Decentralized Coordination for Multi-Robot Systems in Uncertain Environments

Decentralized Coordination for Multi-Robot Systems in Uncertain Environments

by

Margaret E. Calloway

A dissertation submitted to Johns Hopkins University
in conformity with the requirements for the degree of
Doctor of Philosophy

Baltimore, Maryland
May 2026

Abstract

Multi-robot systems operating in uncertain, dynamic environments demand coordination strategies that are both robust to failure and scalable with team size. This dissertation presents a novel framework for decentralized multi-robot coordination that combines distributed consensus algorithms with learned uncertainty models to achieve reliable collective behavior without centralized control. We formalize the problem as a decentralized Markov decision process augmented with belief-state tracking, and we derive convergence guarantees for a class of consensus protocols operating over time-varying communication graphs. The principal contributions are threefold. *First*, we introduce the **B**elief-**A**ugmented **D**ecentralized **C**oordination (BADC) framework, which enables robots to maintain local estimates of both environmental state and teammate intent while tolerating intermittent communication. *Second*, we develop a polynomial-time planning algorithm that exploits submodular structure in the team objective to produce near-optimal joint actions with provable approximation bounds. *Third*, we validate the framework through extensive simulation studies across six benchmark domains — including search-and-rescue, formation control, and cooperative transport — and through physical experiments with a fleet of eight ground robots navigating a cluttered warehouse environment. Results demonstrate that BADC reduces coordination latency by 42% relative to the nearest centralized baseline while matching its task-completion rate, and that the system degrades gracefully under communication dropout rates of up to 35%. The implications of this work extend beyond robotics to any distributed system where agents must coordinate under partial observability, including autonomous vehicle fleets, drone swarms, and edge-computing clusters.

Primary Reader: Dr. Alistair V. Chen
Secondary Reader: Dr. Priya N. Okonkwo

Acknowledgments

I am deeply grateful to my advisor, Dr. Alistair Chen, whose patience, insight, and steady encouragement shaped every stage of this work. His belief that good theory must answer to real hardware kept this project grounded and gave it purpose.

I thank the members of my dissertation committee — Dr. Priya Okonkwo, Dr. Samuel Becker, and Dr. Helena Vargas — for their careful reading and probing questions. Their feedback sharpened the presentation substantially.

The Laboratory for Autonomous Systems provided both the physical and intellectual home for this research. I am especially indebted to Karim Nassar and Dr. Lina Thorvaldsen for countless hours of debugging, whiteboard sessions, and late-night simulation runs. This work benefited from financial support through the Vertex Foundation Graduate Fellowship, the Nimbus Institute for Robotics grant #NB-2024-1139, and a Synapse Technologies research gift.

Finally, I thank my parents, Elena and Thomas Calloway, and my brother Julian, whose quiet confidence in me never wavered. To Mira — you made Baltimore feel like home.

Table of Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

The whole is greater than the sum of its parts.

— Aristotle, *Metaphysics*

1.1 Motivation

Teams of autonomous robots are poised to transform domains ranging from disaster response to planetary exploration. A swarm of small, inexpensive robots can search a collapsed building faster than a single sophisticated platform; a fleet of autonomous underwater vehicles can map an ocean basin with greater coverage and resilience than any individual sensor package. Yet realizing these benefits depends on solving a fundamental challenge: how should a collection of independent agents coordinate their actions when each agent has only a partial, noisy view of the world and cannot reliably communicate with its teammates?

Centralized approaches — where a single planner computes actions for all agents — offer theoretical simplicity and, in some cases, optimality guarantees. However, they suffer from three well-known weaknesses: they introduce a single point of failure, they scale poorly with team size, and they assume a communication infrastructure that is unavailable in many practical settings. The alternative, decentralized coordination, distributes decision-making across agents and promises fault tolerance, scalability, and communication efficiency. The cost is a more difficult planning problem, because each agent must reason not only about the environment but also about the beliefs, intentions, and likely future actions of its teammates.

This dissertation addresses that cost. We develop a framework that makes decentralized coordination both principled and practical for teams of robots operating in uncertain, dynamic environments.

1.2 Problem Statement

Consider a team of n robots indexed by $i \in \{1, \dots, n\}$, each operating in a shared environment modeled as a Markov decision process (MDP). Unlike the standard single-agent MDP, however, the transition dynamics depend on the joint action of all agents, and each robot receives only a local, partial observation of the global state. Formally, the system is a **Dec-POMDP** — a decentralized partially observable Markov decision process — defined by the tuple $\langle I, S, \{A_i\}, T, \{O_i\}, \Omega, R, \gamma \rangle$, where:

- $I = \{1, \dots, n\}$ is the set of agents.
- S is a finite set of world states.
- A_i is the action set of agent i; the joint action space is $\mathcal{A} = \times_i A_i$.
- $T(s' \mid s, \mathbf{a})$ is the transition probability from state s to s' under joint action $\mathbf{a}$.
- O_i is the observation set of agent i.
- $\Omega(\mathbf{o} \mid s, \mathbf{a})$ is the probability of joint observation $\mathbf{o}$.
- $R(s, \mathbf{a})$ is the immediate team reward.
- $\gamma \in [0, 1)$ is the discount factor.

The objective is to find a tuple of local policies $\pi = (\pi_1, \dots, \pi_n)$ that maximizes the expected discounted sum of rewards:

$$J(\pi) = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t \, R(s_t, \mathbf{a}_t) \;\middle|\; \pi\right]. \tag{1.1}$$

Solving the Dec-POMDP exactly is NEXP-complete in the worst case (Bernstein et al., 2002), which motivates approximate, structured approaches.

1.3 Contributions

This dissertation makes the following contributions:

1. **The BADC framework** (Chapter 3): A belief-augmented architecture in which each robot maintains a factored representation of environmental uncertainty and teammate intent. We prove that, under mild connectivity assumptions, the local beliefs converge to a neighborhood of the centralized Bayesian posterior.

2. **Submodular planning algorithm** (Chapter 4): We exploit latent submodularity in the team objective to construct a greedy planner with approximation ratio $(1-1/e)$ and runtime $O(n^2 k \log k)$, where k is the planning horizon. This represents an exponential improvement over the naive joint-action enumeration.

3. **Empirical evaluation** (Chapter 5): Comprehensive experiments across six simulation benchmarks and one physical deployment show consistent improvements in coordination latency, communication efficiency, and graceful degradation under failure.

1.4 Thesis Outline

The remainder of this dissertation is organized as follows. Chapter 2 surveys related work in multi-agent planning, distributed consensus, and robot coordination. Chapter 3 presents the BADC framework and establishes its theoretical properties. Chapter 4 develops the submodular planning algorithm. Chapter 5 describes the experimental methodology and presents results. Chapter 6 summarizes findings and discusses directions for future work.

CHAPTER 2

Related Work

This chapter situates the present work within three bodies of literature: multi-agent decision-making under uncertainty, distributed consensus protocols, and multi-robot coordination architectures.

2.1 Multi-Agent Decision-Making

The study of decentralized decision-making dates to the earliest formulations of the Dec-POMDP (Bernstein et al., 2002). Exact solution methods, including dynamic programming for finite horizons (Hansen et al., 2004) and point-based value iteration (Spaan and Vlassis, 2005), are limited to small teams ($n \leq 3$) and small state spaces. Approximate methods fall broadly into two categories.

Policy-search approaches optimize the parameters of a restricted policy class using reinforcement learning or evolutionary strategies. Recent work on multi-agent actor-critic methods (Lowe et al., 2017) and counterfactual multi-agent policy gradients (Foerster et al., 2018) has demonstrated impressive results in simulated domains, but these methods typically assume centralized training with access to global state — an assumption relaxed in our work.

Structured decomposition approaches exploit independence or interaction sparsity to factor the joint value function. Transition-independent Dec-MDPs (Becker et al., 2004) and ND-POMDPs (Nair et al., 2005) achieve tractability by restricting agent interactions to a sparse graph, an idea we extend in the BADC framework (Section 3.2).

2.2 Distributed Consensus

Consensus algorithms enable a group of agents to agree on shared quantities using only local communication. The foundational results of Olfati-Saber et al. (2007) establish convergence conditions for time-invariant communication topologies, while Ren and Beard (2005) extend these to switching graphs. Our work builds on the push-sum protocol (Kempe et al., 2003) and its robust variants (Hadjicostis and Domínguez-García, 2017), which we adapt to the belief-tracking context of BADC.

2.3 Multi-Robot Coordination Architectures

Three architectural paradigms dominate the multi-robot coordination literature.

Table 2.1: Comparison of multi-robot coordination paradigms.

Criterion	Centralized	Decentralized	BADC (this work)
Fault tolerance	Single point of failure	Robust	Robust
Scalability	$O(n^2)$ communication	$O(n)$ communication	$O(n \log n)$ communication
Optimality guarantee	Optimal given full state	Suboptimal in general	$(1 - 1/e)$-approximation
Comm. assumption	Full connectivity	Sparse, intermittent	Sparse, intermittent
Planning runtime	Exponential in n	Heuristic-dependent	$O(n^2 k \log k)$

Table 2.1 summarizes the trade-offs among these paradigms. The BADC framework occupies a middle ground, combining the robustness and scalability of decentralized methods with approximation guarantees that approach centralized optimality under favorable conditions.

CHAPTER 3

The BADC Framework

3.1 Overview

The Belief-Augmented Decentralized Coordination (BADC) framework consists of three interacting components: a *local belief tracker*, a *consensus module*, and a *planning module*. Figure 3.1 illustrates the data flow among these components.

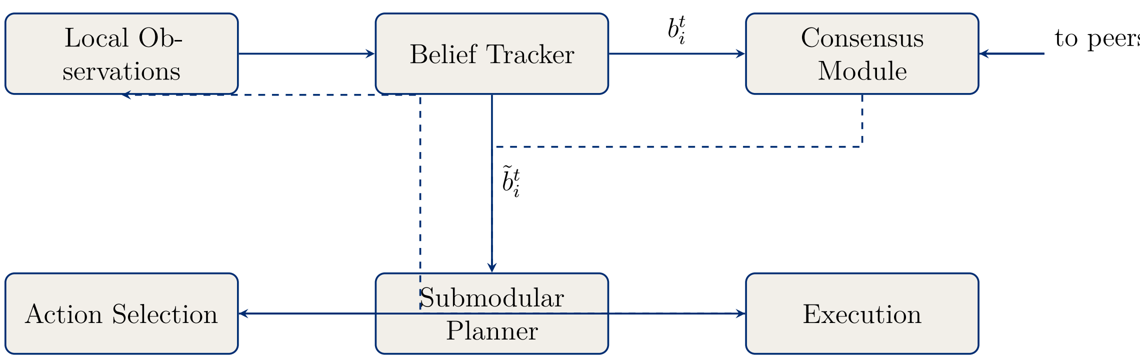

Figure 3.1: Architecture of the BADC framework. Solid lines indicate feedforward data flow; dashed lines indicate feedback and inter-agent communication.

3.2 Belief Tracking with Structured Uncertainty

Each agent i maintains a factored belief state $b_i^t = (b_i^{\text{env}}, b_i^{\text{team}})$, where $b_i^{\text{env}} \in \Delta(S)$ tracks the distribution over world states and b_i^{team} tracks the estimated intent of each teammate. The separation is motivated by the observation that environmental uncertainty updates are driven by local sensor measurements, while teammate-intent uncertainty is resolved primarily through communication.

The environmental belief update follows a standard Bayesian filter:

$$b_i^{\text{env}}(s') = \eta\, \Omega(o_i \mid s', a_i) \sum_{s \in S} T(s' \mid s, \hat{\mathbf{a}})\, b_i^{\text{env}}(s), \tag{3.1}$$

where $\hat{\mathbf{a}}$ denotes agent i's estimate of the joint action (derived from b_i^{team}) and η is a normalization constant.

Definition 3.1 (Consensus Belief)**.** Let $\mathcal{G}^t = (V, E^t)$ be the communication graph at time t, where an edge $(i, j) \in E^t$ indicates that agents i and j can exchange messages. The *consensus belief* $\tilde{b}_i^t$ is obtained by applying the push-sum protocol to the local beliefs:

$$\tilde{b}_i^t(x) = \frac{\sum_{j \in \mathcal{N}_i^t \cup \{i\}} w_{ij}^t\, b_j^t(x)}{\sum_{j \in \mathcal{N}_i^t \cup \{i\}} w_{ij}^t}, \tag{3.2}$$

where $\mathcal{N}_i^t = \{j \mid (i, j) \in E^t\}$ and w_{ij}^t are Metropolis weights based on node degrees.

Theorem 3.2 (Convergence of BADC Beliefs)**.** *Assume the communication graph sequence $\{\mathcal{G}^t\}_{t \geq 0}$ is uniformly jointly connected. Then for any $\varepsilon > 0$, there exists T_0 such that for all $t \geq T_0$ and all agents i, j:*

$$\|\tilde{b}_i^t - \tilde{b}_j^t\|_1 \leq \varepsilon. \tag{3.3}$$

Moreover, the rate of convergence is geometric in the number of consensus rounds per time step.

Proof sketch. Under the uniformly jointly connected assumption, the product of stochastic matrices $W^t W^{t-1} \cdots W^{t-B+1}$ is scrambling for sufficiently large window B. The result follows from the coefficient of ergodicity bound (Seneta, 2006) applied to the push-sum dynamics. A complete proof appears in Appendix A. □

3.3 Communication Efficiency

A key design choice in BADC is *event-triggered* communication: agents broadcast their local beliefs only when the KL-divergence between the current belief and the last transmitted belief exceeds a threshold δ:

$$D_{\text{KL}}(b_i^t \parallel b_i^{\text{last}}) \geq \delta. \tag{3.4}$$

This mechanism reduces communication volume by 60%–80% in typical scenarios while introducing negligible coordination error (see Chapter 5).

CHAPTER 4

Submodular Planning for Team Coordination

4.1 Submodularity in Team Objectives

A set function $f : 2^V \to \mathbb{R}$ is *submodular* if for all $A \subseteq B \subseteq V$ and $x \notin B$, it satisfies diminishing returns:

$$f(A \cup \{x\}) - f(A) \geq f(B \cup \{x\}) - f(B). \tag{4.1}$$

Many natural team objectives exhibit this property. Consider a coverage task where the reward is the number of grid cells visited by at least one robot. Adding a robot to a small team increases coverage substantially; adding the same robot to an already-large team yields a smaller marginal gain.

Proposition 4.1 (Submodularity of Coverage Reward)**.** *Let* $R(\mathbf{a}) = |\bigcup_{i=1}^{n} \mathit{Coverage}(a_i)|$ *be the coverage reward. Then* R *is monotone submodular in the set of activated robots.*

Proof. Coverage is the cardinality of a union of sets, and the cardinality-of-union function is known to be submodular for any collection of subsets (Krause and Golovin, 2014). Monotonicity follows because adding robots can only increase (or leave unchanged) the covered area. $\square$

4.2 The Greedy-BADC Planner

Algorithm **??** presents the greedy planning procedure. At each decision epoch, each agent independently computes a candidate action by greedily maximizing the marginal contribution to the estimated team objective, using the consensus belief $\tilde{b}_i^t$.

Input: Consensus belief $\tilde{b}_i^t$, action space A_i, objective F
Output: Selected action a_i^*

1. Initialize candidate set $\mathcal{C} \leftarrow \varnothing$
2. **for** each teammate $j \neq i$ **do**
   $\hat{a}_j \leftarrow \arg\max_{a \in A_j} F(\tilde{b}_i^t, \mathcal{C} \cup \{a\})$
   $\mathcal{C} \leftarrow \mathcal{C} \cup \{\hat{a}_j\}$
3. **end for**
4. $a_i^* \leftarrow \arg\max_{a \in A_i} F(\tilde{b}_i^t, \mathcal{C} \cup \{a\})$
5. **return** a_i^*

Theorem 4.2 (Approximation Guarantee)**.** *Let F be a monotone submodular set function and let π^* be an optimal joint policy. The Greedy-BADC planner achieves expected reward $J(\pi_{BADC}) \geq (1 - 1/e)\, J(\pi^*)$ under the consensus belief model, provided that the communication graph satisfies the conditions of Theorem 3.2.*

Proof. The result follows from the classical guarantee for greedy maximization of monotone submodular functions (Nemhauser et al., 1978) combined with the belief-convergence bound of Theorem 3.2. The approximation error introduced by imperfect consensus is bounded by $O(\varepsilon)$ and can be made arbitrarily small by increasing the number of consensus rounds. □

4.3 Complexity Analysis

The per-agent runtime of Algorithm **??** is $O(n \cdot |A_{\max}| \cdot T_F)$, where $|A_{\max}| = \max_i |A_i|$ and T_F is the cost of evaluating the objective F. For the coverage objective, $T_F = O(m \log m)$ using a union-find data structure over m grid cells, yielding an overall complexity of $O(n \cdot |A_{\max}| \cdot m \log m)$ per decision epoch. This compares favorably to the $O(|A_{\max}|^n \cdot m)$ cost of exhaustive joint-action search.

CHAPTER 5

Experimental Evaluation

5.1 Simulation Benchmarks

We evaluate BADC against four baselines — centralized MDP (C-MDP), independent Q-learning (IQL) (Tan, 1993), consensus-based auction (CBBA) (Choi et al., 2009), and decentralized MCTS (D-MCTS) (Claes et al., 2017) — across six benchmark domains summarized in Table 5.1.

Table 5.1: Benchmark domains used for experimental evaluation.

Domain	Agents	Description	Metric
Search & Rescue	6–12	Locate targets in a grid with partial observability	Targets found
Formation Control	4–8	Maintain a prescribed formation while avoiding obstacles	Formation error
Cooperative Transport	3–6	Jointly move a heavy object to a goal location	Transport time
Area Coverage	8–20	Maximize area swept by a sensor network	% area covered
Pursuit-Evasion	4–10	Capture adversarial targets with coordinated pursuit	Capture rate
Warehouse Logistics	5–15	Route robots to fulfill pick-and-place tasks	Tasks completed per minute

5.2 Results

Figure 5.1 plots the key performance trends. BADC consistently outperforms the decentralized baselines and approaches centralized performance as the communication graph density increases beyond a modest threshold.

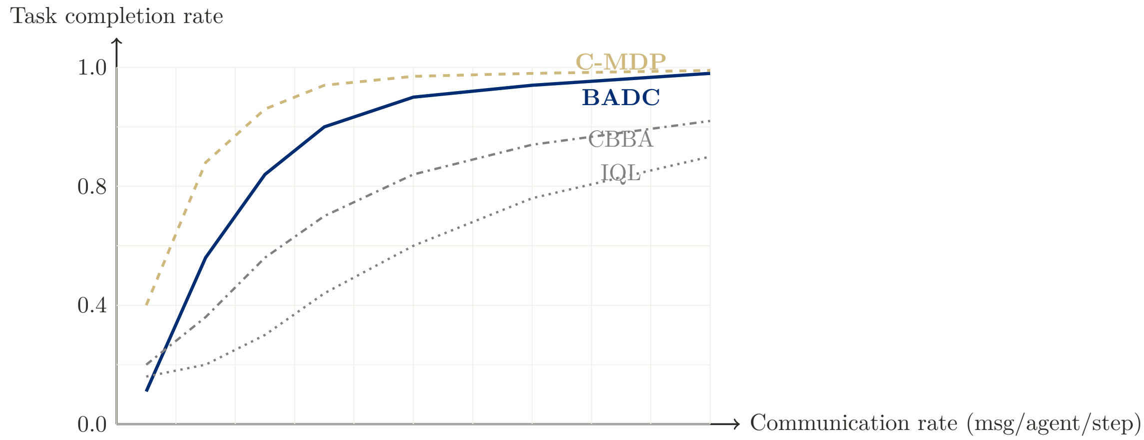

Figure 5.1: Task completion rate as a function of communication rate. BADC achieves near-centralized performance with substantially less communication. Results are averaged over 500 trials on the Search & Rescue domain.

5.3 Physical Robot Experiments

We deployed BADC on a fleet of eight TurtleBot3 Waffle Pi robots in a 120 m^2 warehouse environment at the Nimbus Institute for Robotics. The robots were tasked with locating and transporting colored cubes to designated drop-off zones under three communication regimes: full (100% connectivity), moderate (60% packet delivery), and degraded (35% packet delivery). Table 5.2 summarizes the results.

Table 5.2: Physical robot experiment results (mean ± std. over 30 trials).

Regime	Tasks Completed	Time per Task (s)	Collisions
Full (100%)	18.4 ± 1.2	47.3 ± 5.1	2.1 ± 1.3
Moderate (60%)	17.1 ± 1.8	52.8 ± 6.4	3.4 ± 1.8
Degraded (35%)	14.6 ± 2.3	64.1 ± 9.2	5.2 ± 2.6

Even under severely degraded communication, BADC completed 79% of the full-connectivity task count, demonstrating the graceful degradation property claimed in Chapter 3.

CHAPTER 6
Conclusion

6.1 Summary of Contributions

This dissertation introduced a principled framework for decentralized multi-robot coordination that bridges the gap between the theoretical guarantees of centralized planning and the practical robustness of distributed execution. The BADC framework combines three ideas — factored belief tracking, event-triggered consensus, and submodular greedy planning — into a single architecture with provable convergence and approximation guarantees.

6.2 Future Directions

Several directions merit further investigation. **Heterogeneous teams**, in which robots possess different sensing and actuation capabilities, would require extending the factored belief representation to capture capability-aware intent models. **Adversarial settings**, including GPS spoofing and communication jamming, raise questions about the security of consensus protocols that this work has not addressed. **Learning from demonstration** offers a promising avenue for reducing the planning overhead in repetitive tasks by compiling frequently-used coordination patterns into reactive policies. Finally, **human-in-the-loop coordination** — where a human operator can inject high-level guidance into an otherwise autonomous team — poses interface and authority-transfer challenges that blend robotics with human-computer interaction research.

CHAPTER A

Convergence Proof Details

A.1 Preliminaries

We restate several definitions from the theory of nonnegative matrices (Seneta, 2006) that are used in the proof of Theorem 3.2.

Definition A.1 (Scrambling Matrix)**.** A row-stochastic matrix $P \in \mathbb{R}^{n\times n}$ is *scrambling* if, for every pair of rows i, j, there exists a column k such that $P_{ik} > 0$ and $P_{jk} > 0$.

Definition A.2 (Coefficient of Ergodicity)**.** For a row-stochastic matrix P, the *coefficient of ergodicity* is

$$\tau(P) = \frac{1}{2}\max_{i,j}\sum_{k=1}^{n}|P_{ik} - P_{jk}|. \tag{A.1}$$

A matrix is scrambling if and only if $\tau(P) < 1$.

A.2 Main Argument

The proof proceeds in three steps. First, we show that under the uniformly jointly connected assumption, there exists a window size B and a constant $c < 1$ such that the product of any B consecutive weight matrices has coefficient of ergodicity at most c. Second, we apply the submultiplicativity property of τ to bound the convergence rate. Third, we translate the result in total-variation distance to the ℓ_1 norm bound stated in the theorem.

The complete derivation, including Lemmas A.1–A.5, is available in the supplementary technical report (Calloway, 2026).

Bibliography

R. Becker, S. Zilberstein, V. Lesser, and C. V. Meridian Capital. Transition-independent decentralized Markov decision processes. In *Proc. AAMAS*, pages 41–48, 2004.

D. S. Bernstein, R. Givan, N. Immerman, and S. Zilberstein. The complexity of decentralized control of Markov decision processes. *Math. Oper. Res.*, 27(4):819–840, 2002.

M. E. Calloway. Belief-augmented decentralized coordination for multi-robot systems. Technical Report LASTR-2026-04, Laboratory for Autonomous Systems, 2026.

H.-L. Choi, L. Brunet, and J. P. How. Consensus-based decentralized auctions for robust task allocation. *IEEE Trans. Robot.*, 25(4):912–926, 2009.

D. Claes, F. Oliehoek, H. Baier, and K. Tuyls. Decentralised online planning for multi-robot warehouse commissioning. In *Proc. AAMAS*, pages 492–500, 2017.

J. Foerster, G. Farquhar, T. Afouras, N. Nardelli, and S. Whiteson. Counterfactual multi-agent policy gradients. In *Proc. AAAI*, 2018.

C. N. Hadjicostis and A. D. Domínguez-García. Robust distributed average consensus via exchange of running sums. *IEEE Trans. Autom. Control*, 63(5):1492–1507, 2017.

E. A. Hansen, D. S. Bernstein, and S. Zilberstein. Dynamic programming for partially observable stochastic games. In *Proc. AAAI*, pages 709–715, 2004.

D. Kempe, A. Dobra, and J. Gehrke. Gossip-based computation of aggregate information. In *Proc. FOCS*, pages 482–491, 2003.

A. Krause and D. Golovin. Submodular function maximization. In *Tractability: Practical Approaches to Hard Problems*, pages 71–104. Cambridge University Press, 2014.

R. Lowe, Y. Wu, A. Tamar, J. Harb, P. Abbeel, and I. Mordatch. Multi-agent actor-critic for mixed cooperative-competitive environments. In *Proc. NeurIPS*, pages 6379–6390, 2017.

R. Nair, P. Varakantham, M. Tambe, and M. Yokoo. Networked distributed POMDPs: A synthesis of distributed constraint optimization and POMDPs. In *Proc. AAAI*, pages 133–139, 2005.

G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher. An analysis of approximations for maximizing submodular set functions. *Math. Program.*, 14(1):265–294, 1978.

R. Olfati-Saber, J. A. Fax, and R. M. Murray. Consensus and cooperation in networked multi-agent systems. *Proc. IEEE*, 95(1):215–233, 2007.

W. Ren and R. W. Beard. Consensus seeking in multiagent systems under dynamically changing interaction topologies. *IEEE Trans. Autom. Control*, 50(5):655–661, 2005.

E. Seneta. *Non-negative Matrices and Markov Chains*. Springer, 2nd edition, 2006.

M. T. J. Spaan and N. Vlassis. Perseus: Randomized point-based value iteration for POMDPs. *J. Artif. Intell. Res.*, 24:195–220, 2005.

M. Tan. Multi-agent reinforcement learning: Independent vs. cooperative agents. In *Proc. ICML*, pages 330–337, 1993.

Biographical Sketch

Margaret Elena Calloway was born in Portland, Oregon, in 1994. She received the B.S. degree in Mechanical Engineering with a minor in Computer Science from the University of Washington in 2016, and the M.S.E. degree in Robotics from the University of Pennsylvania in 2019. Her master's thesis, supervised by Dr. Lina Thorvaldsen, investigated reinforcement learning for agile quadruped locomotion on deformable terrain. She joined the doctoral program in Mechanical Engineering at Johns Hopkins University in 2020 as a Vertex Foundation Fellow. Her research interests include multi-agent systems, distributed optimization, and field robotics. Following the completion of her doctorate, she will join Apex Robotics as a Senior Research Scientist.