Effects of Lookahead and Communication Dynamics on Multi-Agent Potential-Field Rover Navigation

Timothy Flavin
University of Tulsa
Tulsa, United States
Timmy-Flavin@utulsa.edu

Joe Shymanski
University of Tulsa
Tulsa, United States
Joe-Shymanski@utulsa.edu

Sandip Sen
University of Tulsa
Tulsa, United States
Sandip-Sen@utulsa.edu

ABSTRACT

In this work, we investigate how predictive lookahead and communication impact multi-agent stochastic artificial potential field navigation (SPF) for rovers. We choose SPF to balance exploration, collision avoidance, and hazard evasion using terrain data derived from real HiRISE Mars terrain data. We empirically evaluate how varying lookahead horizons, prior belief map inaccuracies, and communication frequency impact effective spatial coverage. Our key results show that while minimal communication ($p = 0.05$) successfully bounds lookahead prediction error, naive SPF controllers struggle to use predictive information effectively. Additionally, we found that lookahead combined with environment stochasticity act as an effective local-minimum escape strategy when policy noise is not an option. Ultimately, this work provides a computationally efficient engine and first benchmark for testing multi-agent coverage on real world terrain. Our results suggest lookahead with sparse communication is sufficient for teammate modeling, but taking advantage of that information is beyond simple SPF controllers.

KEYWORDS

Stochastic Potential Field Navigation, Simulated Annealing, Multi-Rover Navigation

ACM Reference Format:
Timothy Flavin, Joe Shymanski, and Sandip Sen. 2026. Effects of Lookahead and Communication Dynamics on Multi-Agent Potential-Field Rover Navigation. In *Appears at the International Workshop on Autonomous Agents and Multi-Agent Systems for Space Applications (MASSpace-26). Held as part of the Workshops at the 25th International Conference on Autonomous Agents and Multiagent Systems., Paphos, Cyprus, May 2026*, IFAAMAS, 5 pages.

1 INTRODUCTION

Future space missions increasingly envision teams of autonomous robotic assets operating in hazardous and partially known environments [6, 15, 19]. Such distributed robotic systems must operate under strict communication constraints, uncertain terrain, and limited sensing. Centralized coordination is often infeasible due to latency and bandwidth limits, making lightweight decentralized decision-making a critical enabler of scalable space exploration.

A fundamental task in such missions is spatial coverage: multiple rovers must efficiently explore and map an unknown region while avoiding hazardous terrain and minimizing redundant traversal [5, 21]. Under partial observability, agents must balance exploration and danger avoidance, anticipate teammate movements to prevent overlap, and compensate for stale or noisy information. In realistic planetary scenarios, prior terrain maps derived from orbital imagery may also contain systematic errors.

Artificial Potential Fields (APFs) provide a computationally lightweight and interpretable mechanism for decentralized navigation [12]. By superimposing attraction and repulsion forces derived from spatial features, agents compute continuous control actions without requiring large neural models. However, classical APF approaches largely assume reactive behavior [12, 17], leaving open questions about the impact of forward prediction and communication.

In this work, we investigate how lookahead simulation and communication constraints interact in multi-agent potential-field rover navigation. We model four rovers exploring a hazardous terrain grid derived from real HiRISE Digital Terrain Models [1, 14], operating under partial observability with probabilistic communication updates. Rover policies are parameterized as weighted superpositions of spatial potential fields and optimized via simulated annealing under varying communication settings. To enable teammate anticipation, agents simulate peer trajectories over configurable lookahead horizons using internal belief states [8, 13]. Controlled noise injected into prior terrain maps models discrepancies between orbital estimates and true surface hazards. This framework allows us to examine three core questions: (1) how lookahead depth affects coordination under imperfect teammate modeling, (2) how prior map inaccuracies interact with predictive coordination strategies, and (3) how communication frequency mitigates or amplifies lookahead mismatch.

Our contributions are threefold:

(1): We introduce a C++ engine for studying decentralized rover coverage under hazardous terrain and communication constraints.

(2): We analyze lookahead dynamics in potential-field navigation, identifying conditions under which prediction improves coordination and conditions in which mismatch degrades performance.

(3): We characterize the interplay between communication reliability, prior map uncertainty, and predictive teammate modeling in multi-agent spatial exploration.

These results provide insight into the design of decentralized autonomy architectures for multi-rover space missions, clarifying when reactive strategies suffice and when predictive modeling or communication becomes essential.

Appears at the International Workshop on Autonomous Agents and Multi-Agent Systems for Space Applications (MASSpace-26). Held as part of the Workshops at the 25th International Conference on Autonomous Agents and Multiagent Systems., S. Chien, G. Picard, I. Zilberstein (Chairs), May 2026, Paphos, Cyprus.

2 BACKGROUND AND RELATED WORK

Our framework intersects three major pillars of multi-agent systems (MAS) research: spatial coverage through artificial potential fields, lookahead dynamics for teammate anticipation, and the effects of communication/transparency on overcoming partial observability.

2.1 Multi-Agent Coverage and Artificial Potential Fields

Spatial coverage and coordination in partially observable environments are frequently addressed using heuristic-driven navigation, most notably Artificial Potential Fields (APF). Originally developed for collision avoidance, APFs have been widely adapted for multi-agent dispersion, where agents are repelled by obstacles and peers but attracted to unobserved frontiers [12, 17]. Recent work has successfully integrated APFs into Multi-Agent Reinforcement Learning (MARL) to solve complex Coverage Path Planning (CPP) problems, using spatial graphs and curiosity-driven intrinsic rewards to minimize coverage gaps [11, 22].

2.2 Lookahead Dynamics and Teammate Anticipation

In decentralized, partially observable settings, agents often rely on internal forward models to anticipate allied movements. Predicting teammate behavior mitigates the "teammate delay" issue caused by concurrent policy updates [4]. Approaches like Higher-Order Gradients (HOG), Off-Policy Action Anticipation (OffPA2), and Policy Mirror Descent with Lookahead formally integrate lookahead mechanisms so agents can explicitly model the future states or actions of their peers [2, 18].

Our work grounds these theoretical lookahead models in a highly parallelized C++ environment, allowing agents to run exact forward-simulations of their peer's policies based on their current belief state. Crucially, we introduce systemic error into these forward models via inaccurate prior belief maps, forcing agents to navigate the discrepancy between predicted environment danger and dynamic hazard realities. This systemic error and policy noise are the only discrepancies between agent lookahead and reality when communication bandwidth is infinite.

2.3 Predictability, Transparency, and Communication Trade-offs

When communication bandwidth is constrained, agents must act transparently for their peers' lookahead models to remain accurate. Recent MARL research emphasizes *alignment-driven intrinsic rewards*, where agents are incentivized to learn behaviors that match their neighbors' expectations [16], naturally breaking coordination symmetries. Similarly, *Predictability Awareness* mechanisms encourage agents to foster soft social conventions; by minimizing the discrepancy between an internal prediction model and actual observations, agents effectively lower the uncertainty of the entire swarm [9]. Our research builds directly on this by quantifying the trade-off: when terrain realities diverge from the shared global map, agents must either rely heavily on explicit communication to resynchronize the swarm or adopt highly predictable local policies to minimize the resulting lookahead drift.

3 METHODOLOGY

3.1 Simulation Framework

We evaluate our hypotheses using our custom environment named *multi-agent-coverage*, a high-performance, batched multi-agent grid-world environment implemented in C++ with OpenMP for parallel execution. The environment simulates agents exploring a 32×32 grid world containing continuous danger/hazard mappings, h, in the range $[-1, 1]$.

Each agent may propose a movement vector $\vec{v} = [\delta x, \delta y]$ which is clamped to a maximum magnitude of 1.0 before being reduced again by the current danger in $\vec{v} = \vec{v}(1 - c \cdot \max(h, 0.0))$. The constant c determines the minimum speed. Every frame that a rover spends on a tile with $h > 0$ has a $p = ch$ probability of getting stuck. Stuck means that the agent's movement speed is zero until the environment resets.

Our C++ architecture ensures zero-copy memory sharing with PyTorch tensors, allowing for exceptionally high throughput during large-scale policy evaluation. This supports rapid rover belief updates and enables efficient instantiation of environment copies from an agent's internal state for future integration with Monte Carlo Search [3] or Prioritized Replay [20].

3.2 Stochastic Potential Field Navigation

Agent behavior is governed by a parameterized force-based policy. Agents compute an action vector through a linear combination of dynamic spatial features. The force exerted by grid cell i and the resulting weighted superposition are defined as:

$$\vec{F}_i = \frac{m_i \cdot \hat{r}_i}{\|\mathbf{r}_i\|^{p_i}} \qquad \mathbf{a}_t = \sum_{k=1}^{K} w_k \, \vec{F}_k \tag{1}$$

where $\hat{r}_i$ is the unit vector to cell i, $\|\mathbf{r}_i\|$ is the Euclidean distance, p_i is the distance exponent, and m_i is the magnitude. The weights $w_k \in [-1, 1]$ represent signed attraction/repulsion for $K = 7$ spatial features: (1) Region Danger Prior, (2) Observed Danger, (3) Visited Status, (4) Agent Locations, (5) Path Recency, (6) Boundaries, and (7) Voronoi Partitions, plus noise to avoid local minima.

Each of these fields, apart from 1 and 6, is subject to partial observability constraints and calculated from imperfect belief states. Self-Path recency sets recently visited cells to a magnitude of 1.0 with exponential decay over time, serving as an anti-pheromone inspired by path-finding algorithms such as Ant Colony Optimization [7]. Voronoi Partition masks unobserved tile magnitudes if they are closer to another agent, allowing a first-order estimation of task responsibility.

3.3 Policy Optimization via Simulated Annealing

To discover optimal potential field parameters, we deploy a Simulated Annealing (SA) hyperparameter search. The search optimizes a sequence of genes representing the specific hyperparameters being evaluated (specifically the weight w_k and the exponent p_i) for each spatial feature. Fitness is defined as the discounted cumulative coverage reward across a batch of environments:

$$f = \sum_{t=0}^{T} \gamma^t \left(10\,\mathbb{I}[\text{done}_t] + o_t - 100\,s_t\right),$$

where $T = 1000$, γ is the discount factor, o_t is the number of newly observed tiles at timestep t, s_t is the number of rovers that become immobilized due to excessive hazard, and $\mathbb{I}[\text{done}_t]$ is an indicator equal to 1 if full coverage is achieved at time t. This fitness directly captures the mission's dual objectives: maximizing spatial discovery while strictly penalizing rover loss due to hazards.

4 EXPERIMENTAL DESIGN

Our experimental maps are constructed from real Digital Terrain Models (DTMs) obtained via the HiRISE mission, centered on a prospective Mars landing site. Sixteen 1km^2 terrain slices were processed into dual mappings: a maximum gradient map and an elevation standard deviation map, serving as proxies for surface steepness and roughness, respectively. To assess the intersection of transparency, lookahead, and communication, we benchmark optimized rovers across all environments on a matrix of conditions:

(1) Communication Levels: We evaluate full, partial, and zero-communication scenarios. Communication consists of [x,y] location and "is_alive" updates to other rovers.

(2) Lookahead Depth: Agents simulate peer movements for 0 to 10 time-steps (up to 300 meters traversal depending on terrain) before selecting actions. With Lookahead (LA) enabled, agents update unseen teammate locations using their internal belief state.

(3) Prior Map Noise: To simulate real-world inaccuracies, controlled Gaussian noise is injected into the expectation maps during zero-communication runs as shown in Figure (2).

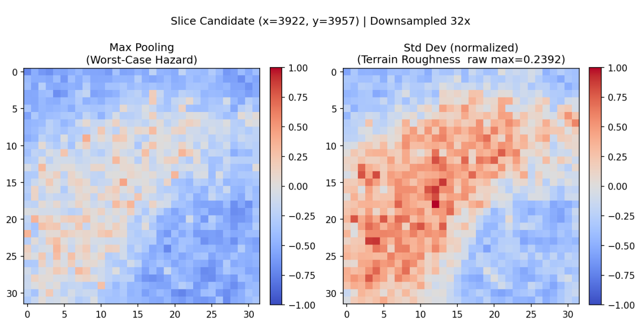

Figure 1: Example terrain slice (1024 × 1024 m). Each cell corresponds to a max-pooled $32\,\mathrm{m}^2$ region derived from HiRISE DTM data.

By masking which features are available to a particular rover's SPF navigation function, we create three baseline policies trained with no communication or lookahead enabled. **Global** and **Random** have access to all variables including ground-truth and partially observed variables. **Global** starts with a noisy Voronoi policy and **Random** begins with random uniform weights over all features. We also mask out all but the local variables for **Local**, and we train variants via simulated annealing specifically for lookahead and communication $p = 0.05$: **Local+LA** and **Random+LA**.

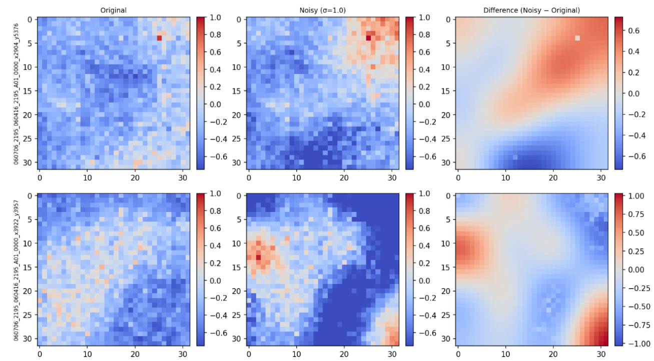

Figure 2: Two map examples with (left) original max-pooled gradient map, (center) noise-perturbed prior map, and (right) the injected Gaussian noise field.

We aim to test several hypotheses about the impact of lookahead, communication, and policy entropy on team performance:

Hypothesis 1: As policy noise weight increase, lookahead prediction will grow significantly correlating with degraded performance. Figures (3, 4).

Hypothesis 2: As the prior belief map of terrain danger becomes more noisy, lookahead will become more destructive. Figure (5).

Hypothesis 3: Lookahead accuracy will improve significantly with occasional correction from communication, improving team performance. Figure (6).

5 RESULTS

To evaluate our hypotheses, we analyzed both coverage performance and ally prediction error across varying lookahead steps and noise conditions.

Lookahead under Imperfect Teammate Modeling (Hypothesis 1): Figures 3 and 4 demonstrate that without the occasional resetting (1/20 chance) to ground truth, lookahead introduced mean squared error and reduced performance for already stochastic policies, but lookahead helped deterministic policies performance. This is because a local minimum in the potential field only changes if an agent's belief about its teammates changes.

Impact of Prior Map Mismatch (Hypothesis 2): Figure 5 illustrates the effect of injecting noise into the agents' prior beliefs. Agents with access to the global potential fields with no lookahead (blue/green solid) are less negatively impacted than local (orange) agents with lookahead (dotted lines). It is worth noting in the prediction error graphs that global agents view the global state, but their local beliefs are not updated besides outside of comms usage.

Communication and Lookahead Utility (Hypothesis 3): Figure 6 shows that although $p = 0.05$ information sharing is enough to bound lookahead error regardless of the number of steps (1-10), naive SPF controllers fail to capitalize on the lookahead information, despite its accuracy.

6 DISCUSSION

While lookahead with zero communication increased MSE compared to stationary imagined teammates, we found the mean absolute error to be around 30% lower at lookahead 1 than 0. This

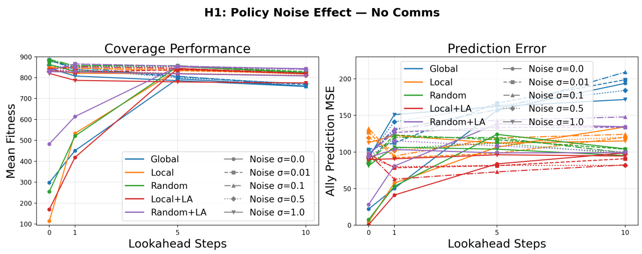

Figure 3: Policy Noise Weight No Comms: Color is the simulated annealing genome type, line-style is the noise level. Left is fitness (higher better) and right is ally location prediction error (lower better)

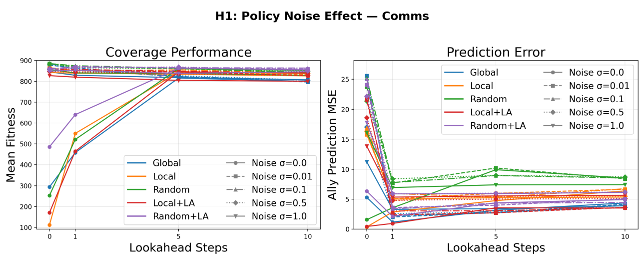

Figure 4: H1: Policy Noise Weight With Comms: The same scenario as Figure (3) but the agents have a 5 percent chance of sharing their up-to-date location with another agent at each frame. It can be seen that slight communication grounds all lookahead by limiting the random walk distance of compounding errors.

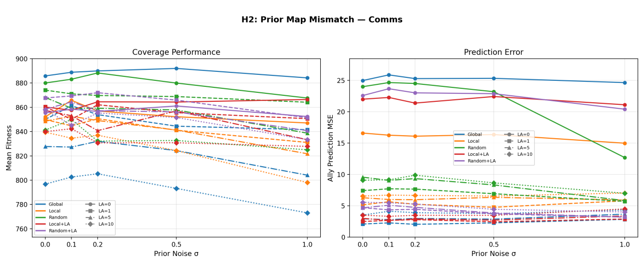

Figure 5: H2: Prior Map Noise Injection: The x axis is the level of injected noise, Color is simulated annealing genome and Line-Type is lookahead. Solid lines (less lookahead) are less impacted in terms of performance

property appears to originate from property ungrounded forward models suffer due to compounding spatial uncertainty ($\propto t^{3/2}$ for integrated random walks [10]) which rapidly exceeds the agent's actual possible traversal distance ($\propto t$)

For potential field navigation, noise can be necessary to push an agent out of areas with a zero gradient. We expected this noise to worsen prediction but that hypothesis was not supported. Interestingly, lookahead appears to act as an effective mechanism for perturbing the agent from a local minimum. We hypothesize that it is unlikely for all 4 agents in the belief state to be in local

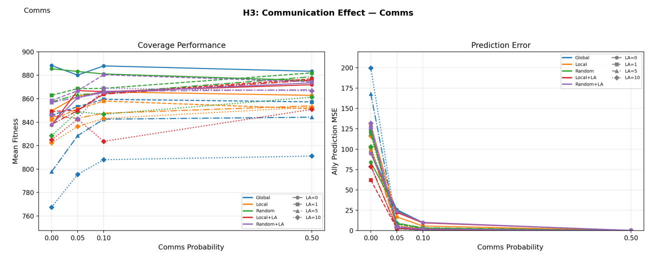

Figure 6: H3 Communication Probability: X-axis is communication probability, color is genome type, and line-style is lookahead. We see that Local (orange) and Local trained with lookahead (red) benefit the most from information sharing, but that 5 percent is sufficient.

minimums simultaneously, and that environment stochasticity as imagined agents "get stuck" with some probability are enough to perturb the potential field so that local minimums are transient.

We found that prior mismatch with communication probability above 0.05 had a negligible impact on performance. Across genomes, the distance drop-off exponent for danger was often 2 or 3, forcing the agent to only care about local danger because only local hazards can cause the agent harm. Because agent's act locally, far away updates are not very impactful beyond agent position updates. For communication probability, somewhere between 1/20 and the time horizon of 1/1000 leads to degradation of lookahead accuracy, but the frequency required to keep lookahead in check is relatively low.

It is important to note the scale limitations of our study: our 32×32 grid is quite course. Fine-grained prior noise or more space to wander may impact the magnitude's of our results.

7 CONCLUSIONS AND FUTURE WORK

This work demonstrates minimal communication ($p = 0.05$) combined with basic agent modeling bounds prediction error regardless of the number of projection steps for simple SPF controllers. While SPF shows robustness to both policy and environment noise, they fail to exploit the predictive information that lookahead provides, by focusing on local information. Potential field navigation is inherently reactive, so changing circumstances are not necessarily destructive to the control scheme.

Future work will involve additional search with teammate models designed for more pointed action selection, such as Monte Carlo Search, or as a lookahead/predictability baseline for deep reinforcement learning approaches. A field like "discovery value" to incentivize agents towards globally relevant objectives may also better highlight the effects of prior map mismatch. We hope that the introduction of a lightweight parallel environment with zero-copy memory sharing and realistic planetary terrain models will serve as a robust testbed for bridging classical heuristic navigation and advanced multi-agent reinforcement learning at a low computational cost. We put forth these initial results as a baseline from which higher-order navigation logic and models can be built and iterated on for a variety of coverage tasks.

REFERENCES

- [1] Ross A Beyer, Oleg Alexandrov, and Scott McMichael. 2018. The Ames Stereo Pipeline: NASA's open source software for deriving and processing terrain data. *Earth and Space Science* 5, 9 (2018), 537–548.
- [2] Ariyan Bighashdel, Daan De Geus, Pavol Jancura, and Gijs Dubbelman. 2024. Off-policy action anticipation in multi-agent reinforcement learning. *Journal of Machine Learning Research* 25, 67 (2024), 1–31.
- [3] Cameron B Browne, Edward Powley, Daniel Whitehouse, Simon M Lucas, Peter I Cowling, Philipp Rohlfshagen, Stephen Tavener, Diego Perez, Spyridon Samothrakis, and Simon Colton. 2012. A survey of monte carlo tree search methods. *IEEE Transactions on Computational Intelligence and AI in games* 4, 1 (2012), 1–43.
- [4] Feng Chen, Xinwei Chen, Rong-Jun Qin, Cong Guan, Lei Yuan, Zongzhang Zhang, and Yang Yu. 2025. Efficient Multi-Agent Cooperation Learning through Teammate Lookahead. *Transactions on Machine Learning Research* (2025).
- [5] Jorge Cortes, Sonia Martinez, Timur Karatas, and Francesco Bullo. 2004. Coverage control for mobile sensing networks. *IEEE Transactions on robotics and Automation* 20, 2 (2004), 243–255.
- [6] Jean-Pierre de la Croix, Federico Rossi, Roland Brockers, Dustin Aguilar, Keenan Albee, Elizabeth Boroson, Abhishek Cauligi, Jeff Delaune, Robert Hewitt, Dima Kogan, et al. 2024. Multi-agent autonomy for space exploration on the cadre lunar technology demonstration. In *2024 IEEE Aerospace Conference*. IEEE, 1–14.
- [7] Marco Dorigo, Mauro Birattari, and Thomas Stutzle. 2006. Ant colony optimization. *IEEE computational intelligence magazine* 1, 4 (2006), 28–39.
- [8] Jakob Foerster, Gregory Farquhar, Triantafyllos Afouras, Nantas Nardelli, and Shimon Whiteson. 2018. Counterfactual multi-agent policy gradients. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 32.
- [9] Roman Chiva Gil, Daniel Jarne Ornia, Khaled A Mustafa, and Javier Alonso Mora. 2024. Predictability Awareness for Efficient and Robust Multi-Agent Coordination. *arXiv preprint arXiv:2411.06223* (2024).
- [10] Claude Godrèche and Jean-Marc Luck. 2022. Record statistics of integrated random walks and the random acceleration process. *Journal of Statistical Physics* 186, 1 (2022), 4.
- [11] M. Hasan and R. Niyogi. 2025. Efficient Multi-Agent Exploration in Area Coverage Under Spatial and Resource Constraints. In *Proc. of ICAART*, Vol. 3. 1278–1287.
- [12] O. Khatib. 1986. Real-time obstacle avoidance for manipulators and mobile robots. *The International Journal of Robotics Research* (1986).
- [13] Ryan Lowe, Yi I Wu, Aviv Tamar, Jean Harb, OpenAI Pieter Abbeel, and Igor Mordatch. 2017. Multi-agent actor-critic for mixed cooperative-competitive environments. *Advances in neural information processing systems* 30 (2017).
- [14] Alfred S McEwen, Eric M Eliason, James W Bergstrom, Nathan T Bridges, Candice J Hansen, W Alan Delamere, John A Grant, Virginia C Gulick, Kenneth E Herkenhoff, Laszlo Keszthelyi, et al. 2007. Mars reconnaissance orbiter's high resolution imaging science experiment (HiRISE). *Journal of Geophysical Research: Planets* 112, E5 (2007).
- [15] Sharan Nayak, Grace Lim, Federico Rossi, Michael Otte, and Jean-Pierre de la Croix. 2025. Multi-robot exploration for the CADRE mission. *Autonomous Robots* 49, 2 (2025), 17.
- [16] Kamal K Ndousse, Douglas Eck, Sergey Levine, and Natasha Jaques. 2021. Emergent social learning via multi-agent reinforcement learning. In *International conference on machine learning*. PMLR, 7991–8004.
- [17] A. Pertzovsky, R. Stern, A. Felner, and R. Zivan. 2025. Enhancing Lifelong Multi-Agent Path-finding by Using Artificial Potential Fields. In *Proc. of AAMAS*.
- [18] Kimon Protopapas and Anas Barakat. 2024. Policy mirror descent with lookahead. *Advances in Neural Information Processing Systems* 37 (2024), 26443–26481.
- [19] Gregg Rabideau, Joseph Russino, Andrew Branch, Nihal Dhamani, Tiago Stegun Vaquero, Steve Chien, Jean-Pierre de la Croix, and Federico Rossi. 2025. Planning, scheduling, and execution on the Moon: the CADRE technology demonstration mission. *arXiv preprint arXiv:2502.14803* (2025).
- [20] Tom Schaul, John Quan, Ioannis Antonoglou, and David Silver. 2015. Prioritized experience replay. *arXiv preprint arXiv:1511.05952* (2015).
- [21] Mac Schwager, Daniela Rus, and Jean-Jacques Slotine. 2009. Decentralized, adaptive coverage control for networked robots. *The International Journal of Robotics Research* 28, 3 (2009), 357–375.
- [22] Yuanbo Zhu, Guangjie Han, Chuan Lin, Fan Zhang, and Yun Hou. 2025. Multi-USV Coverage Path Planning Using Spatial Graph Multi-Actor-Attention-Critic Reinforcement Learning Framework With Operator Pooling. *IEEE Transactions on Mobile Computing* (2025).