

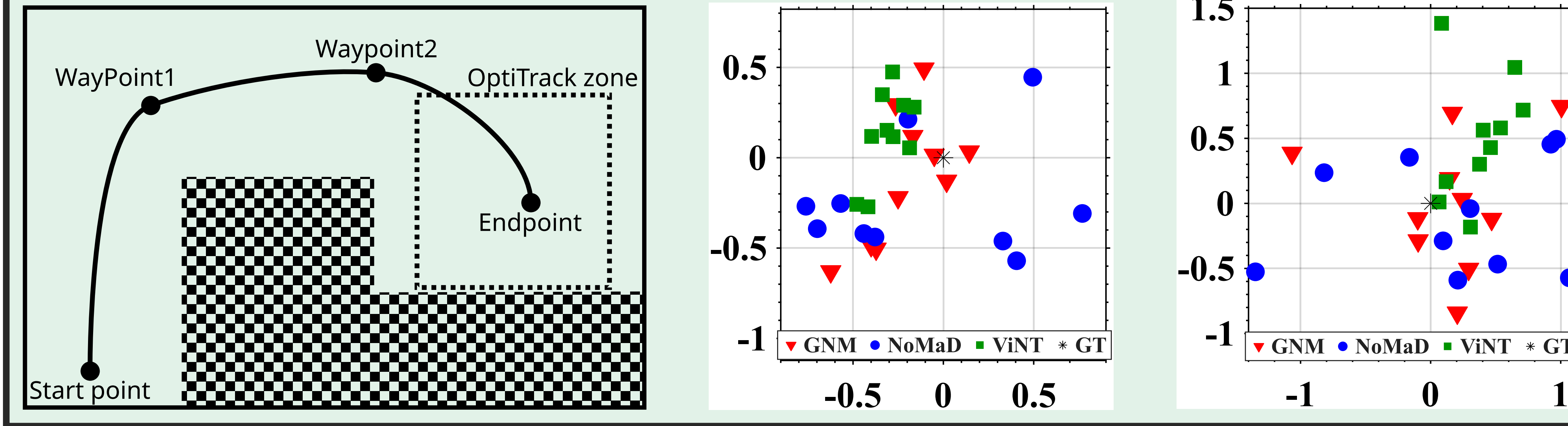
Main aspects

- Create trajectory
- Indoor: AD-R1M and Pioneer 3-AT
- Outdoor: Husarion Panther
- Simulation (Isaac Sim): Clearpath Jackal
- Visual Navigation Transformers (ViNT) [3]
- Navigation with Goal Masked Diffusion (NoMaD) [4]
- General Navigation Model (GNM) [2]

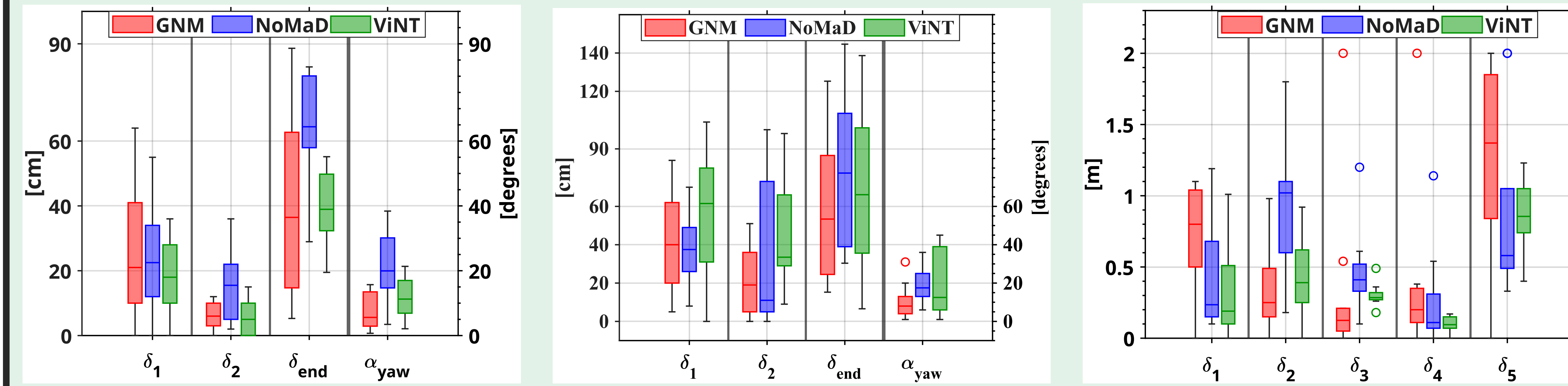
Robot platforms



Indoor (reference path and endpoint distributions)



Displacement error (Indoor & Outdoor)



Numerical results for indoor

Minimum	δ_1	δ_2	δ_{end}	α_{yaw}
GNM	0.0	0.0	5.3	0.7
NoMaD	0.0	2.0	29.0	3.5
ViNT	0.0	0.0	19.5	2.1
Average	δ_1	δ_2	δ_{end}	α_{yaw}
GNM	26.9	6.5	39.1	7.6
NoMaD	23.2	16.6	64.7	20.9
ViNT	17.7	6.0	40.3	11.2
Maximum	δ_1	δ_2	δ_{end}	α_{yaw}
GNM	64.0	12.0	88.6	15.7
NoMaD	55.0	36.1	82.9	38.4
ViNT	36.0	15.0	55.8	21.4

Minimum	δ_1	δ_2	δ_{end}	α_{yaw}
GNM	5	0	15.3	1
NoMaD	8	0	30.4	6
ViNT	0	9	6.6	1
Average	δ_1	δ_2	δ_{end}	α_{yaw}
GNM	43.2	21.6	59.7	10.6
NoMaD	37.6	32.4	79.4	19
ViNT	54.9	43	68.5	18.4
Maximum	δ_1	δ_2	δ_{end}	α_{yaw}
GNM	84	51	125.3	31
NoMaD	70	100	144.6	36
ViNT	104	98	138.7	45

	Memory [GB]	Inference [sec]	Power [W]
GNM	0.63	0.28	11.4
NoMaD	1.29	1.16	12.7
ViNT	0.95	0.65	12.1

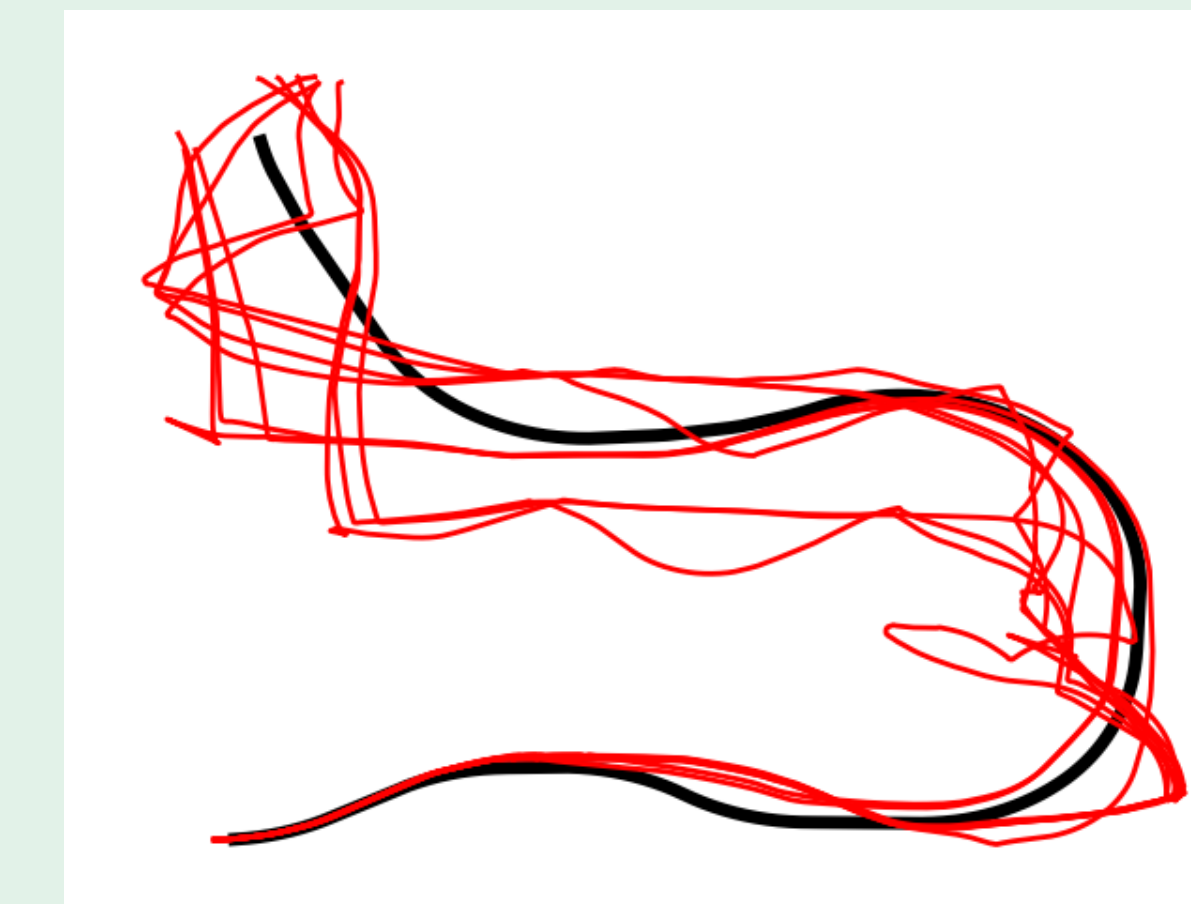
References

This paper was financially supported by the Romanian National Authority for Scientific Research, project nr. PN-IV-P7-7.1-PTE-2024-0105 and by the Technical University of Cluj-Napoca for PhD students.

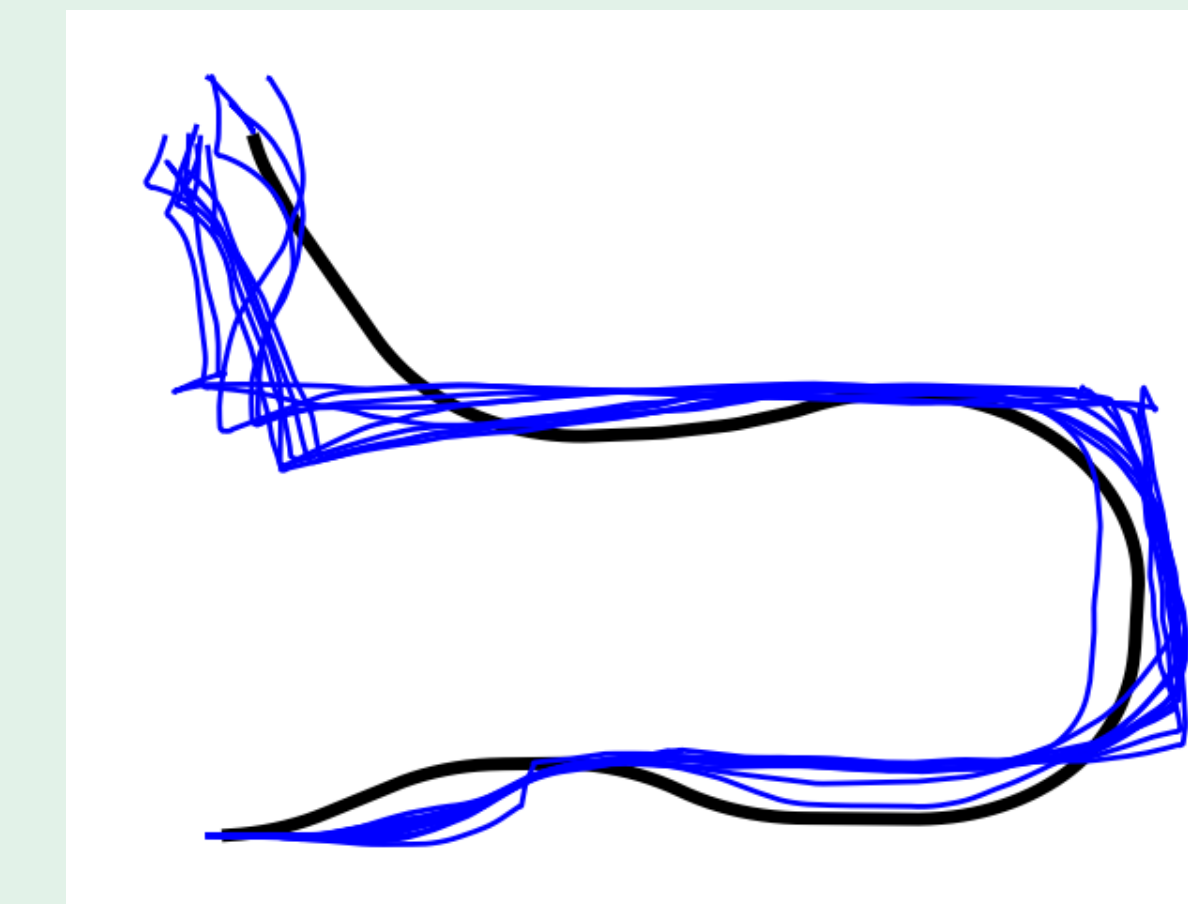
This paper was published in [1]. For more information: <https://molnarszilard.github.io/> or <http://rocon.utcluj.ro/>.

- [1] Szilárd Molnár and Levente Tamás. Vision-based Navigation Models for Autonomous Rovers: Experimental Analysis and Comparison. In *2026 12th International Conference on Control, Decision and Information Technologies (CoDIT)*, pages 2223–2228, 2026.
- [2] Dhruv Shah, Ajay Sridhar, Arjun Bhorkar, Noriaki Hirose, and Sergey Levine. GNM: A General Navigation Model to Drive Any Robot. In *IEEE International Conference on Robotics and Automation, ICRA 2023, London, UK, May 29 - June 2, 2023*, pages 7226–7233. IEEE, 2023.
- [3] Dhruv Shah, Ajay Sridhar, Nitish Dashora, Kyle Stachowicz, et al. ViNT: A Foundation Model for Visual Navigation. In Jie Tan, Marc Toussaint, and Kourosh Darvish, editors, *Conference on Robot Learning, CoRL 2023, 6-9 November 2023, Atlanta, GA, USA*, volume 229 of *Proceedings of Machine Learning Research*, pages 711–733. PMLR, 2023.
- [4] Ajay Sridhar, Dhruv Shah, Catherine Glossop, and Sergey Levine. NoMaD: Goal Masked Diffusion Policies for Navigation and Exploration. In *IEEE International Conference on Robotics and Automation, ICRA 2024, Yokohama, Japan, May 13-17, 2024*, pages 63–70. IEEE, 2024.

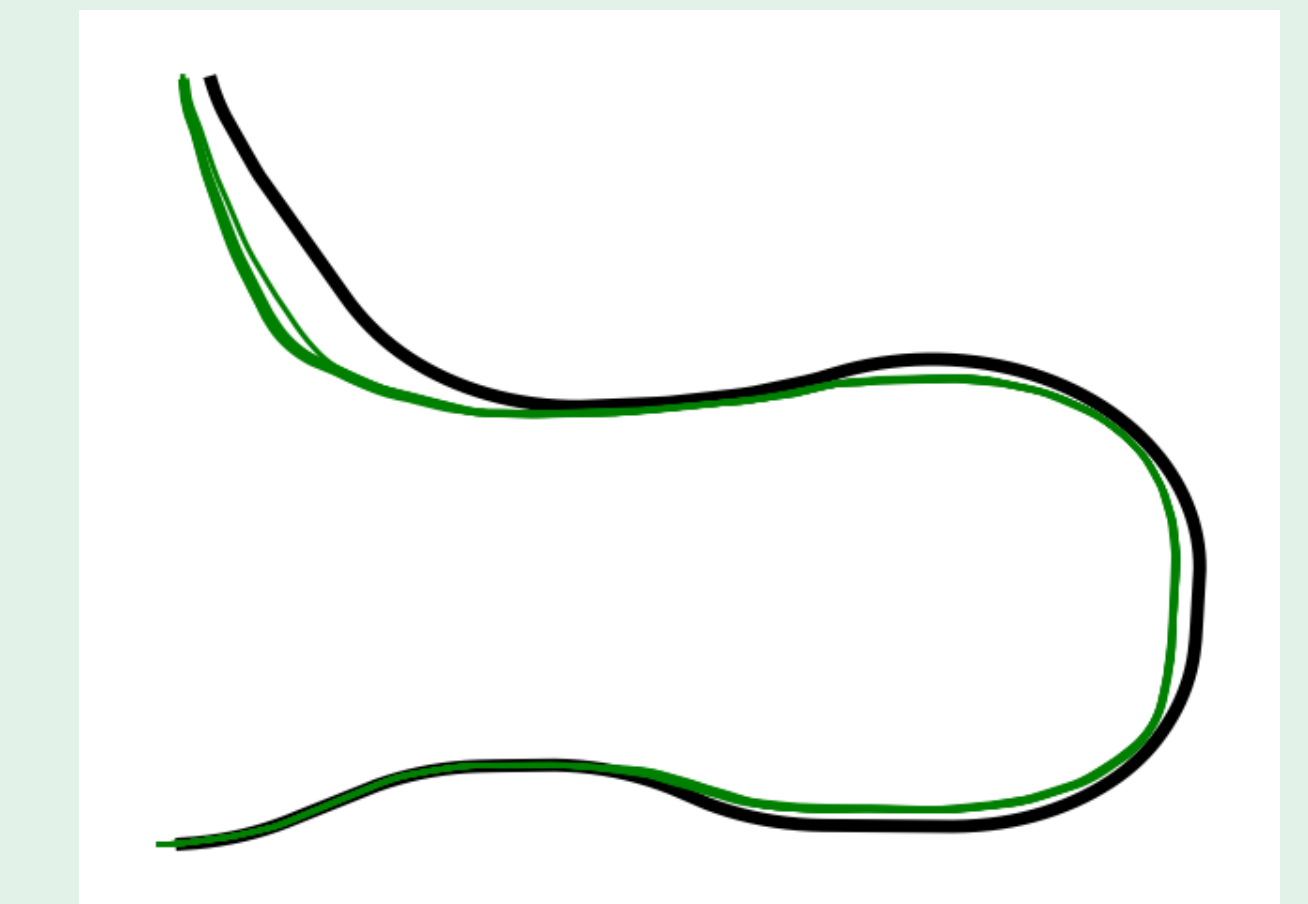
Simulation path overlap



GNM



NoMaD



ViNT