Extending Operational Mission Lifetimes of Free-Flying Space Robots via Hypernetwork-Based Multi-Task GNC Controller

Ricard Marsal I Castan[1], Cédric Pradalier[2], and Miguel Olivares-Mendez[1]

***Abstract*— Space free-flying robots must sustain diverse, evolving mission profiles, from rendezvous and docking to inspection and emergency stabilization, across operational lifetimes spanning years or decades. The prevailing approach maintains a library of purpose-built controllers, one per mission phase, which is architecturally brittle, memory-intensive, and unable to adapt to unanticipated conditions. We present HYPER-GNC, a hypernetwork-based Multi-Task Reinforcement Learning (MTRL) framework that lets a single, compact controller master a full suite of Guidance, Navigation, and Control (GNC) tasks without retraining. By projecting physics informed semantic embeddings onto a continuous task manifold and mapping them to policy weights via a hypernetwork, HYPER-GNC resolves conflicting gradient dynamics while enabling zero-shot task composition: synthesizing novel mission configurations at inference by combining learned physical priorities, without any retraining or gradient steps. We validate on JAXA's Int-Ball2 in 6-DOF simulation and on a physical satellite emulator performing stabilization, docking, velocity tracking, inspection, and obstacle navigation. HYPER-GNC consistently outperforms single-task experts and multi-task baselines.**

I. Introduction

Free-flying robots aboard the ISS must dock, inspect equipment, track crew, and navigate cluttered corridors over lifetimes of years with no possibility of controller re-upload. Today's approach, a library of purpose-built controllers, one per phase, scales poorly, wastes on-board memory, and cannot adapt to unanticipated scenarios.

Reinforcement Learning (RL) offers a data-driven alternative [1], [2], [3], but single-task policies fail under slight objective shifts [4]. Multi-Task RL (MTRL) trains a shared policy but suffers from conflicting gradients [5] and discrete task IDs that prevent interpolation. Hypernetworks [6] generate task-specific weights from a context vector. HYPER-GNC extends this to 6-DOF Int-Ball2 GNC with a *physics-informed continuous* semantic manifold. Contributions: (1) single-model policy mastering four orbital GNC profiles simultaneously; (2) zero-shot task composition to unseen profiles; (3) sim-to-real transfer on physical hardware.

II. HYPER-GNC: Framework

Physics-informed semantic embeddings. Each mission profile is encoded as:

$$e_k = [\alpha_{\text{pos}},\ \alpha_{\text{att}},\ \alpha_{\text{vel}},\ \alpha_{\text{brake}},\ \alpha_{\text{safe}}] \in [0,1]^5, \qquad (1)$$

where each component maps to a classical GNC objective [7], [8]: position/attitude error, velocity tracking, braking, and obstacle repulsion. Unlike learned codes, e_k is set by the mission designer, interpretable to operations teams. Training samples e_k uniformly over $[0,1]^5$, acting as parameter-space data augmentation over all blended priorities. Zero-shot composition is immediate: setting $\alpha_{\text{safe}}{=}1$ on Docking gives *Docking with Obstacles*; dropping α_{att} gives *Point Navigation*, no retraining required.

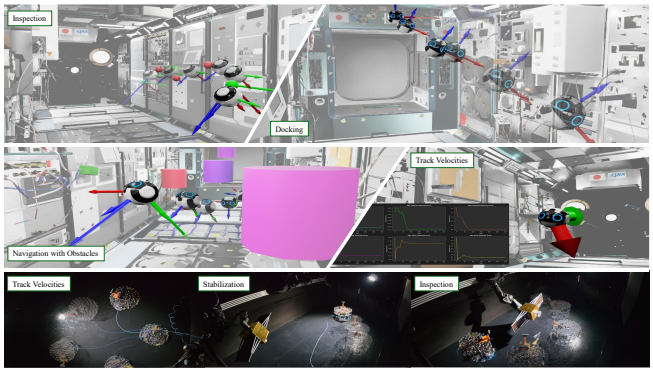

Fig. 1. **Overview.** HYPER-GNC on Int-Ball2 (simulation, top) and physical satellite emulator (bottom).

Hypernetwork architecture. A shared actor-critic is weight-modulated per task via:

$$\mathbf{W}_{\text{dyn}}(e_k) = \mathbf{W}_{\text{base}}(1 + \Psi_{\text{s}}^{W}) + \Psi_{\text{a}}^{W}, \qquad (2)$$

$$\Psi(e_k) = \mathbf{q}(e_k) \times \mathbf{k}(e_k)/\sqrt{d_{\text{out}}}, \qquad (3)$$

where $\mathbf{q}, \mathbf{k} \in \mathbb{R}^{32}$ are hypernetwork outputs. Rank-1 factorization gives a $16\times$ parameter reduction vs. full-rank, analogous to LoRA [9], essential for flight hardware [10]. Keys and Queries initialize at $\sigma{=}10^{-4}$ with scale offset $+1$, so training begins from identity modulation [11].

Training. JAXA's Int-Ball2 [12] (6-DOF, continuous) is the simulation platform; a 5 kg Floating Platform (3-DOF, binary thrusters) serves as physical testbed. Four tasks train simultaneously in 4,096 Isaac Lab [13] environments with PPO [14] and domain randomization [15] over CoM, mass, and wrenches.

III. Experimental Validation

To demonstrate HYPER-GNC's suitability for long-duration space missions, we evaluate its ability to master multiple operational phases simultaneously and adapt to unforeseen scenarios without structural changes or retraining. We compare against **Single-Task RL**, standard **MTRL**, **PCGrad** [16], and **Multi-Critic** [17].

Simulation. HYPER-GNC outperforms or matches baselines across trained profiles (Table II). Crucially, this is achieved with a single compact model, avoiding the memory bloat of

[1]University of Luxembourg
[2]Georgia Tech

TABLE I

SIM-TO-REAL RESULTS. PERFORMANCE OF HYPER-GNC WITH AND WITHOUT DOMAIN RANDOMIZATION ON THE PHYSICAL PLATFORM.

Methods	*Docking* $e_p[m]\downarrow$	$e_o[rad]\downarrow$	*Track Velocities* $e_v[m\,s^{-1}]\downarrow$	$e_\omega[rad\,s^{-1}]\downarrow$	*Inspection* $e_o[rad]\downarrow$	$SR[\%]\uparrow$	*Navigation with Obstacles* $e_p[m]\downarrow$	*Stabilization* $e_v[m\,s^{-1}]\downarrow$	$e_\omega[rad\,s^{-1}]\downarrow$	*Point Navigation* $e_p[m]\downarrow$
HYPER-GNC w/o Rand	0.092 (±0.039)	0.198 (±0.172)	0.183 (±0.029)	**0.152** (±0.033)	0.125 (±0.006)	**100**	2.136 (±0.636)	0.135 (±0.125)	0.155 (±0.147)	0.093 (±0.005)
HYPER-GNC	**0.028** (±0.019)	**0.128** (±0.011)	**0.169** (±0.015)	0.157 (±0.107)	**0.082** (±0.021)	**100**	**0.087** (±0.006)	**0.077** (±0.043)	**0.110** (±0.066)	**0.094** (±0.004)

TABLE II

INT-BALL2 FOUR GNC TASKS COMPARISON PERFORMANCE.

Methods	*Docking* $e_p[m]\downarrow$	$e_o[rad]\downarrow$	*Track Velocities* $e_v[m\,s^{-1}]\downarrow$	$e_\omega[rad\,s^{-1}]\downarrow$	*Inspection* $e_o[rad]\downarrow$	$SR[\%]\uparrow$	*Navigation with Obstacles* $e_p[m]\downarrow$
Single-Task RL	0.017 (±0.008)	0.041 (±0.022)	0.106 (±0.053)	0.178 (±0.069)	0.622 (±0.151)	85.46	0.056 (±0.025)
Multi-Task RL	0.036 (±0.020)	0.126 (±0.067)	0.233 (±0.094)	0.386 (±0.033)	0.661 (±0.156)	62.50	0.115 (±0.076)
PCGrad	0.032 (±0.014)	0.136 (±0.063)	0.281 (±0.110)	0.403 (±0.032)	**0.400** (±0.491)	60.35	0.113 (±0.091)
Multi-Critic	0.026 (±0.005)	0.046 (±0.027)	0.187 (±0.088)	0.246 (±0.065)	0.585 (±0.136)	85.00	0.060 (±0.034)
HYPER-GNC (ours)	**0.005** (±0.003)	**0.011** (±0.007)	**0.086** (±0.041)	**0.157** (±0.052)	0.597 (±0.159)	**91.25**	**0.021** (±0.009)

TABLE III

INT-BALL2 PERFORMANCE ON COMPOSITION TASKS.

Methods	*Docking with Obstacles* $e_p[m]\downarrow$	$e_o[rad]\downarrow$	*Stabilization* $e_v[m\,s^{-1}]\downarrow$	$e_\omega[rad\,s^{-1}]\downarrow$	*Point Navigation* $e_p[m]\downarrow$
Single-Task RL	**0.057** (±0.076)	**0.141** (±0.099)	0.001 (±0.002)	0.037 (±0.076)	0.019 (±0.009)
Multi-Task RL	0.079 (±0.235)	0.210 (±0.237)	0.425 (±0.177)	0.680 (±0.056)	0.037 (±0.024)
Multi-Task RL SemEmb	0.081 (±0.184)	0.617 (±0.339)	0.028 (±0.017)	0.484 (±0.289)	0.036 (±0.029)
PCGrad	0.080 (±0.266)	0.210 (±0.249)	0.055 (±0.024)	0.952 (±0.392)	0.032 (±0.014)
PCGrad SemEmb	0.076 (±0.151)	0.799 (±0.543)	0.010 (±0.017)	0.110 (±0.169)	0.045 (±0.048)
Multi-Critic	0.092 (±0.292)	0.206 (±0.341)	**0.000** (±0.001)	**0.001** (±0.020)	0.026 (±0.005)
Multi-Critic SemEmb	0.076 (±0.217)	0.705 (±0.478)	0.003 (±0.004)	0.046 (±0.058)	0.041 (±0.019)
HYPER-GNC (ours)	0.076 (±0.262)	0.182 (±0.233)	0.001 (±0.001)	0.017 (±0.011)	0.017 (±0.010)

maintaining multiple specialized controllers over a satellite's lifespan.

Zero-shot task composition (Table III). Long-life adaptation necessitates handling emergent mission requirements. HYPER-GNC demonstrates zero-shot capabilities on entirely novel profiles blending different objectives (e.g., Stabilization, Point Navigation), matching or exceeding retrained specialists without any gradient updates.

Sim-to-Real. Reliable physical deployment is the ultimate test for long-term orbital autonomy. On our physical emulator (Table I), HYPER-GNC successfully bridges the reality gap, proving its viability for persistent operations.

IV. CONCLUSION

HYPER-GNC enables a single compact GNC controller to span the full mission lifecycle of an orbital free-flying robot. Physics-informed semantic embeddings on a continuous manifold support zero-shot task composition without retraining. A mission operator adjusts only a five-dimensional vector to switch between any mission phase, no re-upload, no ground intervention required.

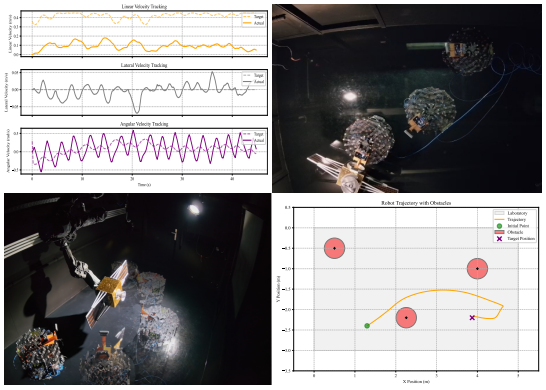

Fig. 2. **HYPER-GNC in the real world.** From left to right, top to bottom, target and actual velocities of the Track Velocities task, docking into a satellite, inspection of a satellite, and navigation through obstacles.

REFERENCES

[1] D. Izzo, M. Märtens, and B. Pan, "A survey on artificial intelligence trends in spacecraft guidance dynamics and control," *Astrodynamics*, vol. 3, no. 4, pp. 287–299, 2019.

[2] B. Gaudet, R. Linares, and R. Furfaro, "Deep reinforcement learning for six degree-of-freedom planetary landing," *Advances in Space Research*, vol. 65, no. 7, pp. 1723–1741, 2020.

[3] M. El-Hariry, A. Richard, V. Muralidharan, M. Geist, and M. Olivares-Mendez, "DRIFT: Deep reinforcement learning for intelligent floating platforms trajectories," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2024, pp. 14 034–14 041.

[4] K. Cobbe, O. Klimov, C. Hesse, T. Kim, and J. Schulman, "Quantifying generalization in reinforcement learning," in *Proceedings of the 36th International Conference on Machine Learning*. PMLR, 2019, pp. 1282–1289.

[5] T. Yu, S. Kumar, A. Gupta, S. Levine, K. Hausman, and C. Finn, "Gradient surgery for multi-task learning," *Advances in Neural Information Processing Systems*, vol. 33, pp. 5824–5836, 2020.

[6] D. Ha, A. Dai, and Q. V. Le, "Hypernetworks," *arXiv preprint arXiv:1609.09106*, 2016.

[7] B. Wie, *Space Vehicle Dynamics and Control*, 2nd ed. AIAA Education Series, 2008.

[8] W. Fehse, *Automated Rendezvous and Docking of Spacecraft*. Cambridge University Press, 2003.

[9] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, and W. Chen, "LoRA: Low-rank adaptation of large language models," 2021.

[10] G. Giuffrida *et al.*, "The phi-sat-1 mission: The first on-board deep neural network demonstrator for satellite earth observation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–14, 2022.

[11] T. He, Z. Zhang, H. Zhang, Z. Zhang, J. Xie, and M. Li, "Bag of tricks for image classification with convolutional neural networks," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[12] S. P. Yamaguchi, T. Yamamoto *et al.*, "Int-Ball2: ISS JEM internal camera robot with increased degree of autonomy — design and initial checkout," *IEEE Robotics and Automation Magazine*, 2024.

[13] M. Mittal, P. Roth, J. Tigue, A. Richard *et al.*, "Isaac lab: A gpu-accelerated simulation framework for multi-modal robot learning," *arXiv preprint arXiv:2511.04831*, 2025.

[14] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," 2017.

[15] J. Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, and P. Abbeel, "Domain randomization for transferring deep neural networks from simulation to the real world," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, pp. 23–30.

[16] T. Yu, S. Kumar, A. Gupta, S. Levine, K. Hausman, and C. Finn, "Gradient surgery for multi-task learning," 2020.

[17] S. Mysore, G. Cheng, Y. Zhao, K. Saenko, and M. Wu, "Multi-critic actor learning: Teaching RL policies to act with style," in *International Conference on Learning Representations*, 2022.