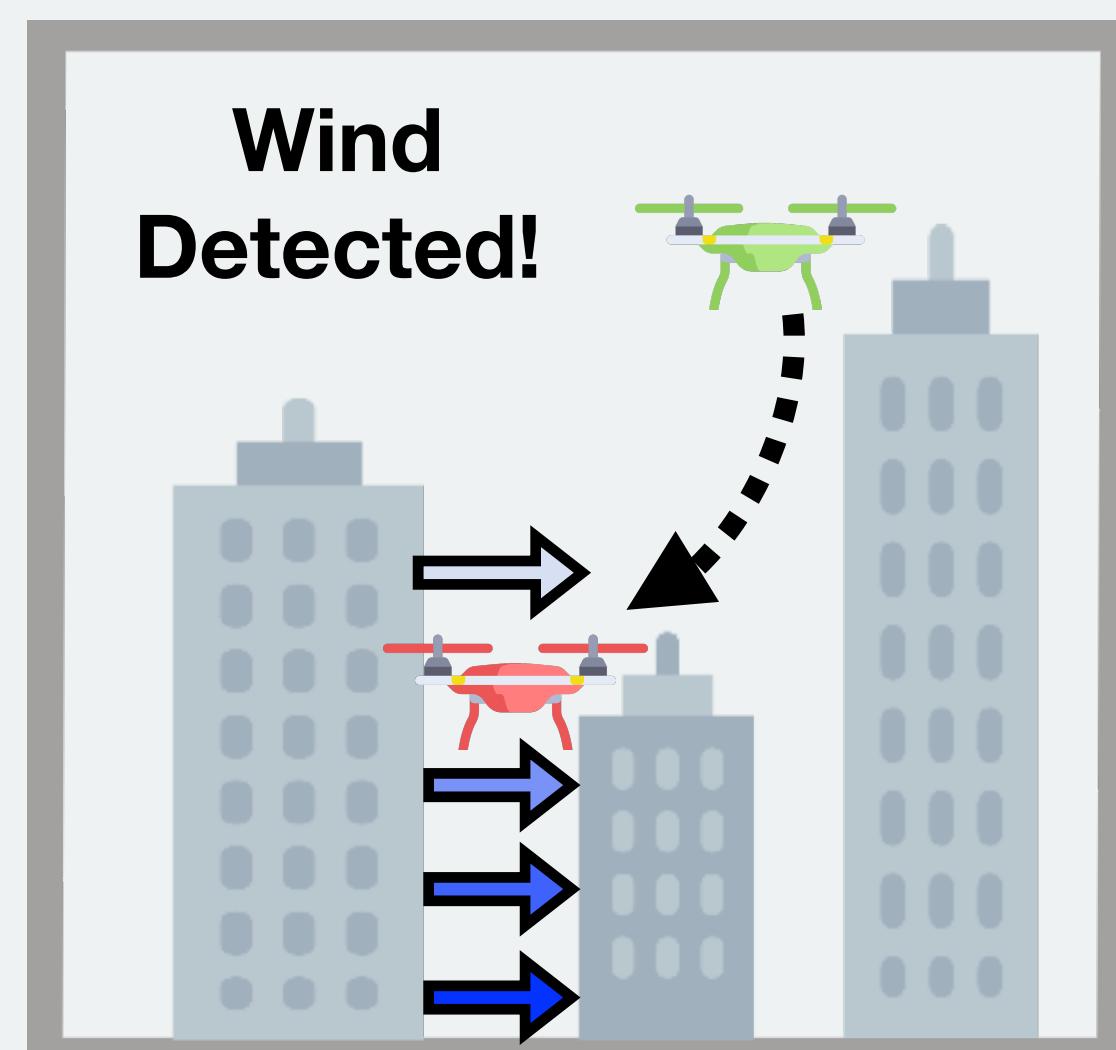


1. Define Safe Backup Set

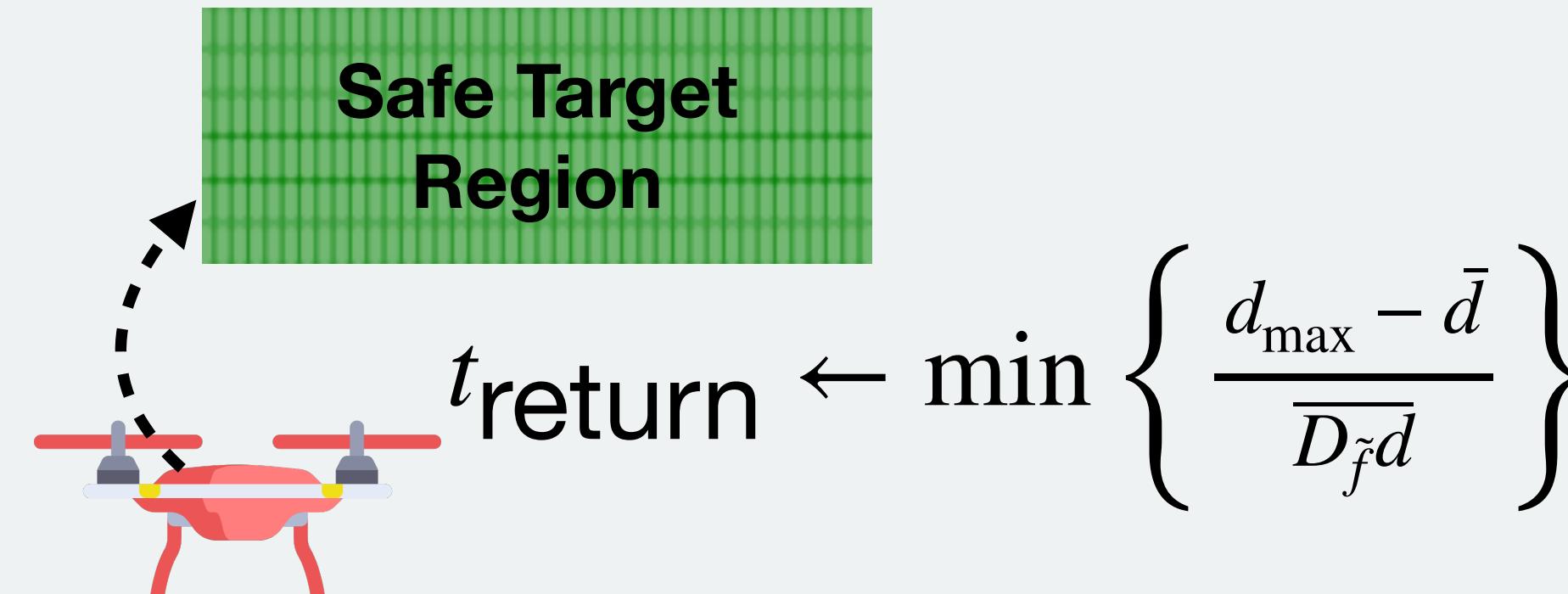
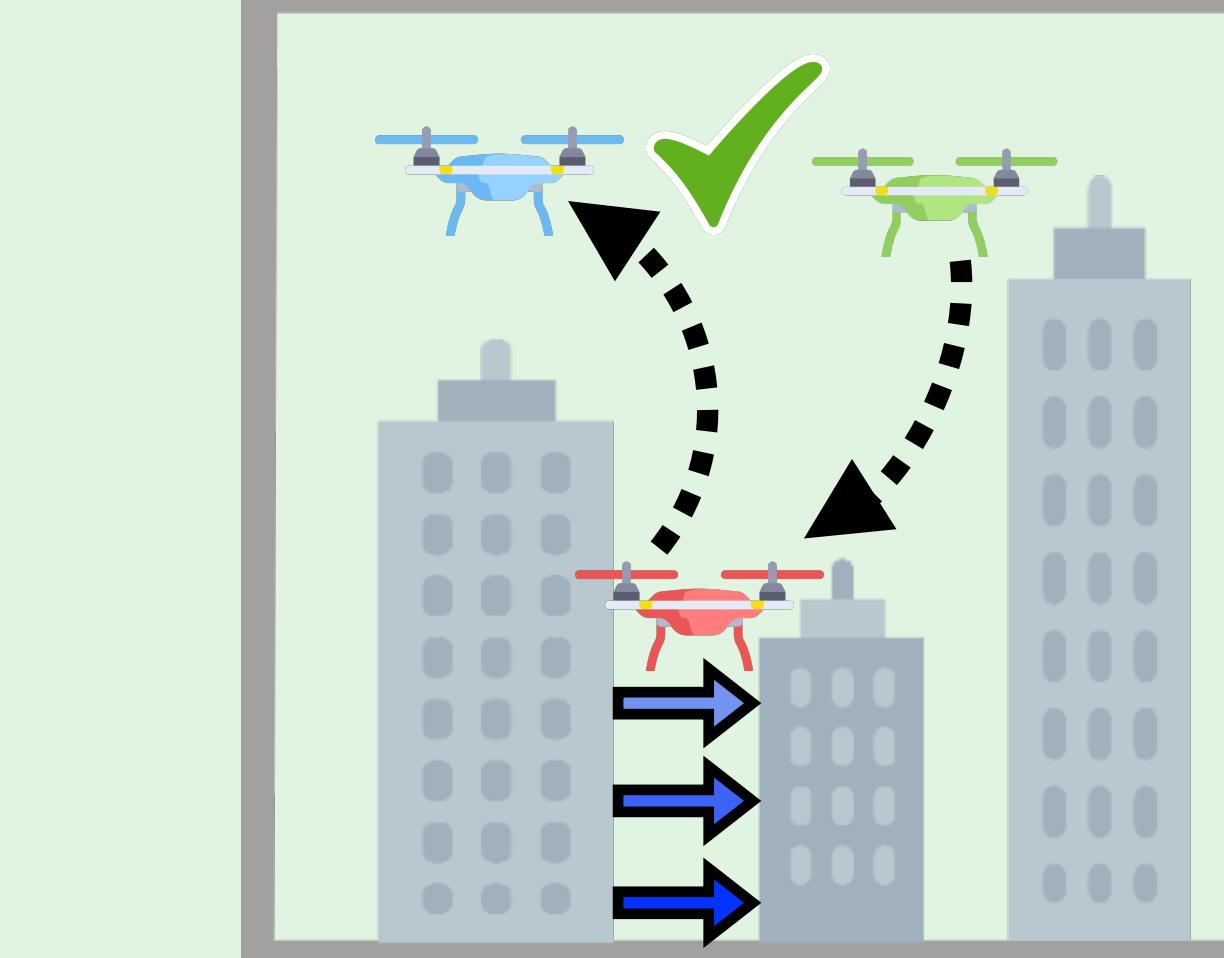
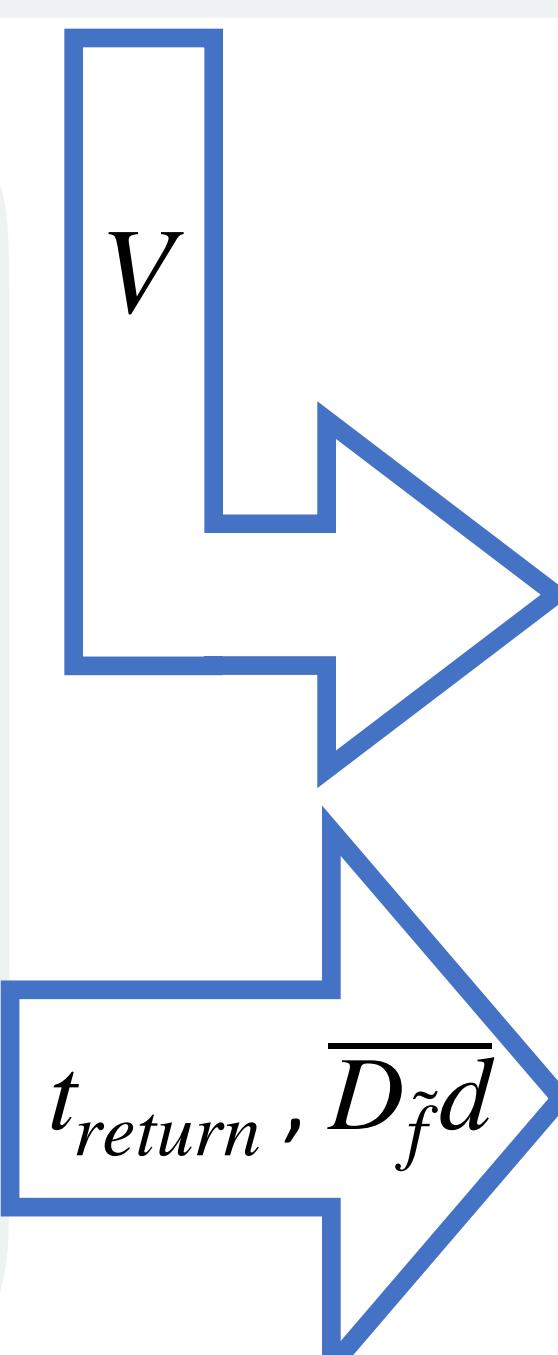
## SPACE2TIME: Offline

 $V(x, t, \bar{d})$ 

3. Learn/Compute Value Functions

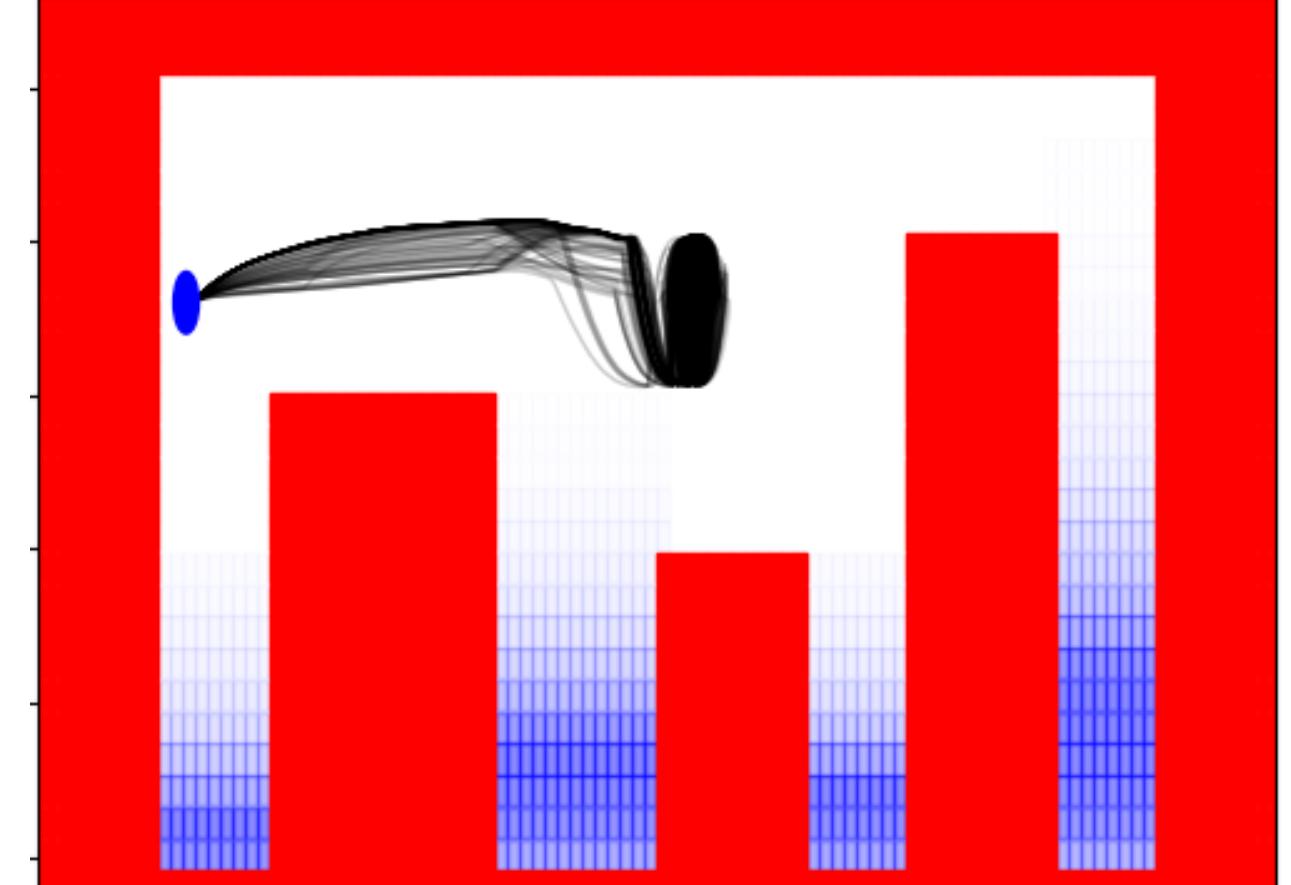
1. Sense  $\bar{d}, \bar{D}_{\bar{f}\bar{d}}$ 

## SPACE2TIME: Online $\text{Hz}$

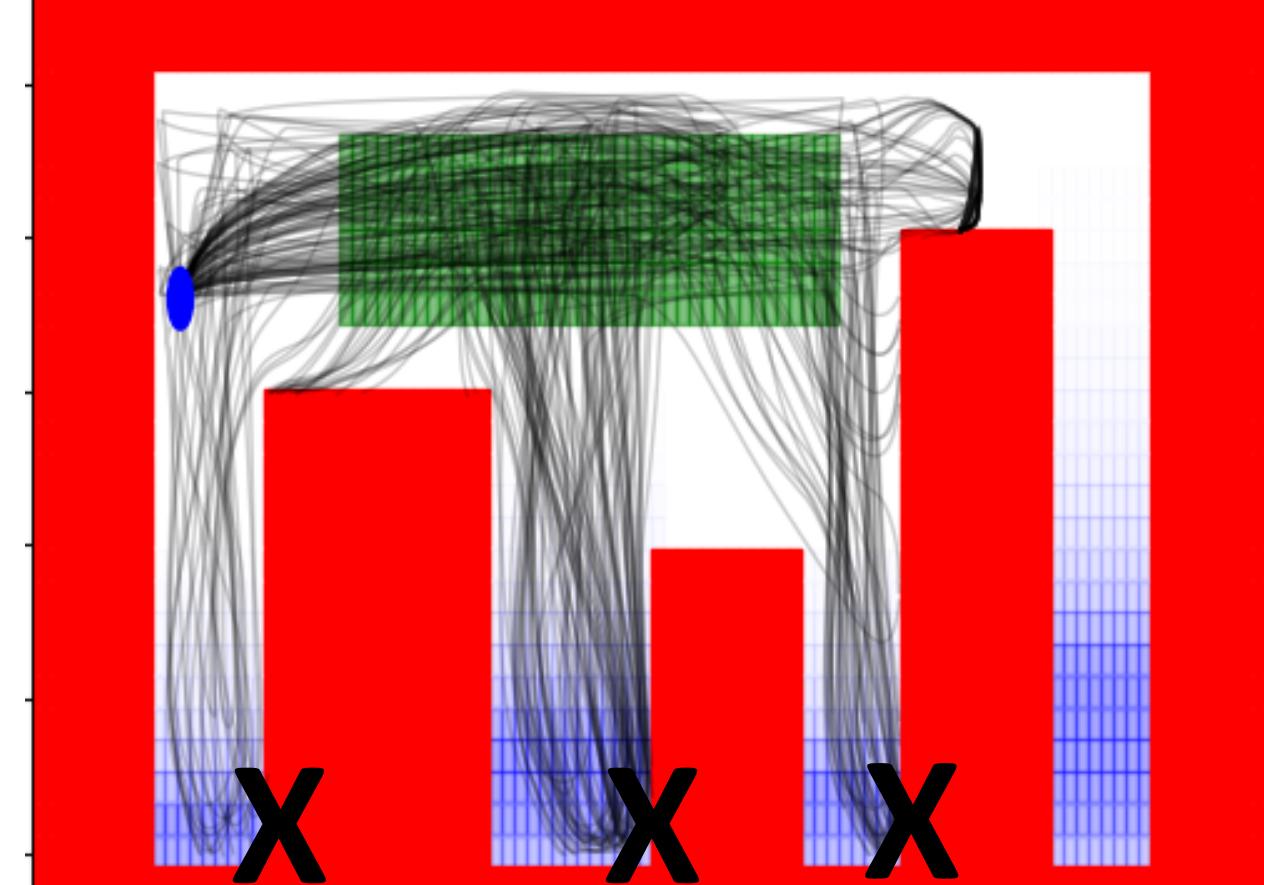
2. Compute  $t_{\text{return}}$ : time to return to safe set3. Minimally invasive safety filter using:  $V(x, t_{\text{return}}, \bar{D}_{\bar{f}\bar{d}})$ 

## Results:

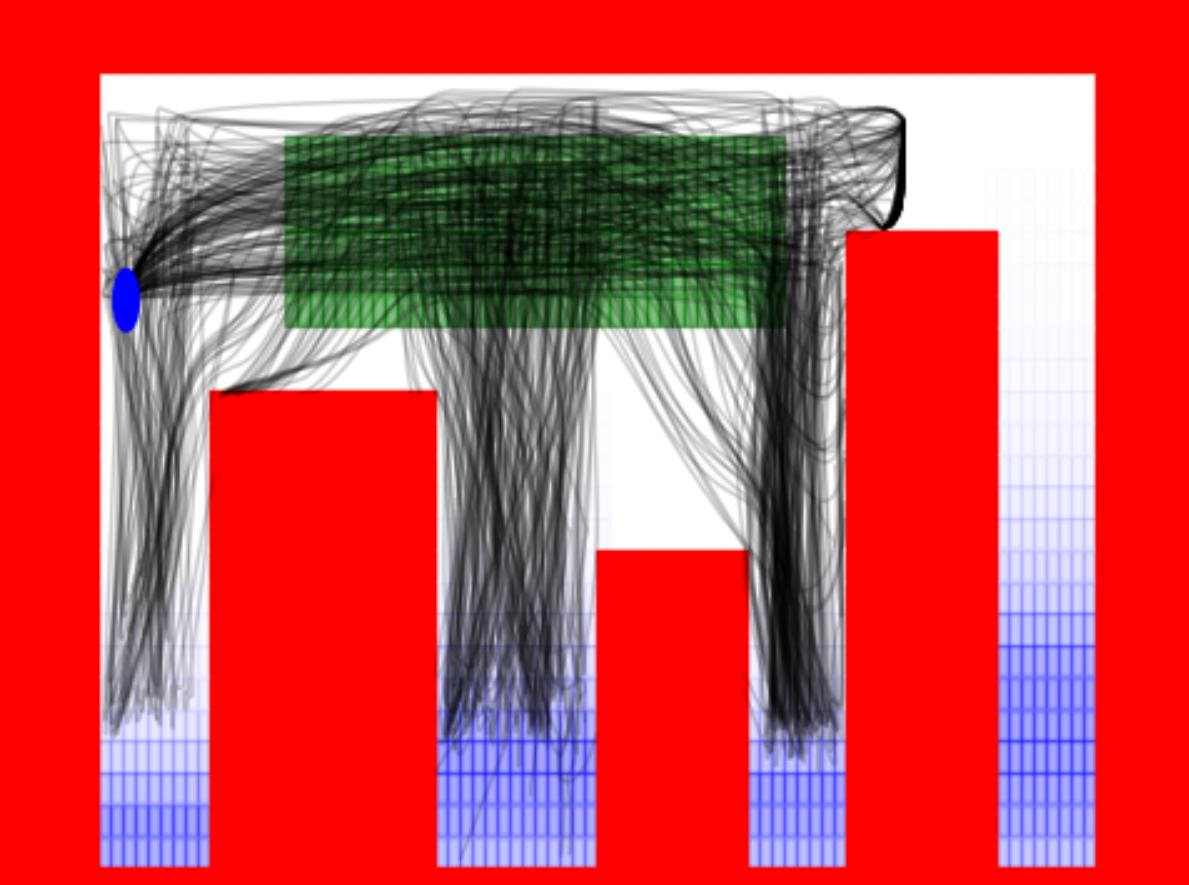
- **Naive Worst Case:** Safety filter assuming worst case disturbance magnitude bounds
- **Naive Disturbance Switching:** Switching safety filter to use disturbance bounds compatible with most recent observations

 $X = \text{Crash}$ 

Naive Worst Case



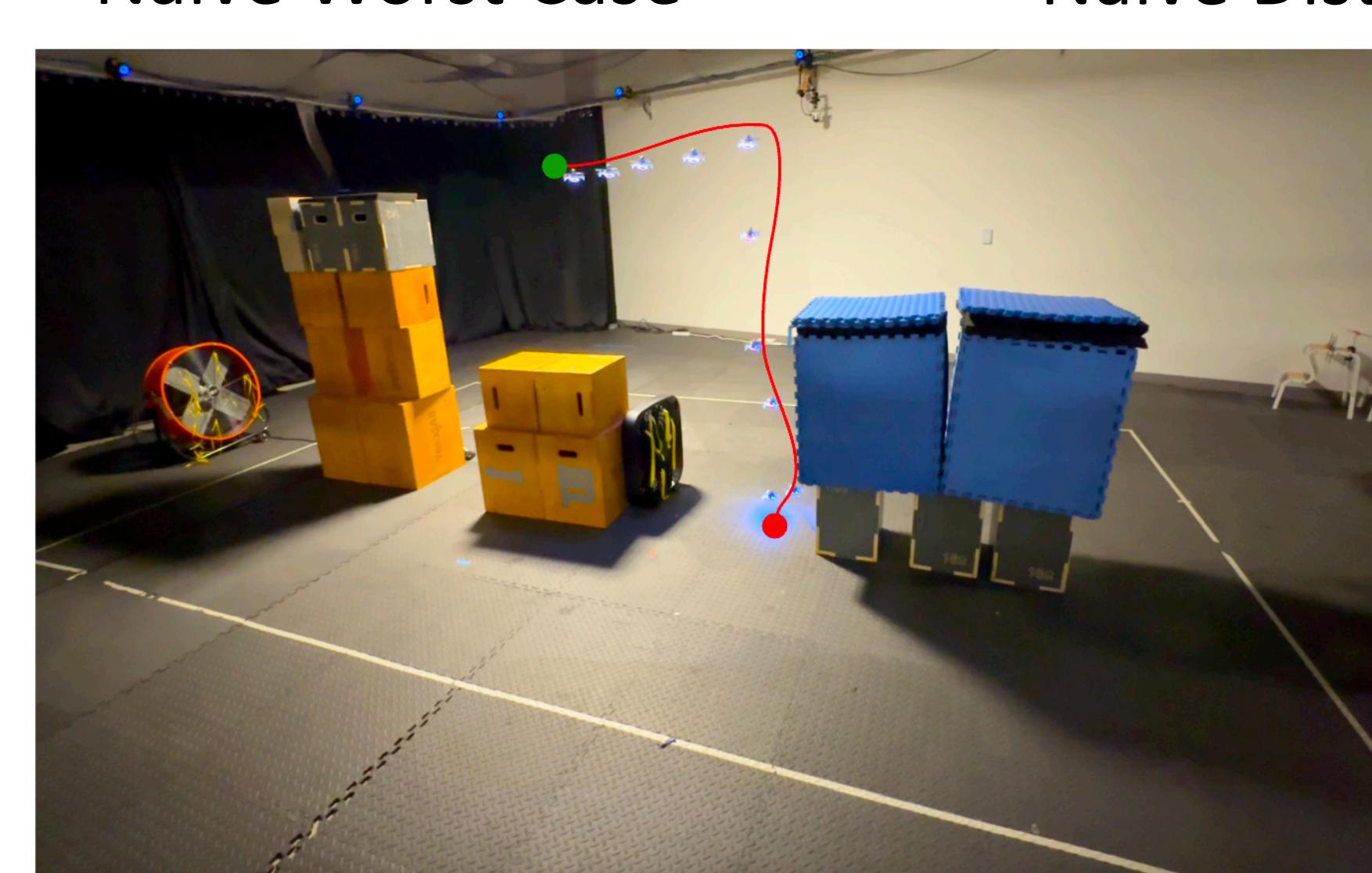
Naive Disturbance Switching



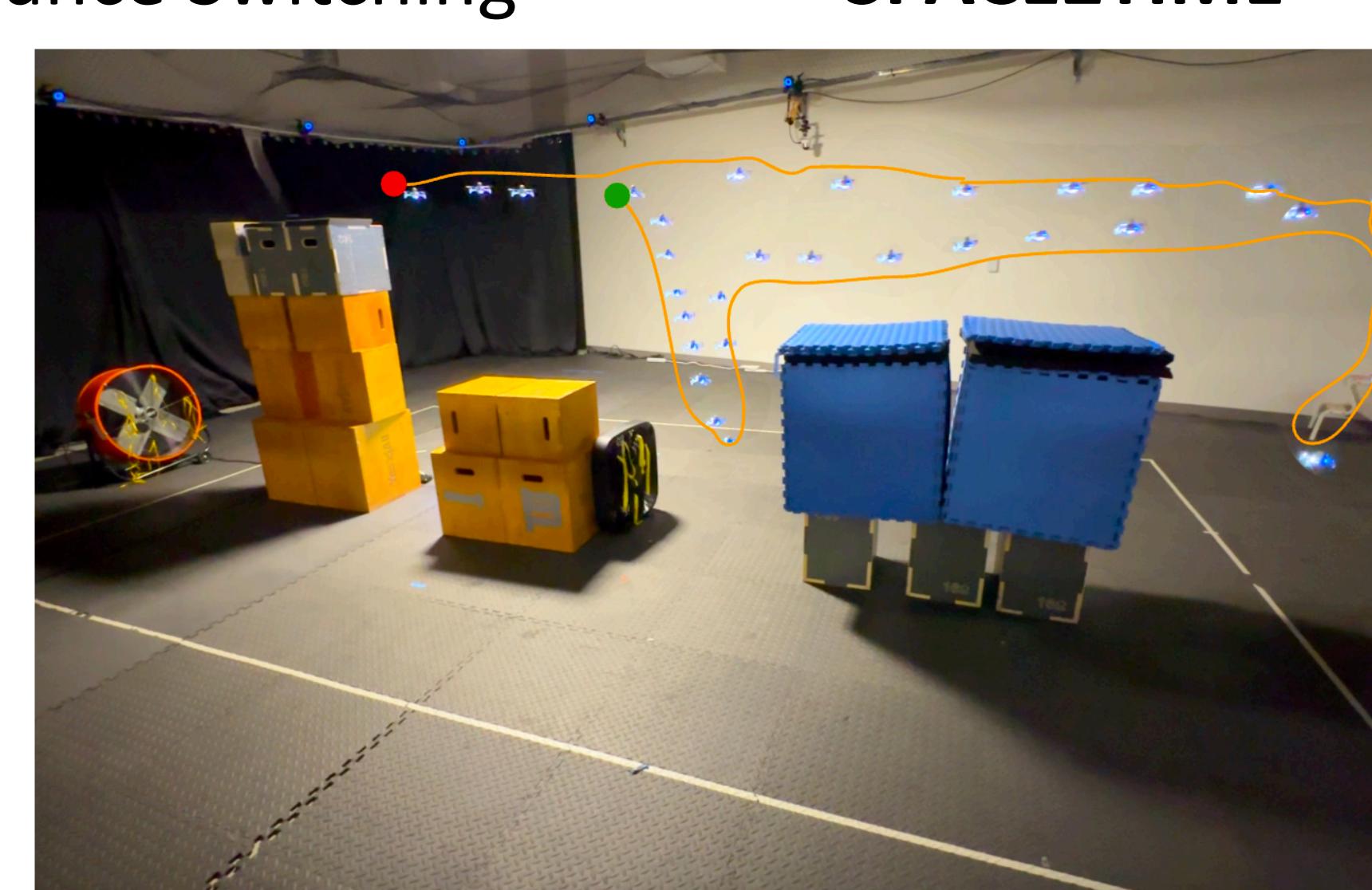
SPACE2TIME

**Table of Simulation Experiments:** Metrics are generated over 100 random environments with 10 goals and 1000 control steps each.

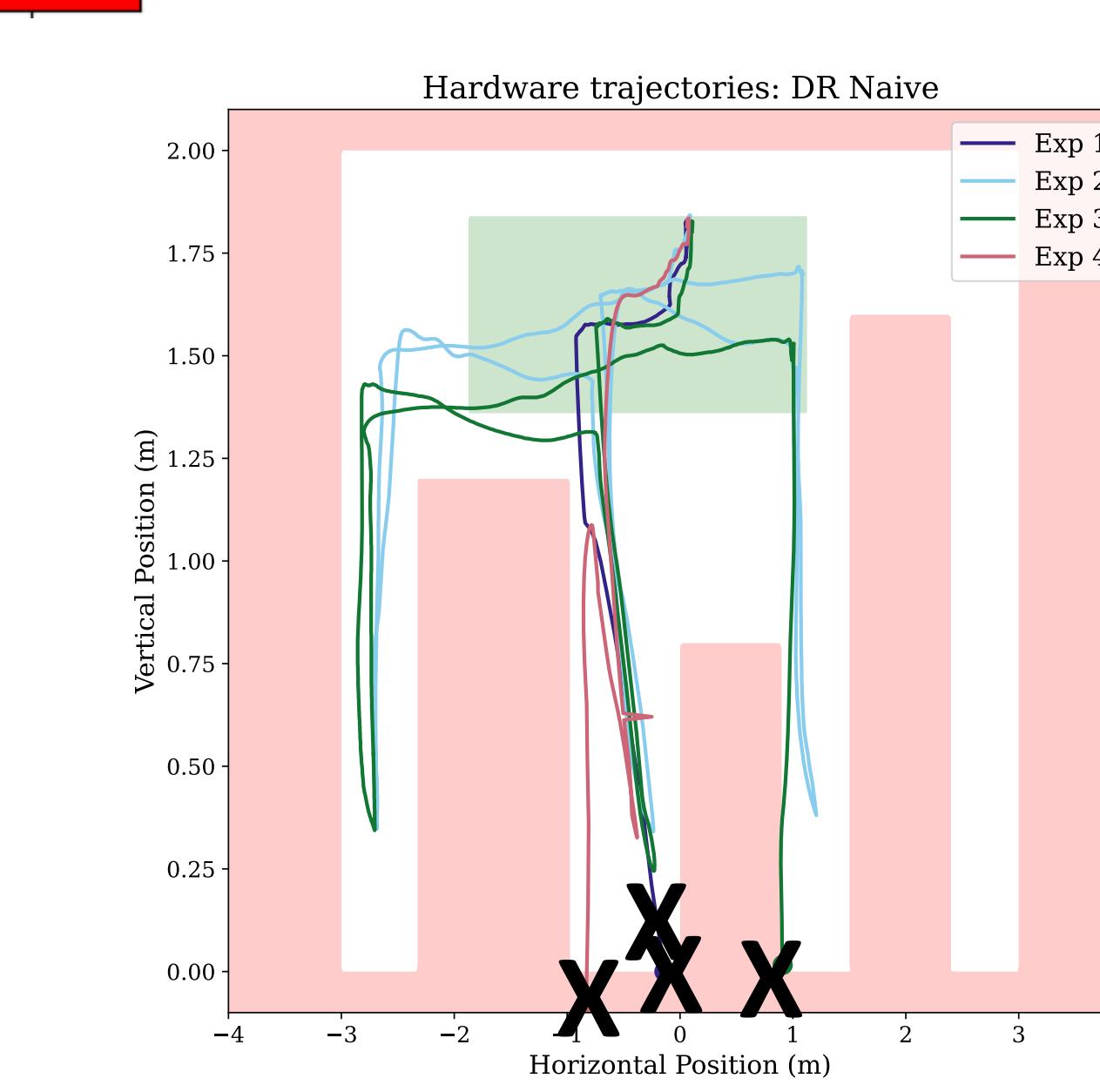
Approach	% Safety Violations ↓	Mean Goal Distance ↓	Mean Trajectory Length ↑
HJR Naive	96%	1.09	415
HJR Naive Worstcase	0%	1.95	1000
DeepReach Naive	90%	1.19	472
<b>DeepReach Ours</b>	34%	1.02	781
<b>HJR Ours</b>	2%	0.78	993



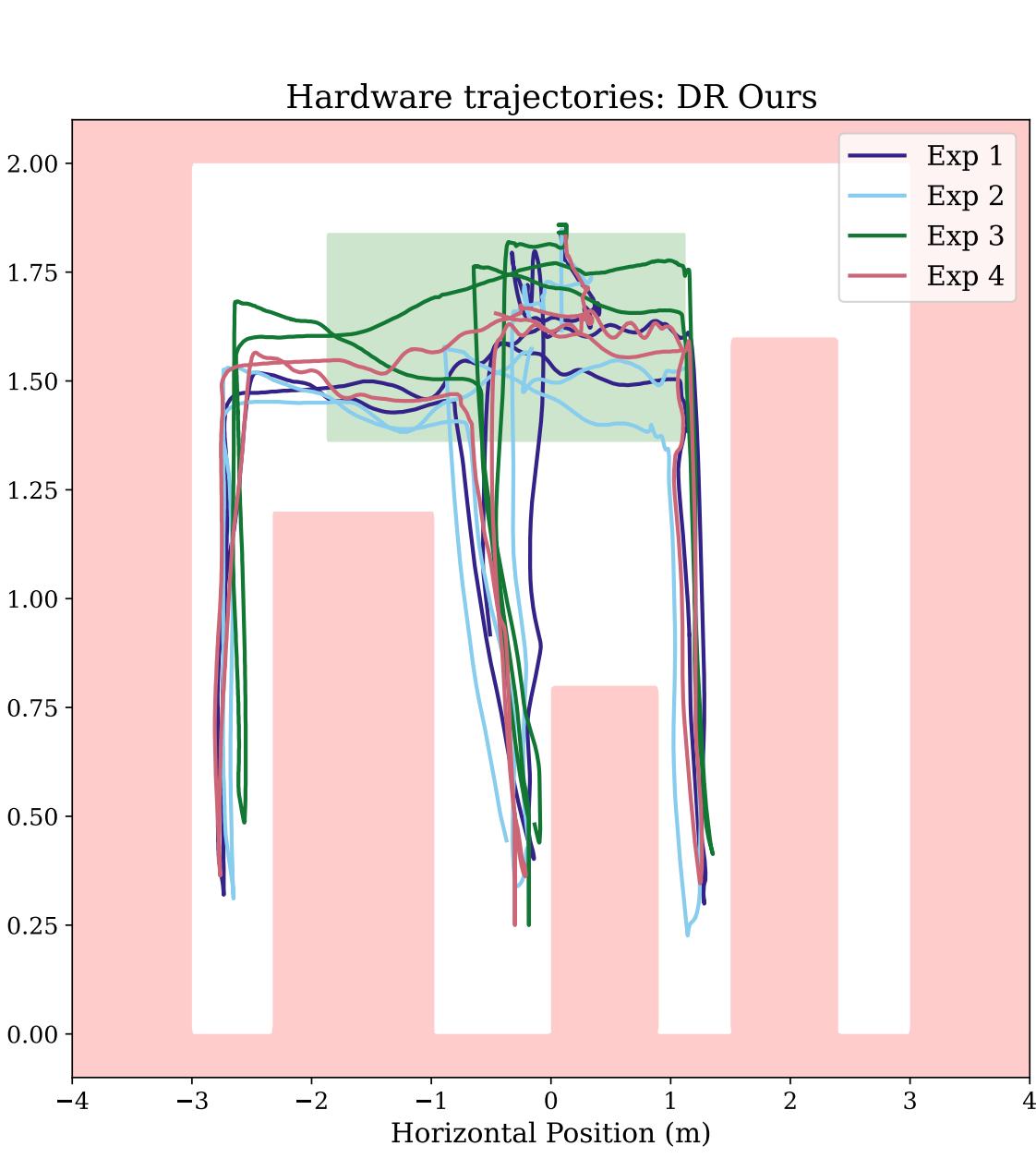
Naive Disturbance Switching



SPACE2TIME



Naive Disturbance Switching



SPACE2TIME