

The Unified Autonomy Stack

Toward a Blueprint for Generalizable Robot Autonomy

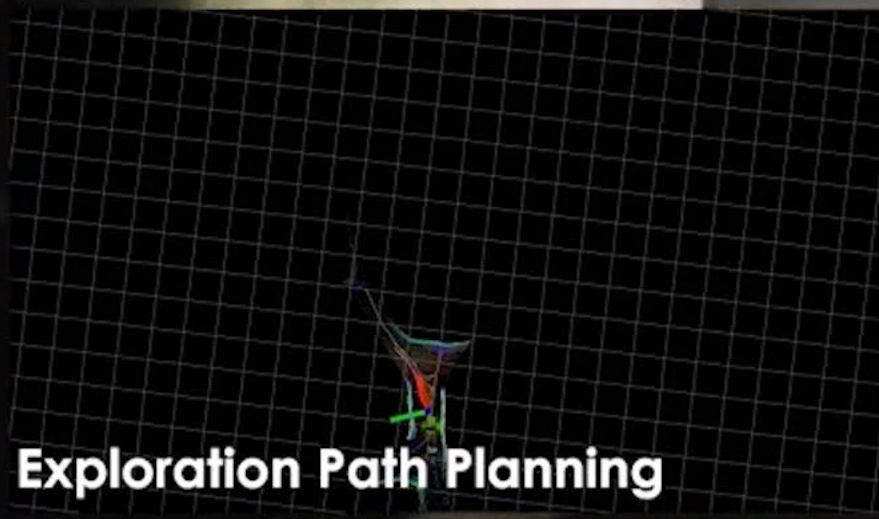
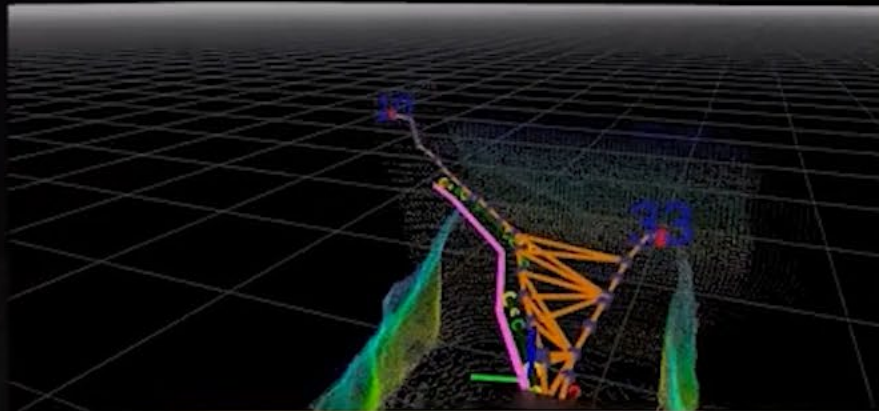
Mihir Dharmadhikari, Nikhil Khedekar, Mihir Kulkarni, Morten Nissov, Martin Jacquet, Angelos Zacharia, Marvin Harms, Albert Gassol Puigjaner, Philipp Weiss, Kostas Alexis



Thermal Camera



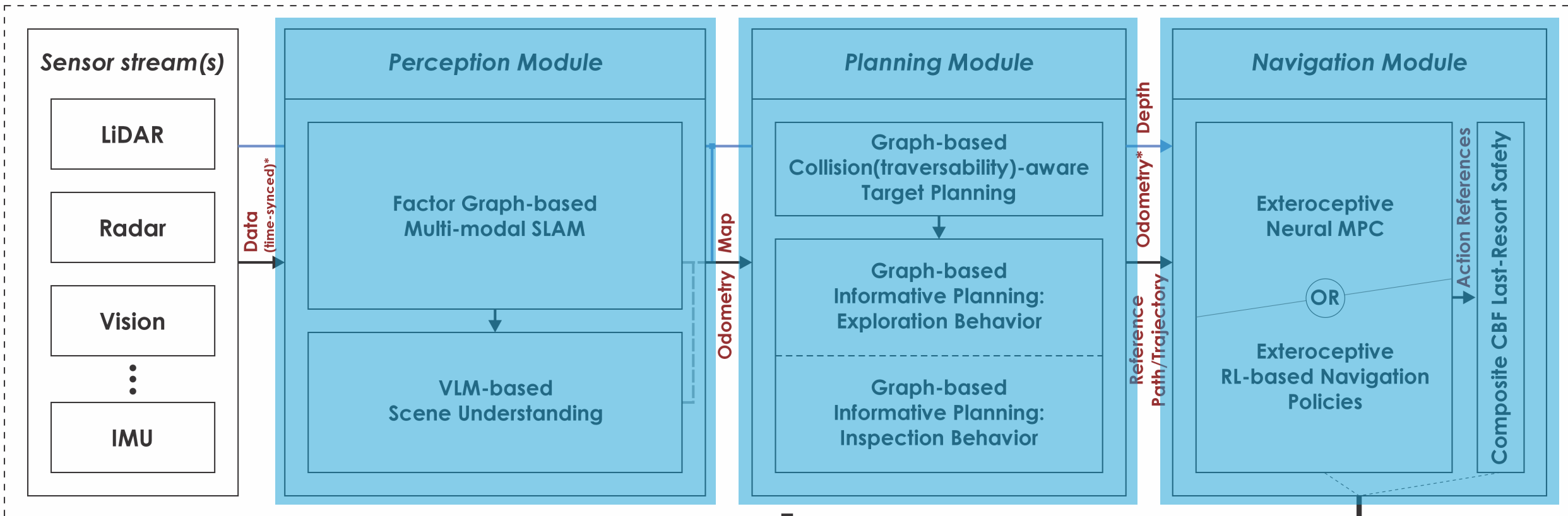
Local Exploration Planner



Exploration Path Planning

Driving Question

Can we consolidate autonomy across morphologies into a single unified stack?



Robots integrate necessary sensing & compute (e.g., UniPilot)

Action References to Low-level Controllers (Autopilots)



* Underwater extension being finalized

Abstraction Layers

Robot Abstraction Layer

What the stack needs to know about the robot:

- robot type — multicopter, rotorcraft, legged, rover
- motion constraints — size, kinematics, traversability
- sensor suite — LiDAR, radar, cameras, IMU with extrinsics & FoVs
- command interface — acceleration (drones) / velocity (ground)
- which safety layers are active

Mission Abstraction Layer

What we want the robot to do:

- task — Target, Exploration, or Inspection
- mission bounds — area, height limits, endurance
- reasoning toggles — VLM Q&A, semantic mapping
- tracking targets — objects-of-interest, language prompts
- behavior switching — e.g. exploration → inspection on completion

Utilized Robot Morphologies



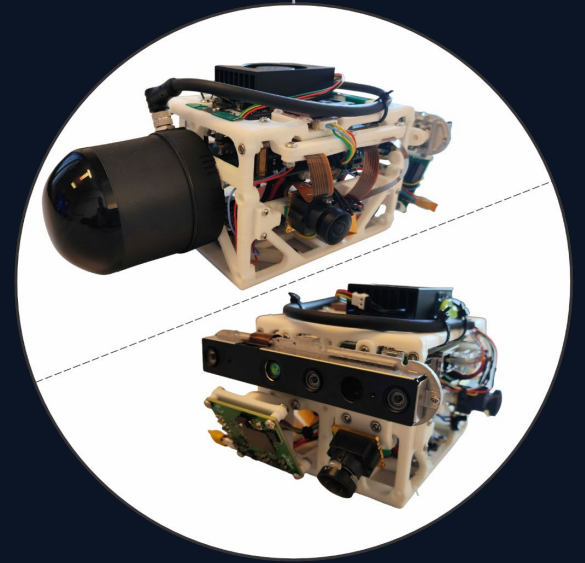
AR-1



AR-2



GR-1

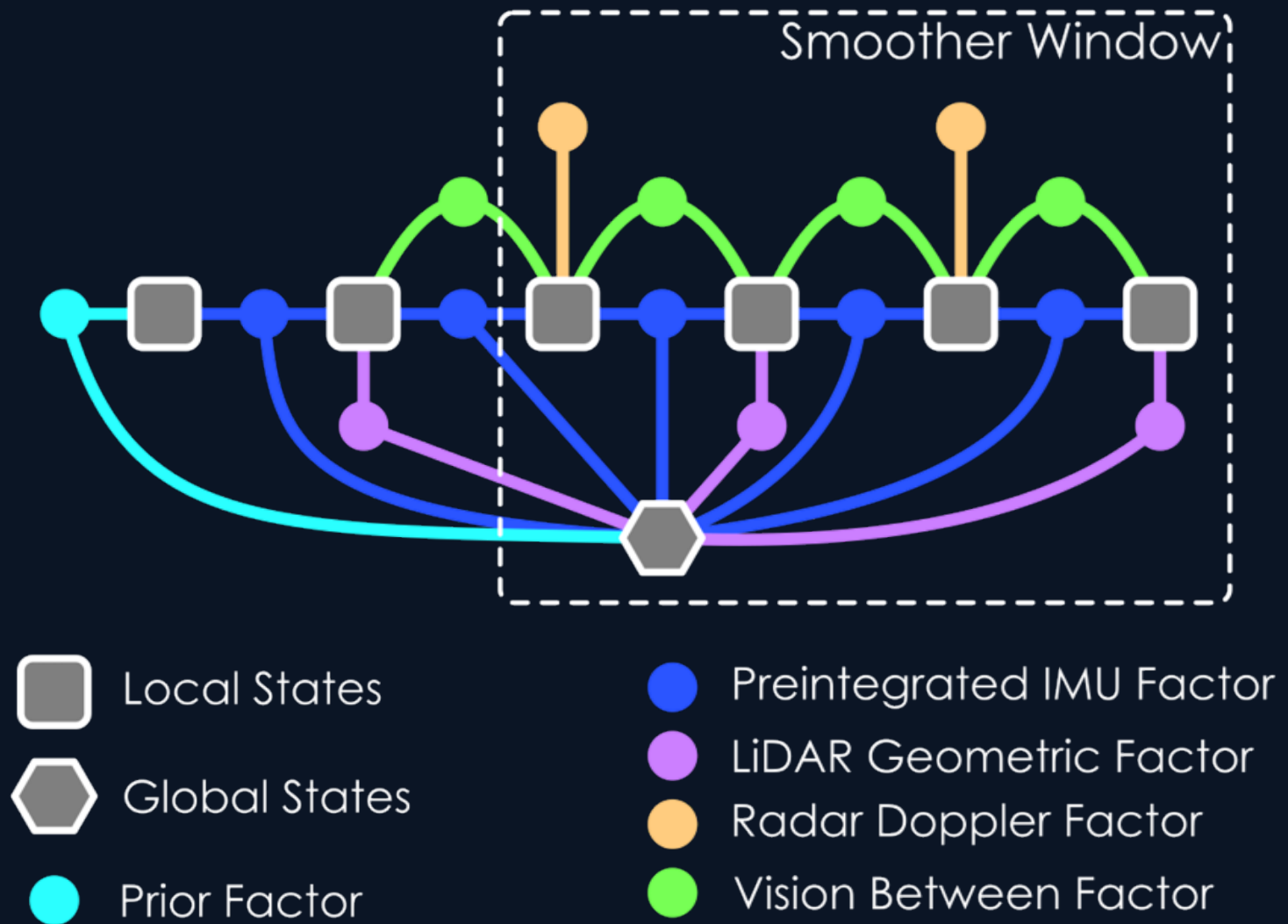
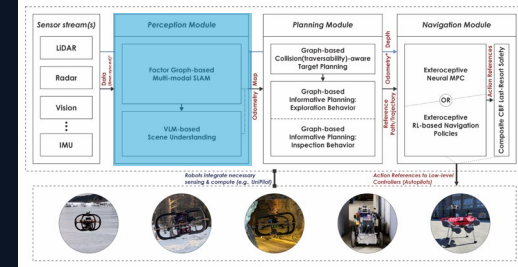


UniPilot

LiDAR	Ouster OS0-128
Radar	TI IWR6843AOP
Vision	FLIR Blackfly S 0.4MP
IMU	VectorNav VN-100
Compute	Khadas VIM4

LiDAR	RoboSense Airy or Hesai JT-128
Radar	TI IWR6843AOP-based unit
Vision	3 x MIPI Vision Components IMX296
IMU	VectorNav VN-100
Compute	NVIDIA OrinNX

Mimosa-X



Urban Tunnel - Fyllingsdal

Urban Tunnel - Runehamar

Frozen Lake

University Campus with Fog

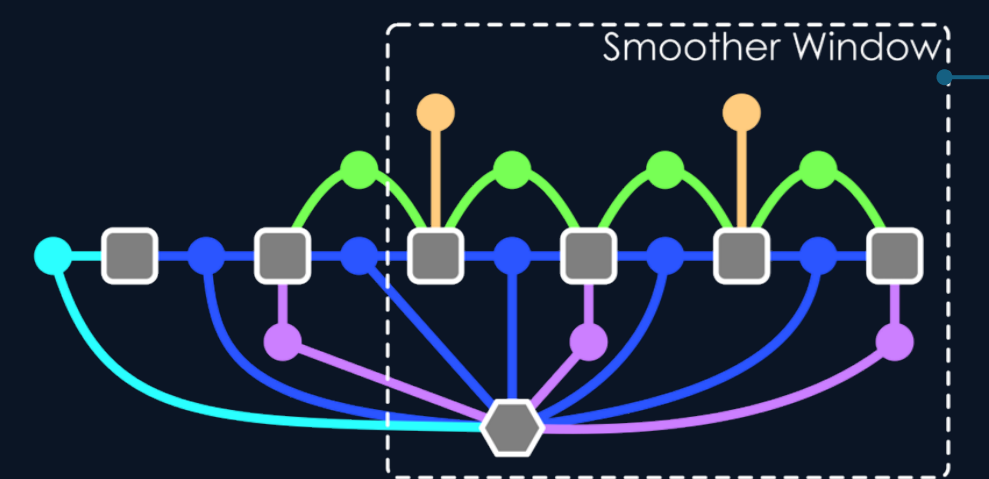
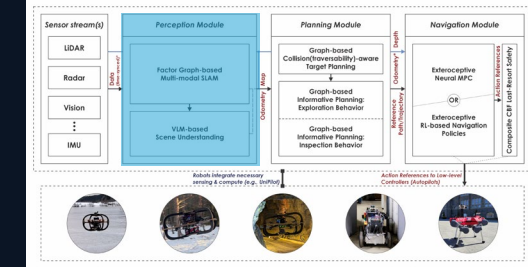
Mimosa-X Evaluation

	Fyllingsdal Tunnel	Runehamar Tunnel	Frozen Lake	Campus Fog	
Length [m]	1275.378	1444.151	826.054	669.609	
ATE [m] / RTE ₁₀ [%]	FAST-LIO2	×	6.578 / 1.891	×	
	FAST-LIVO2	4.930 / 2.852	7.881 / 1.805	×	
	GaRLIO	—	×	×	
	AF-RLIO	×	×	×	
	ROVIO	42.851 / 20.295	62.375 / 29.806	3.112 / 7.177	×
	OpenVINS	13.710 / 6.271	13.337 / 5.443	1.985 / 3.700	×
	Ours - LI	×	7.183 / 1.762	18.661 / 7.840	×
	Ours - RI	21.489 / 6.245	×	10.189 / 30.868	14.075 / 6.189
	Ours - VI	×	×	7.728 / 7.989	×
	Ours - LVI	×	7.012 / 1.763	10.598 / 4.633	×
	Ours - RVI	20.123 / 5.906	×	9.141 / 7.589	17.737 / 6.351
	Ours - LRI	3.899 / 1.649	7.094 / 1.787	10.562 / 6.900	7.324 / 4.240
	Ours - LRVI	3.872 / 1.642	7.048 / 1.789	10.030 / 8.609	8.345 / 4.231

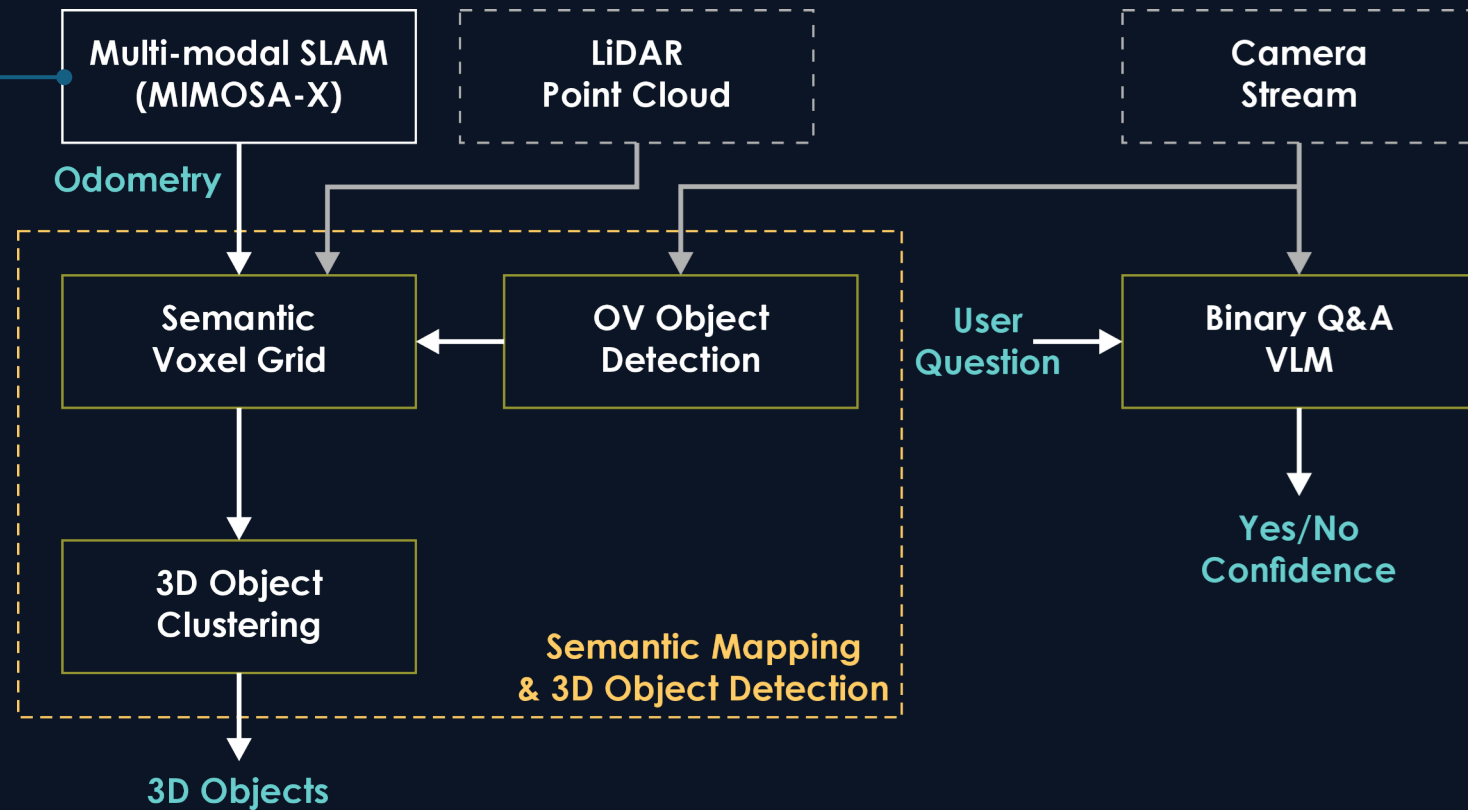
Method failure due to ATE > 5% is indicated by × and due to inability to generalize to sensor configuration (radar being at 25Hz) is indicated by —.

For verifiability and reproducibility, the full implementation and the dataset involved in these studies are openly released.

VLM Scene Reasoning



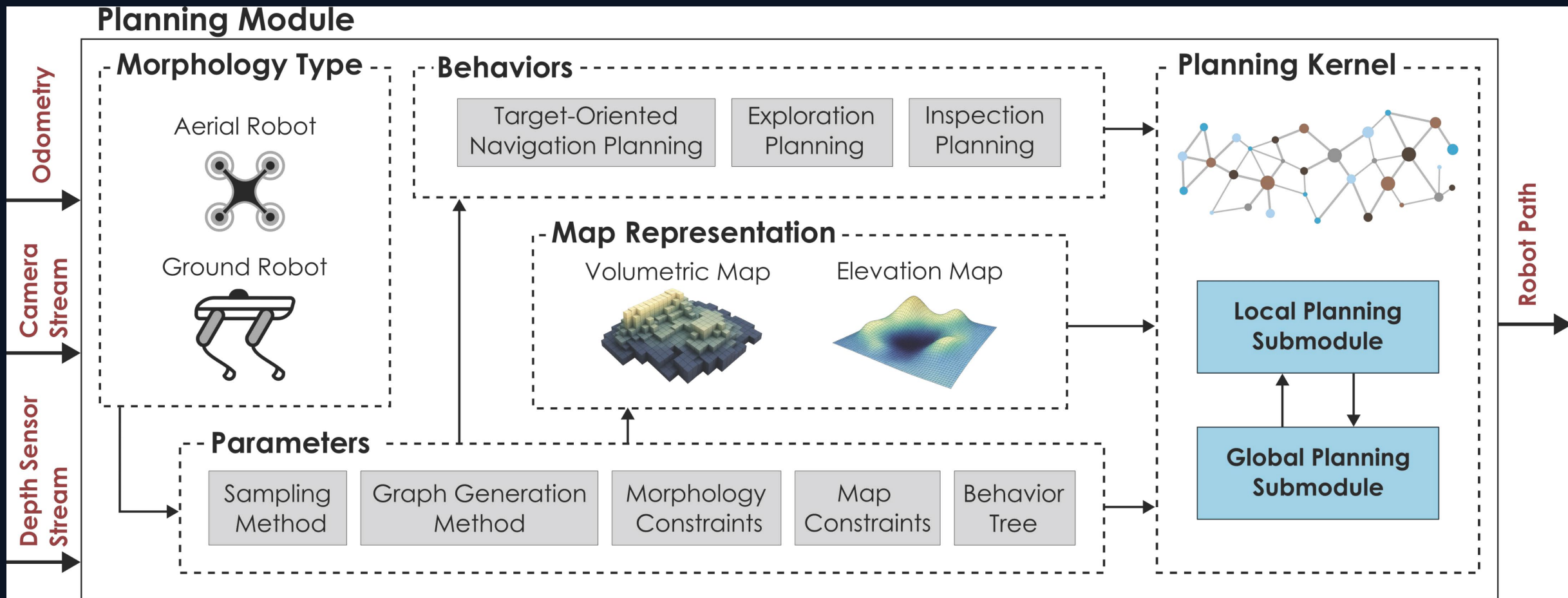
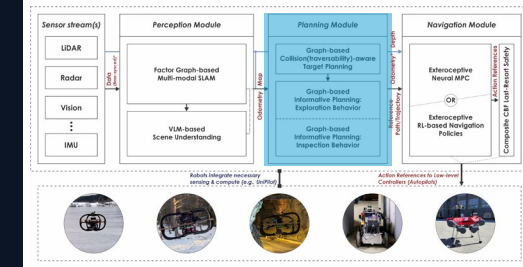
- Local States
- Global States
- Prior Factor
- Preintegrated IMU Factor
- LiDAR Geometric Factor
- Radar Doppler Factor
- Vision Between Factor



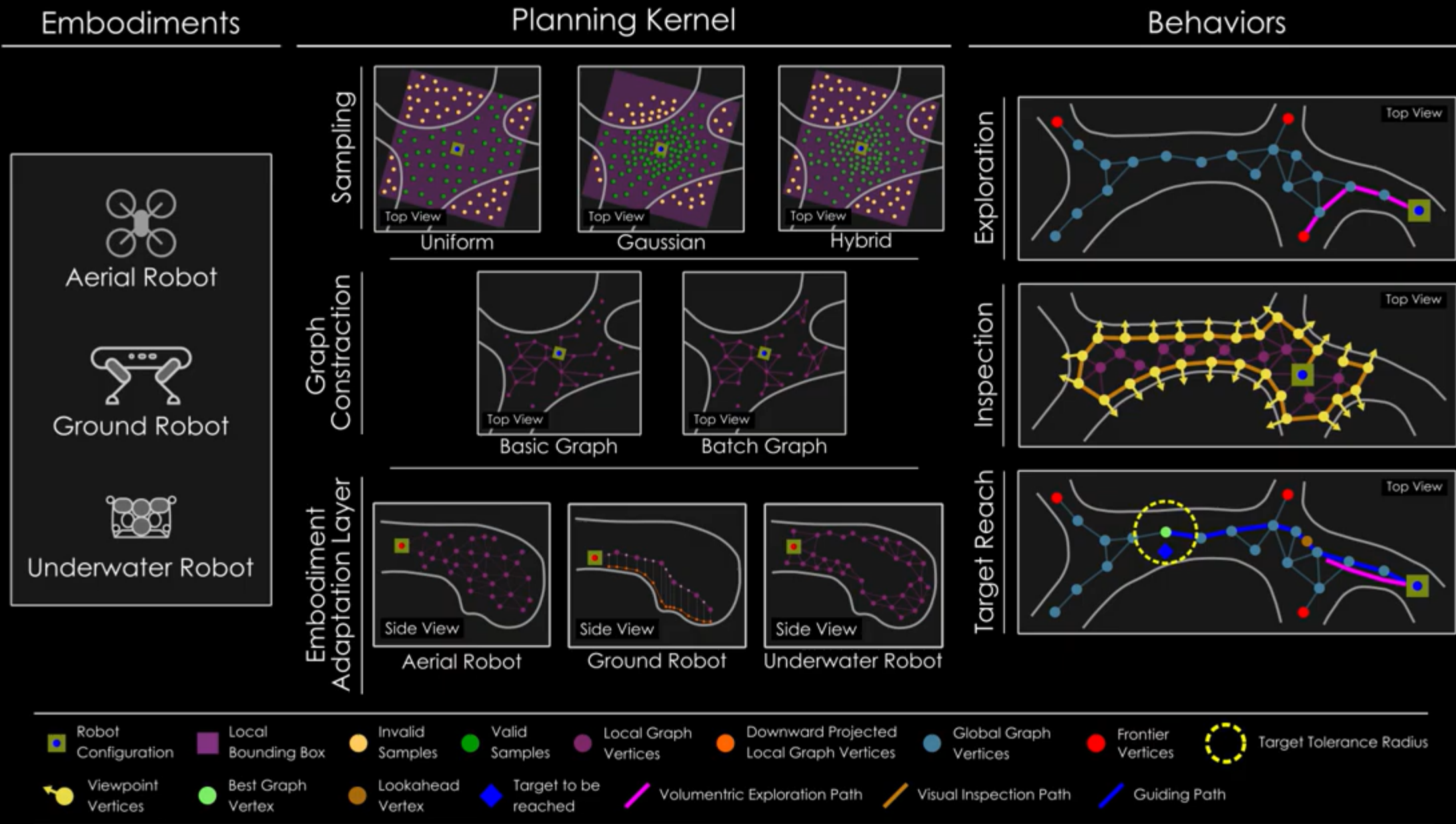
University Campus with Fog - Handheld UniPilot

University Campus Exploration - GR-1

Multi-behavior Planning



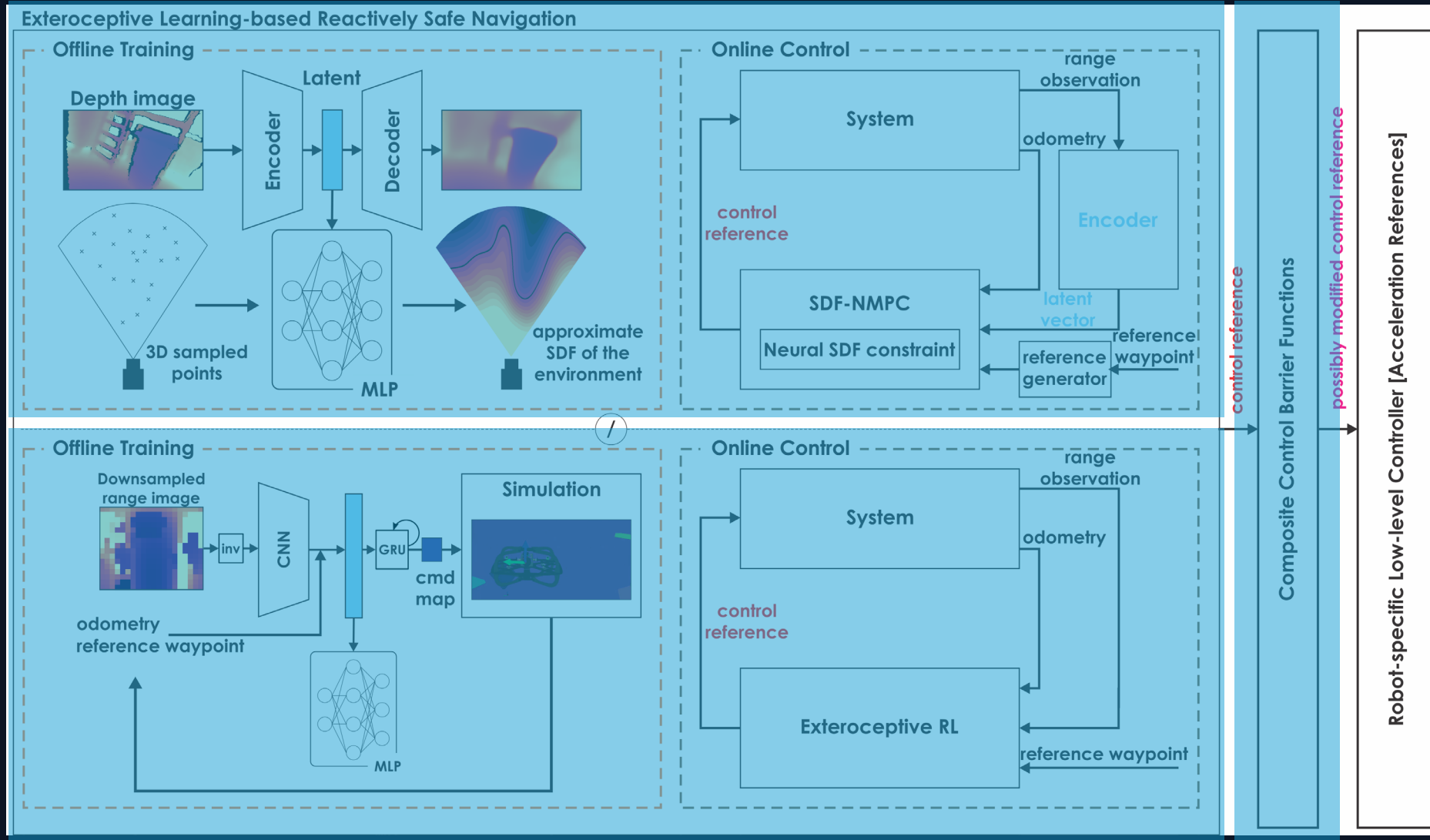
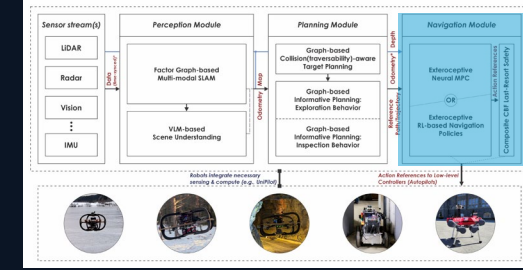
Method Overview



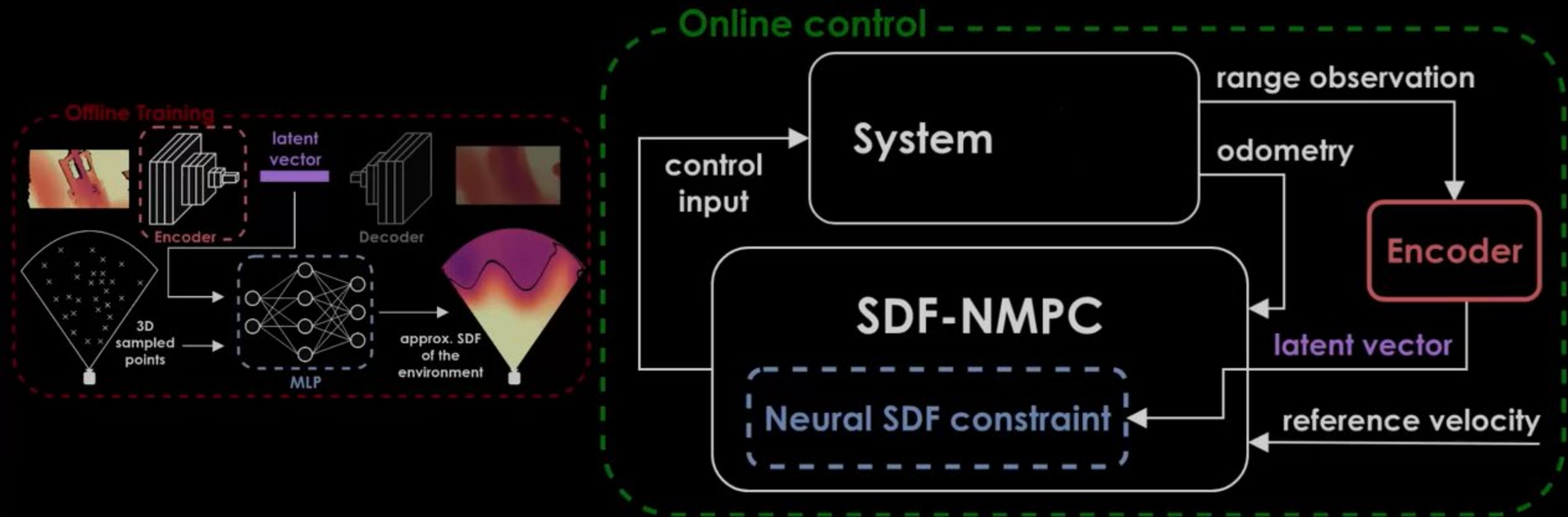
Simulation Studies

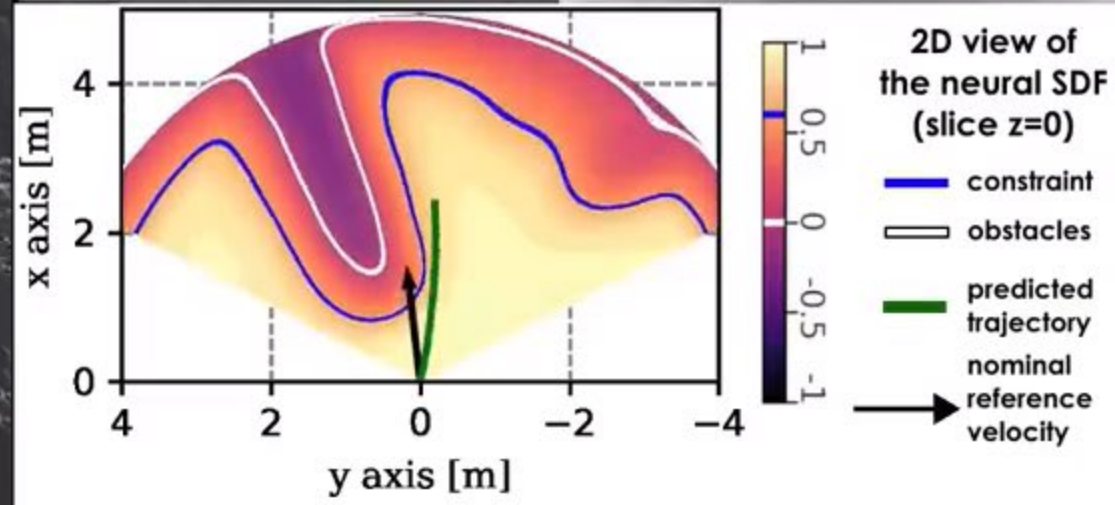
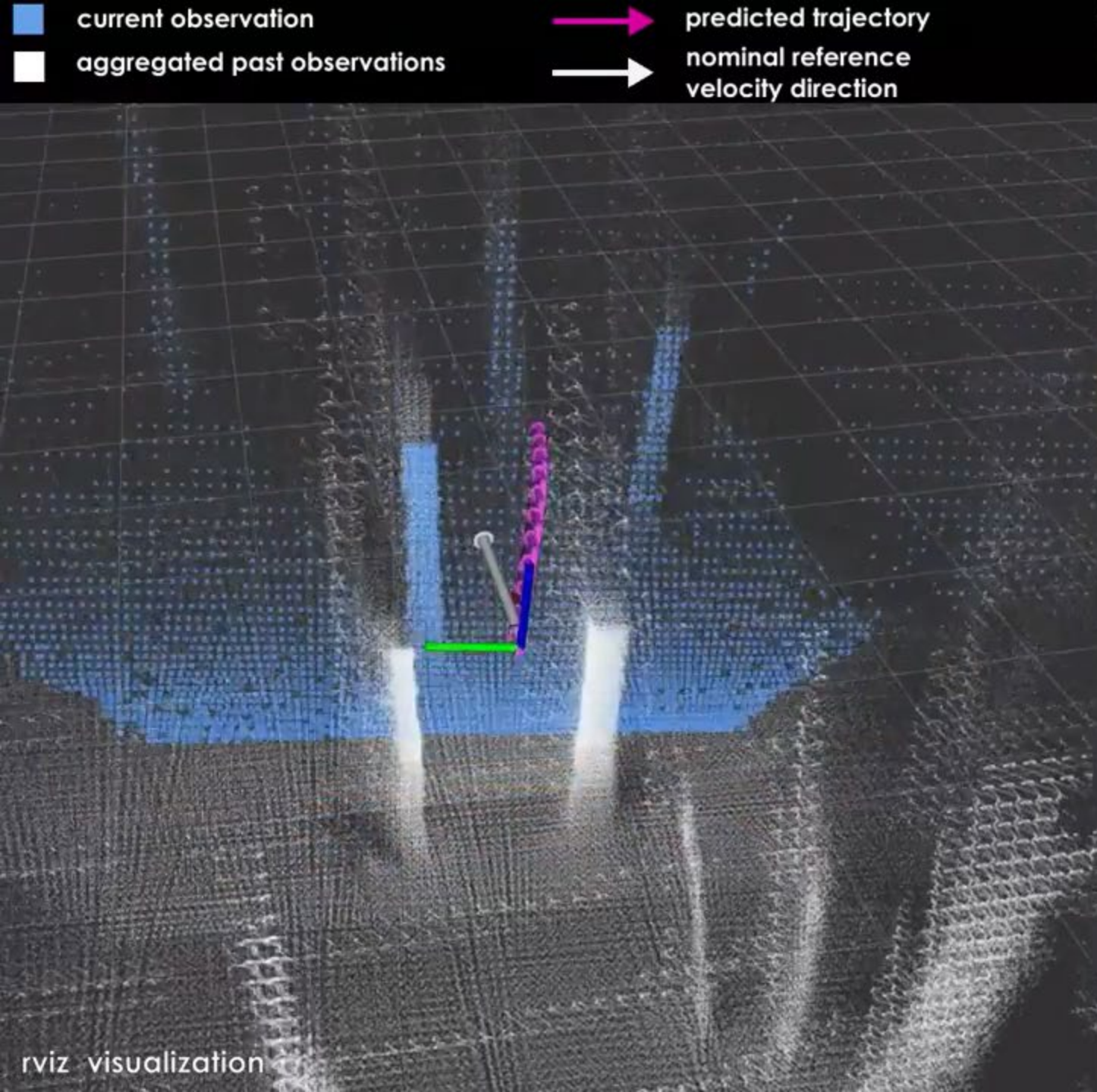
Field Experiments

Multi-layered Navigation

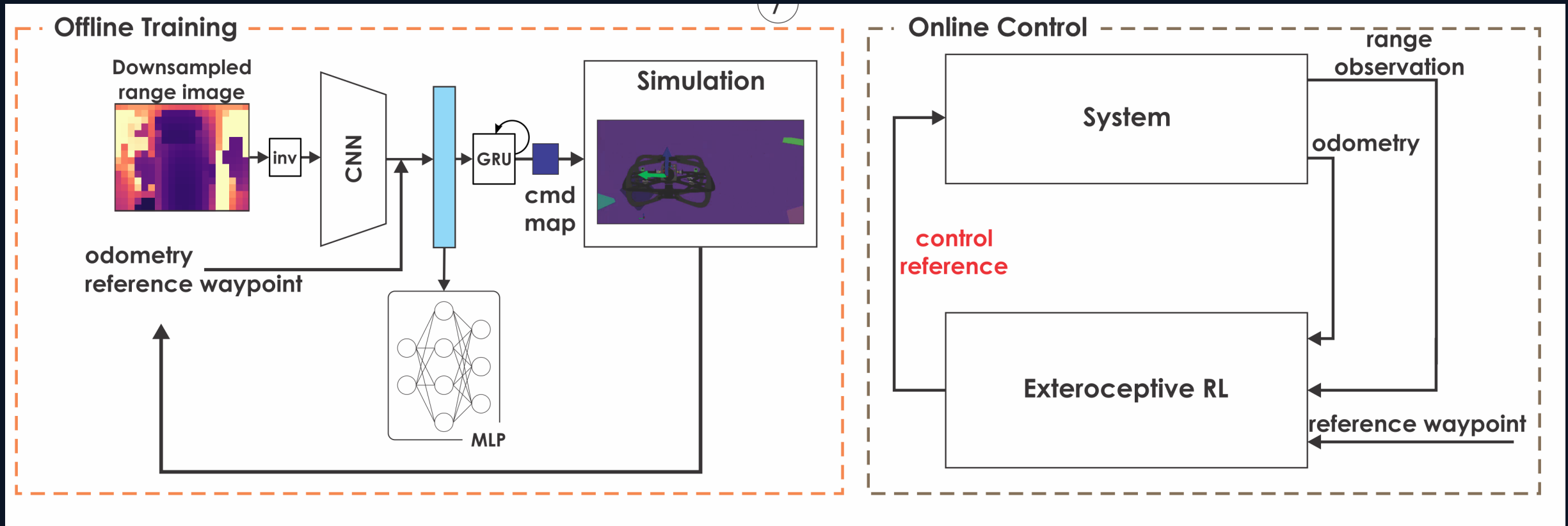


Method Overview



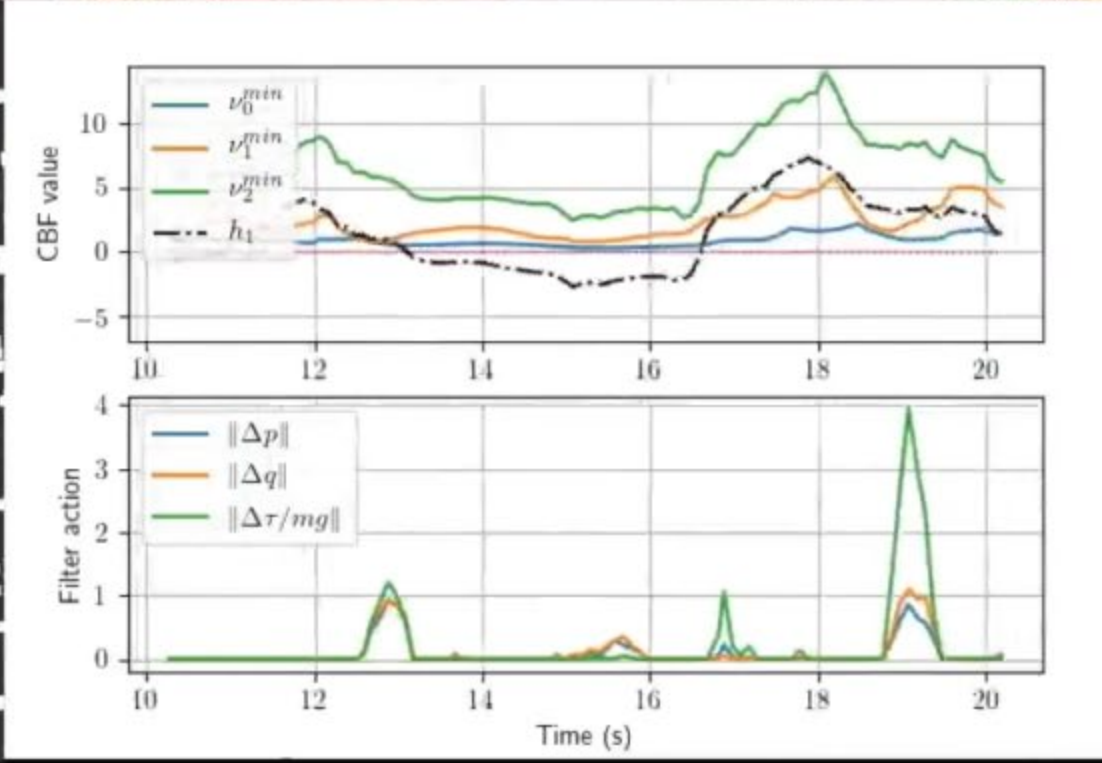
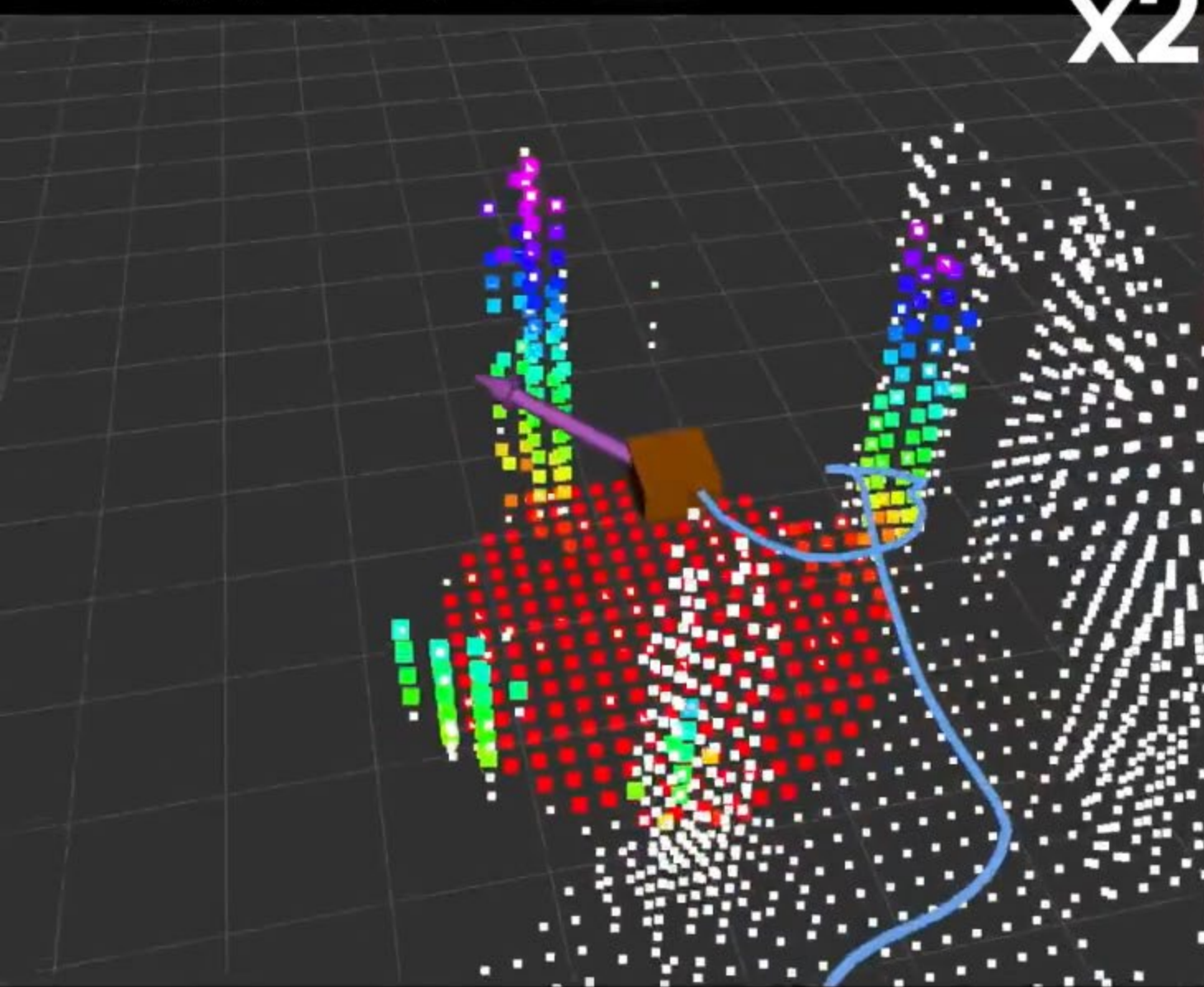


Exteroceptive Deep Reinforcement Learning

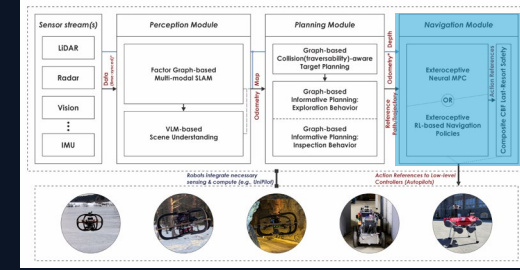


- reference velocity from adversarial policy trying to collide
- current obstacles considered by the safety filter, z-colored
- aggregated history of observations

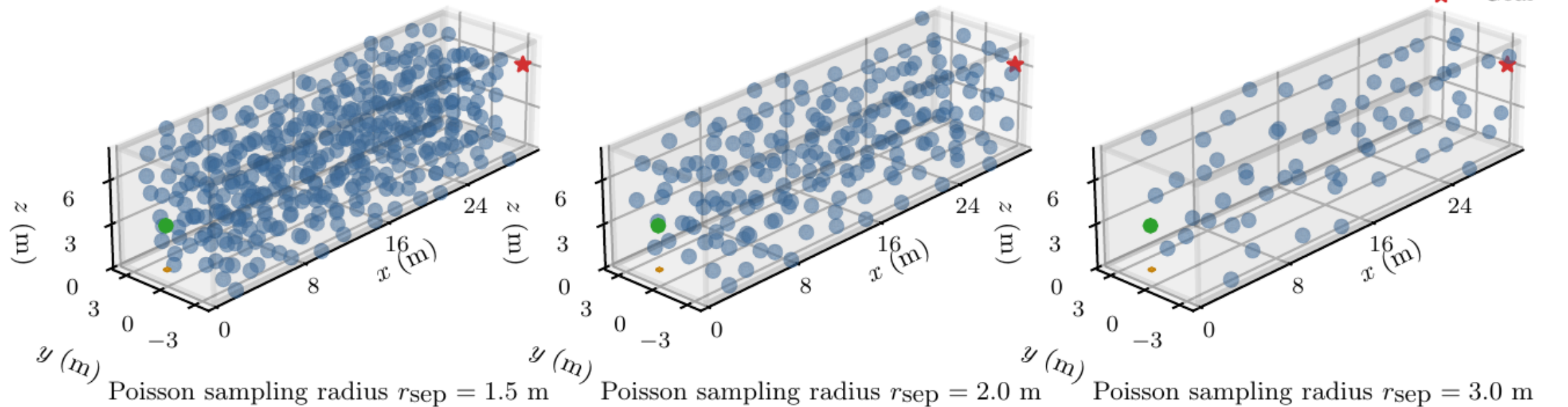
x2



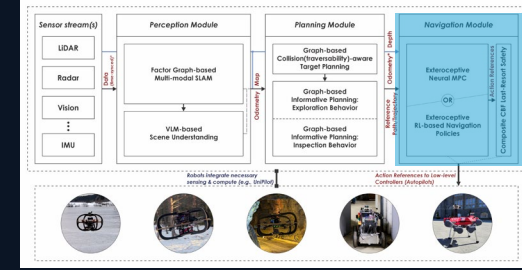
Multi-layered Navigation



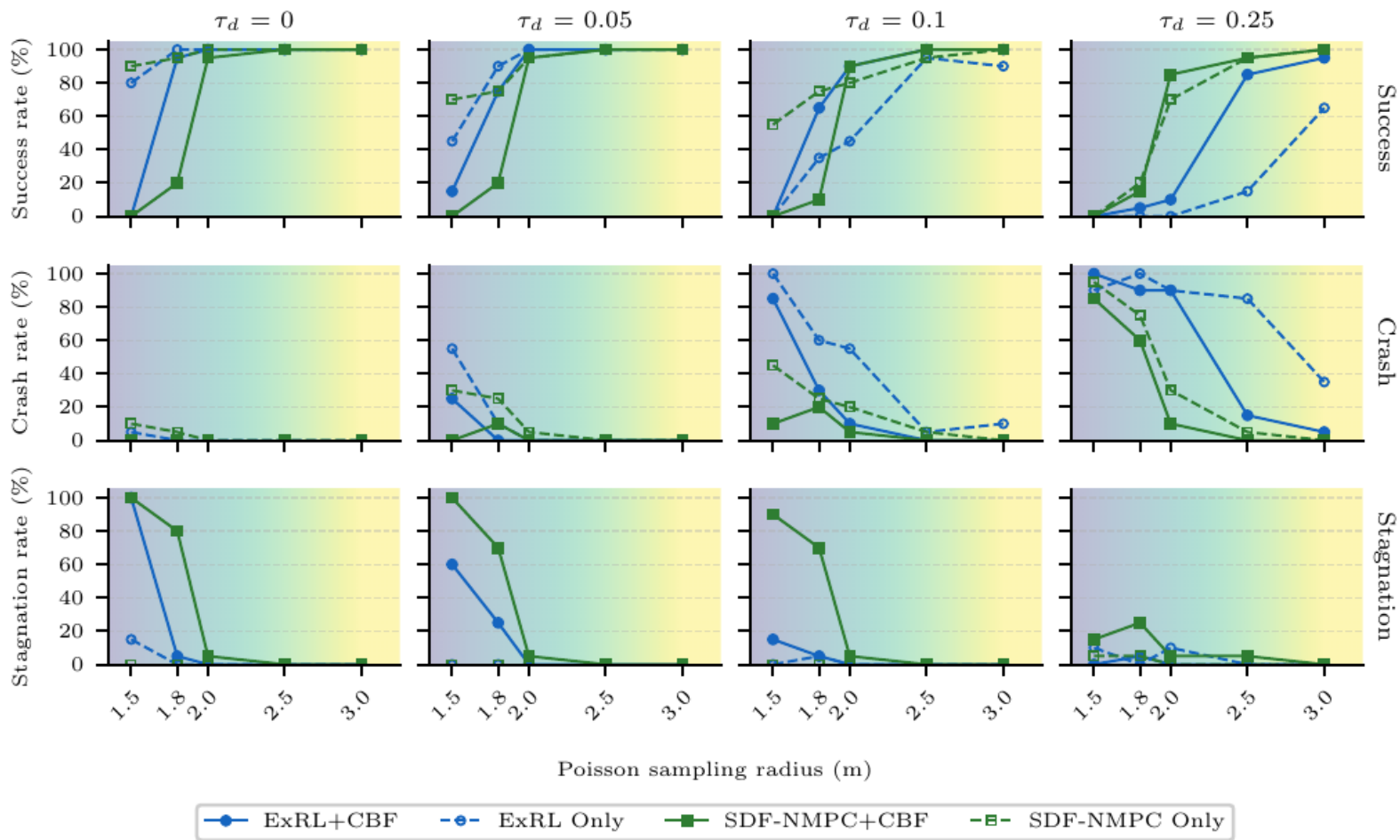
Simulation Environment Densities



Multi-layered Navigation



Ablation Study — Outcome Rates vs. Obstacle Density



Setup

Simulated quadrotor in Gazebo with independent low-pass filter (with time constant τ_d) on commanded accelerations + yaw rate simulating non-instantaneous reference tracking

Result

At realistic $\tau_d \leq 0.1$:

- SDF-NMPC degrades gracefully/ExRL degrades globally due to sim2real mismatch
- CBF substantially reduces collisions however "X"+CBF inherits upstream characteristics

Forest - Waypoint Navigation

Setup

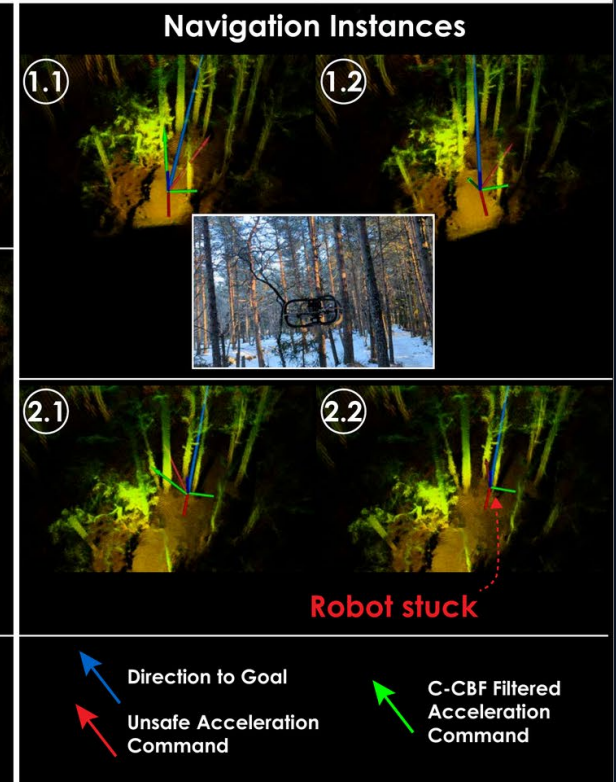
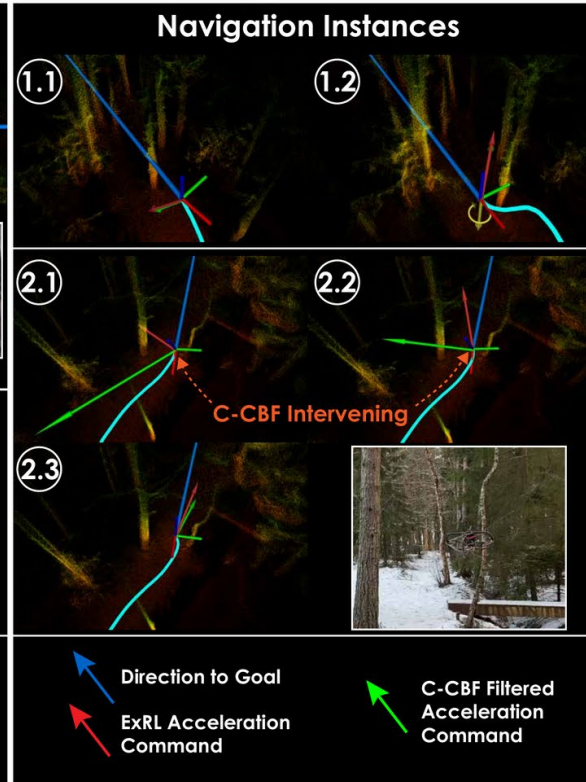
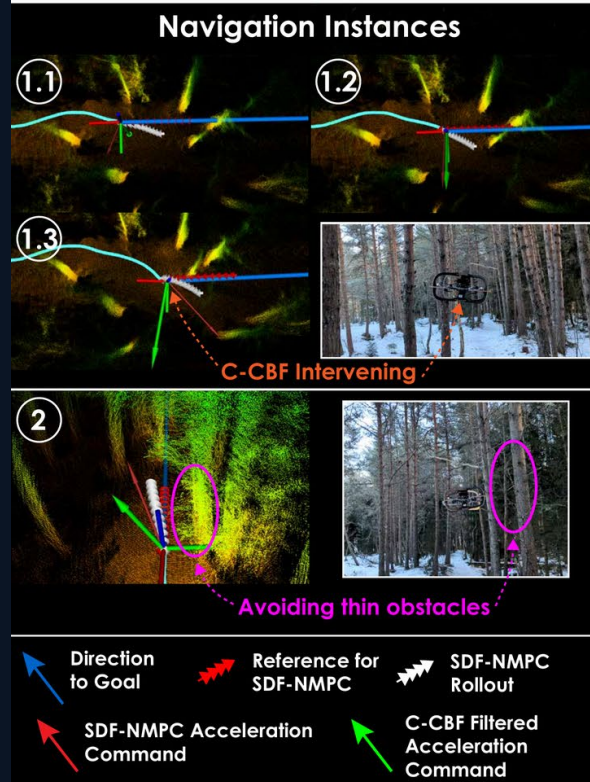
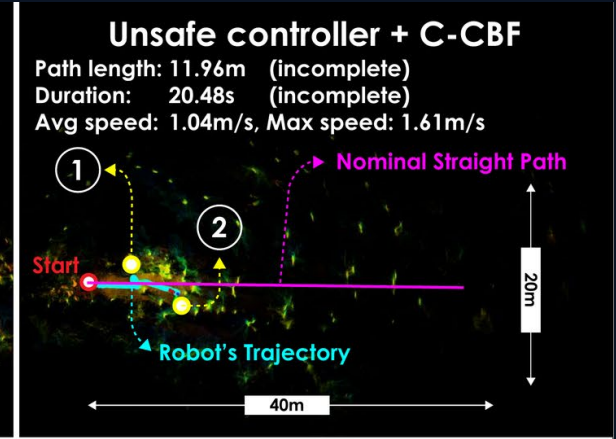
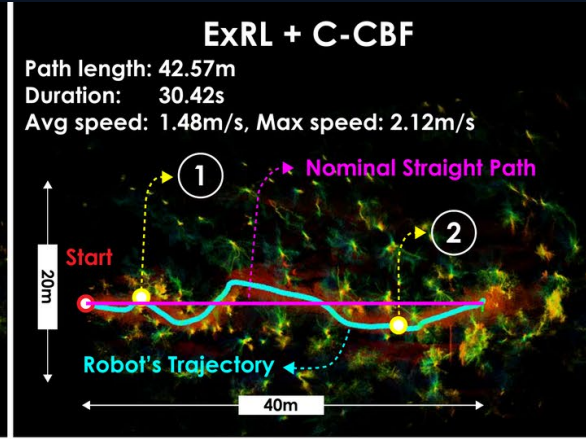
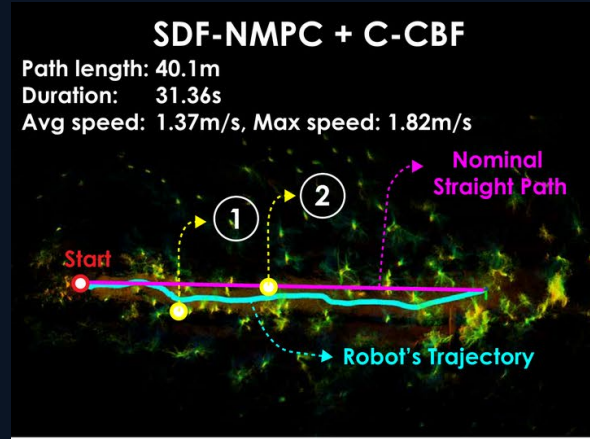
AR2 commanded to a waypoint 38 m ahead through trees and branches — map-based planning disabled.

Result

NMPC+CCBF and ExRL+CCBF both reach the goal. The CCBF-only baseline stays safe but stagnates.

Style note

NMPC hugs the reference; DRL deviates more but moves faster.



Urban - Moving Obstacle

Navigation Module: SDF-NMPC + C-CBF

Urban - Moving Obstacle

Navigation Module: SDF-NMPC + C-CBF

UStack Experimental Studies

Indicative Examples

Full List of Experiments

Purpose	Environment	Robot
SLAM Validation	Fyllingsdal	AR-1
	Runehamar	AR-1
	Campus	Handheld
	Frozen Lake	AR-1
Safety Validation	Forest	AR-2
		AR-2
		AR-2
	Campus	AR-2
		AR-2
Full-Stack Validation	Forest	AR-2
		AR-2
	Mine	AR-2
		AR-2
		GR-1
	Ship	AR-2
	Campus	GR-1

Underground Mine

Navigation Module: SDF-NMPC + C-CBF

Forest

Navigation Module: SDF-NMPC + C-CBF

Ship Cargo Hold

Underground Mine

University Campus

Thank you

Open-source: https://github.com/ntnu-arl/unified_autonomy_stack

Documentation: https://ntnu-arl.github.io/unified_autonomy_stack/

