Mutual Information-based Gradient-ascent Control for Distributed Robotics

by

Brian John Julian

B.S., Cornell University (2005)

S.M., Massachusetts Institute of Technology (2009)

Submitted to the Department of
Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2013

Author .....
Department of Electrical Engineering and Computer Science
August 30, 2013

Certified by .....
Daniela L. Rus
Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by ...
Leslie A. Kolodziejski
Chair, Department Committee on Graduate Students

This work is sponsored by the Department of the Air Force under Air Force contract number FA8721-05-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the United States Government.

Mutual Information-based Gradient-ascent Control for Distributed Robotics

by

Brian John Julian

Submitted to the Department of Electrical Engineering and Computer Science
on August 30, 2013, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This thesis presents the derivation, analysis, and implementation of a novel class of decentralized mutual information-based gradient-ascent controllers that continuously move robots equipped with sensors to better observe their environment. We begin with the fundamental problem of deploying a single ground robot equipped with a range sensor and tasked to build an occupancy grid map. The desired explorative behaviors of the robot for occupancy grid mapping highlight the correlation between the information content and the spatial realization of the robot's range measurements. We prove that any occupancy grid controller tasked to maximize a mutual information reward function is eventually attracted to unexplored space, i.e., areas of highest uncertainty. We show that mutual information encodes geometric relationships that are fundamental to robot control and yields geometrically relevant reward surfaces on which robots can navigate.

Taking inspiration from geometric-based approaches to distributed robot coordination, we show that many multi-robot inference tasks can be cast in terms of an optimization problem. This optimization problem defines the task of minimizing the conditional entropy associated with the robots' inferred beliefs of the environment, which is equivalent to maximizing the mutual information between the environment state and the robots' next joint observation. Given simple robot dynamics and few probabilistic assumptions, none of which involve Gaussianity, we derive a gradient-ascent solution approach to these optimization problems that is convergent between sensor observations and locally optimal. More formally, we invoke LaSalle's Invariance Principle to prove that, given enough time between consecutive joint observations, robots following the gradient of mutual information will converge to goal positions that locally maximize the expected information gain resulting from the next observation.

We show that the algorithmic implementation of the generalized gradient-ascent controller is not readily distributed among multiple robots, and thus sample-based methods are introduced to distributively approximate the likelihoods of the robots' joint observations. Not only are the involved non-parametric representations com-

patible with any type of Bayesian filter, but the computational complexities of the resulting decentralized controllers are independent with respect to the number of robots. Concerning the distributed approximations, we give two example consensus-based algorithms that run on an undirected network graph. The first consensus-based algorithm approximates discrete measurement probabilities, while the second approximates continuous likelihood distributions. We show that these anytime approximations provably converge to the correct values on a static and connected network graph without knowledge of the number of robots in the network or the corresponding graph's topology.

Lastly, we incorporate the resulting consensus-based algorithms into both a hardware system and a simulation environment to allow for decentralized controller evaluation under non-ideal network settings. For the hardware experiments, the task is to infer the state of a bounded, planar environment by deploying five quadrotor flying robots with simulated sensors in both indoor and outdoor settings. For the numerical simulations, Monte Carlo-based analyses are performed for 100 robots, where each robot is simulated on an independent computer node within a computer cluster system. Simulations are also performed for 1000 robots using a single workstation computer equipped with a multicore GPU-enabled graphics card. The results from both the hardware experiments and numerical simulations validate our theoretical and computational claims throughout the thesis.

Thesis Supervisor: Daniela L. Rus
Title: Professor of Electrical Engineering and Computer Science

Acknowledgements

This thesis would not have been possible without the support from those around me. I thank first and foremost my advisor Prof. Daniela Rus, whose steadfast confidence in my work kept me excited and motivated from day one. Her creativity and enthusiasm are unmatched, and her ability to evaluate the potential of an idea is nothing short of visionary. I am truly lucky to be a part of her academic family.

I am also deeply indebted to the professors of my thesis committee. Everyone knows (or should know!) how important Prof. Leslie Kaelbling's contributions are to the general robotics research community, but I want her to know how important her contributions are to my thesis. The same applies to Prof. John Leonard, who is one of the few researchers that maintains a realistically high threshold when considering whether or not a problem is "solved" - something that an applications-focused engineer such as myself greatly respects. Lastly, I have thoroughly enjoyed Prof. Mac Schwager's mentorship over the entirety of my graduate career, and I am a better person because agreed to be on my committee.

There are many other faculty and colleagues that I wish to thank. The joy of being part of the Distributed Robotics Laboratory can be directly attributed to Prof. Carrick Detweiler, Prof. Steve Smith, Prof. Marty Vona, Prof. Elizabeth Basha, Prof. Nora Ayanian, Dr. Kyle Gilpin, Dr. Iuliu Vasilescu, Dr. Ross Knepper, Dr. Danny Feldman, Capt. Wil Selby, Stephanie Gil, Andy Marchese, and Danny Soltero. Special recognition goes to Dr. Marek Doniec for defending first and thus lighting a fire under me to graduate. Before doing so, however, it was an honor to be a teaching assistant under Prof. Sertac Karaman and Prof. Jon How. My graduate career would have been incomplete without this valuable teaching experience.

I thank all my collaborators at the German Aerospace Center (DLR) for their technical guidance and amazingly witty humor. I have the highest respect for Dr. Michael Angermann and Dr. Patrick Robertson, who have become two of my closest friends despite being half a world away. The same applies to Michael Lichtenstern and Martin Frassl, for whom I wish great success in their own academic journey.

I also thank MIT Lincoln Laboratory for graciously supporting my graduate education. Special recognition goes to:

- My group leaders Dr. Donald Coe, Dr. Steve Forman, Craig Perini, and James Ingraham for providing critical feedback of my performance as a student and as an engineer.
- My division leaders Dr. Eli Niewood, Dr. Bob Davis, and Dr. Mike Languirand for supporting my educational and professional development over the years.
- My mentors Dr. Dennis Buranek and Andrew Stimac for offering thoughtful reviews of my publications and this thesis.
- My colleagues Michael Boulet, Byron Stanley, Nick Armstrong-Crews, and Michael Carroll for sharing my passion for robotics.

Most importantly, my work would be meaningless without the love from my family. I thank my parents, Kathleen and John, and my sister, Dawn, for their everlasting patience and optimism. Lastly, to my beautiful wife Bina, you inspire me in all that I do.

Contents

1 Introduction **19**

1.0.1 Aerial surveillance use case 25

1.1 Technical approach 26

1.1.1 Scope and limitations 28

1.2 Research contributions 29

1.2.1 Publications 31

1.3 Thesis outline 32

2 Prior works **35**

2.1 Occupancy grid mapping and SLAM 35

2.2 Decentralized inference and coordination 37

2.3 Distributed algorithms for multi-robot systems 39

2.4 Summary 41

3 Background **43**

3.1 Mathematical notation 43

3.2 Mathematical preliminaries 48

3.2.1 Probability theory 48

3.2.2 Information theory 49

3.2.3 Complexity theory 51

3.3 Summary 51

4 Mutual information-based control for mapping applications **53**

4.1 Eventual attraction to unexplored space 56
4.1.1 The occupancy grid mapping algorithm 56
4.1.2 Main technical assumptions 59
4.1.3 Identification of mutual information-based behaviors 61
4.1.4 A theorem for attractive behavior 64
4.2 On informative behaviors 67
4.2.1 Algebraic properties of mutual information 67
4.2.2 Expected values of future beliefs 69
4.2.3 A proof for attractive behavior 70
4.3 Computing mutual information for mapping 72
4.3.1 Computing the measurement's prior 72
4.3.2 Computing the mutual information reward function 74
4.3.3 Computational complexity 75
4.4 Experiments in mapping 78
4.4.1 Hardware mapping system 79
4.4.2 Mutual information for occupancy grid mapping 80
4.4.3 Mutual information for SLAM 82
4.5 Summary 84
4.6 Proofs 84

5 Gradient-ascent control for multi-robot inference tasks 91
5.1 Problem Formulation 92
5.1.1 Mutual information reward for Bayesian inference 92
5.1.2 Single robot case 95
5.1.3 Centralized multi-robot case with ideal network 96
5.2 Multi-robot coordination 98
5.2.1 Gradient-ascent control 98
5.2.2 Occupancy grid mapping case 102
5.2.3 Sample sets for likely observations 105
5.2.4 Decentralized control 107

5.3 Algorithmic implementations 109
5.3.1 Decentralized standard Bayesian estimation 110
5.3.2 Non-parametric decentralized Bayesian estimation 112
5.4 Experiments in multi-robot systems 115
5.4.1 Quadrotor flying robot system 116
5.4.2 Indoor experiment using a decentralized discrete Bayesian filter 118
5.4.3 Indoor experiment using a decentralized non-parametric Bayesian filter 120
5.4.4 Outdoor experiment using a decentralized non-parametric Bayesian filter 122
5.5 Summary 123
5.6 Proofs 123

6 Distributed algorithms enabling decentralized inference and coordination **131**

6.1 Distributed robot network 133
6.1.1 Communication model 133
6.1.2 Discovery of the maximum in/out degree 135
6.2 Consensus of joint measurement probabilities 136
6.2.1 Distributive approximations for the sampled joint measurement probabilities 137
6.2.2 Algorithmic implementation 140
6.3 Consensus of joint measurement likelihood distributions 142
6.3.1 Distributive approximations for scaled Gaussian joint measurement likelihood distributions 142
6.3.2 Algorithmic implementation 144
6.3.3 Performance guarantees 146
6.4 Parallelized numerical simulations 149

6.4.1 Consensus-only simulations approximating measurement likelihood distributions 149
6.4.2 Control simulations performing consensus for the measurement probabilities 151
6.5 Summary 154
6.6 Proofs 155

7 Conclusions and remarks **161**
7.1 Lessons learned 162
7.2 Future work 163

List of Figures

1-1 Example application using two flying robots to monitor a national park susceptible to forest fires. 20

1-2 Geometric-based multi-robot control for camera coverage. 21

1-3 Single- and multi-robot systems using mutual information-based control to monitor the state of a national park. 26

1-4 The research pillars and foundation that support the contributions of this thesis. 27

4-1 A hallway "dead end" example illustrating the geometric relevance of mutual information reward surfaces. 54

4-2 Dynamic Bayesian network for occupancy grid mapping. 57

4-3 The perceptual ranges for an example robot with a ten beam range sensor. 59

4-4 Dynamic Bayesian network for the two independent theoretical experiments. 63

4-5 An illustration of when a larger amount of entropy does not imply a greater mutual information reward. 66

4-6 An omnidirectional ground robot with 360 degree laser rangefinder for the hardware experiments. 80

4-7 The measurement distribution for the sensor model for various expected measurement values. 81

4-8 The measurement distribution gradient with respect to the robot's position for the sensor model for various expected measurement values. . 81

4-9 The evolution of the occupancy grid map and mutual information reward surface for occupancy grid mapping. 82
4-10 The evolution of the occupancy grid map and mutual information reward surface for SLAM. 83
4-11 The occupancy grid map, entropy map, and mutual information reward surface for SLAM. 84

5-1 A centralized multi-robot system within a probabilistic framework. . . 93
5-2 Dynamic Bayesian network for multi-robot inference tasks. 93
5-3 Robots moving their sensors' field of view towards a region of the environment that corresponds to high uncertainty. 94
5-4 A centralized multi-robot system with an ideal network. 97
5-5 Our approach to distributed inference and coordination. 99
5-6 Non-smooth features of the mutual information reward surface for the narrow beam-based sensor model. 102
5-7 Non-smooth characteristics of the narrow beam-based sensor model. . 103
5-8 Smooth characteristics of the divergent beam-based sensor model. . . 104
5-9 The sampling methodology for creating a robot's local observation sample set. 105
5-10 The sampling methodology for creating the joint observation sample set. 106
5-11 The joint observation sampling methodology using various decentralized Bayesian filters. 114
5-12 A snapshot of the autonomous deployment of five quadrotor flying robots for an outdoor hardware experiment. 115
5-13 Five quadrotor flying robots inferring the state of an indoor environment using a decentralized discrete Bayesian filter. 118
5-14 The decrease in entropy over time of a five robot experiment using a decentralized discrete Bayesian filter. 119
5-15 Five quadrotor flying robots inferring the state of an indoor environment using a decentralized non-parametric Bayesian filter. 121

5-16 The beginning, middle, and end configuration of a five robot experiment over a 10 cell environment. 128
5-17 Five quadrotor flying robots inferring the state of an outdoor environment using a decentralized non-parametric Bayesian filter. 129
5-18 Preliminary deployment in preparation for the outdoor experiment. . 129
5-19 The time evolution of the outdoor experiment. 130

6-1 An unweighted undirected network graph describing the communication capabilities of a three robot system. 133
6-2 One dimensional normalized local measurement likelihood distributions for ten robots with respect to their normalized joint. 149
6-3 The evolution of each robot's normalized joint approximation on a connected network graph. 150
6-4 The evolution of each robot's normalized joint approximation on a star network graph, a chain network graph, and a ring network graph. . . 151
6-5 The position evolution of a simulated system of 100 heterogeneous robots employing a decentralized standard Bayesian filter and a decentralized mutual information-based gradient-ascent controller. 152
6-6 Entropy plots for a simulated system of 100 robots employing decentralized Bayesian filters and decentralized mutual information-based gradient-ascent controllers. 153
6-7 The position evolution of a simulated system of 1000 heterogeneous robots employing a decentralized non-parametric Bayesian filter and a decentralized mutual information-based gradient-ascent controller. . . 154

List of Tables

3.1 Notation used in this thesis. 47

4.1 Parameters for the beam-based proximity mixture sensor model. . . . 79

5.1 Common parameters for the hardware experiments in Section 5.4. . . 117

5.2 Parameters for the indoor experiment in Sections 5.4.2 and 5.4.3. . . 120

5.3 Parameters for the outdoor experiment in Section 5.4.4. 122

List of Algorithms

1 MappingMutualInformationReward() 76

2 MappingMutualInformationRewardGradient() 77

3 BeliefCumulationProduct($\mathcal{I}_{k+1}^{[j]}$) 78

4 MeasurementPrior($j, \mathcal{I}_{k+1}^{[j]}, z, \eta$) 78

5 MeasurementPriorGradient($j, \mathcal{I}_{k+1}^{[j]}, z, \eta$) 79

6 DiscreteEnvironmentStateController($\check{p}_k^{[i]}$) 110

7 ContinuousEnvironmentStateController($\check{p}_k^{[i]}$) 111

8 SequentialImportanceResampling($\check{\mathcal{X}}_k^{[i]}$) 113

9 NonparametricEnvironmentStateController($\check{p}_k^{[i]}$) 113

10 MeasurementProbabilitiesConsensus($\check{\mathcal{Y}}^{[i]}$) 141

11 LikelihoodDistributionConsensus($\check{\mathcal{Y}}^{[i]}$) 145

Chapter 1

Introduction

Multi-robot systems will change the way we sense the world. Consider, for example, the problem of monitoring a national park susceptible to forest fires – a problem often addressed by flying manned aircraft over areas of concern to provide aerial surveillance (Rauste et al., 1999). This and other "traditional mechanisms" (San-Miguel-Ayanz et al., 2005) are human and resource intensive, and thus limited to periods marked with active fires or elevated risk (Kelha et al., 2001). A paradigm shifting solution materialized two decades ago when researchers began employing satellite technology to continuously survey vulnerable areas at a global scale (Justice and Korontzi, 2001). These "near realtime" solutions use a combination of low Earth orbit and geostationary satellites, with the former providing periodic high resolution imagery and the latter providing continuous yet lower resolution imagery (deGroot et al., 2007).

Given this state-of-the-art approach to aerial surveillance, naturally we ask: does there exist a more capable solution that enables both persistent *and* high resolution imagery, especially if the surveillance task is not conducted at a global scale but instead at the scale of, for example, a national park? Considering recent advancements in unmanned aerial vehicle (UAV) technology, e.g., ultra-light UAV imagery (Kung et al., 2011), the answer to our question is undoubtedly *yes* from a *hardware* perspective. However, we believe current limitations in *autonomy* are self-evident from the undeniable absence of fielded multi-robot systems with autonomous surveillance

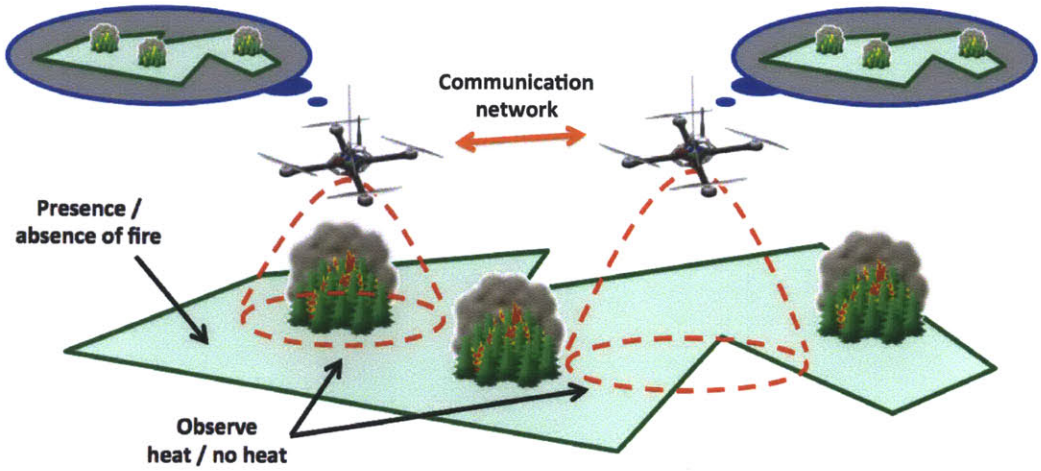


Figure 1-1: Example application using two flying robots to monitor a national park susceptible to forest fires. Each robot's observation provides information on whether or not heat exists within its field of view (red dashed circles), and the robots share this information on the communication network (orange arrow) to improve the quality of their beliefs (blue circles) of the environment state. In this thesis, we address how these robots can reason about where to position their sensors for future observations.

capabilities (Christensen, 2013). To reveal the next paradigm shifting solution to the forest fire and many other information acquisition problems, we must enable these multi-robot systems to efficiently, robustly, and provably learn the environment and autonomously reason where to make future sensor observations, e.g., see Figure 1-1.

To this end, we present the derivation, analysis, and implementation of a novel class of decentralized controllers that continuously move robots equipped with sensors to better observe their environment. Built on an information-theoretic foundation, these controllers aim to maximize the expected information gain of the robots' next sensor observations when paired with a Bayesian filter. They allow for general convergent results, and lead to practical control strategies that account for the limited computational resources of the robots, the decentralized nature of their computation, and the finite bandwidth of their communication network. The controllers are able to dynamically adapt to changing network connectivity and be scalable with respect to the number of robots, enabling robot teams of large size and heterogeneous composition. Lastly, they exhibit anytime performance when augmented with distributed algorithms to approximate the measurement likelihoods for system-wide observations.

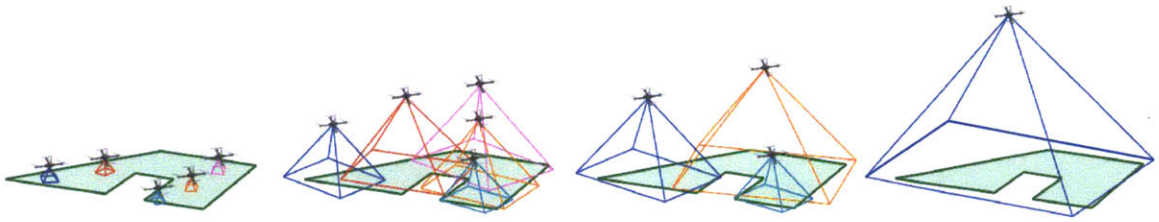


Figure 1-2: Geometric-based multi-robot control for camera coverage in the spirit of (Schwager et al., 2011b). Five flying robots are tasked to cover a bounded environment with the fields of view of their downward facing cameras. *Left to right:* The system of five robots i) is deployed, ii) converges to a locally optimal configuration, iii) compensates for the removal of two robots, and lastly iv) covers the entire environment with the single remaining robot.

This thesis aims to improve on the state-of-the-art approaches for controlling a network of sensing robots for information acquisition tasks. Most existing approaches employ optimization functions that directly relate the robots' positions to geometric features of the environment (Bullo et al., 2009). For example, the camera coverage approach in (Schwager et al., 2009) formulates a system-wide cost function that depends on the projection of the cameras' fields of view onto the environment, e.g., see Figure 1-2. By moving in the negative gradient of the cost function with respect to its own position, each robot employing this *geometric-based* controller reaches a final goal position that is of locally minimal cost with respect to the system-wide cost function. In fact, many geometric-based controllers owe their beginnings to a Voronoi-based approach introduced by Cortés et al. (2004), and can be considered solution approaches to variations on the same optimization problem (Schwager, 2009). Although we are greatly inspired by this unifying geometric-based approach to decentralized robot control, this thesis considers optimization functions that relate the robots' positions to the performance of Bayesian filters designed to infer the environment state.

The proposed *information-based* controllers in this thesis ensure that the resulting decentralized controllers take into account how the robots' next sensor observations improve on their past observations to resolve current uncertainty associated with the inference. As a result, we are interested in controllers that aim to maximize a mutual information reward function over a single step horizon. Mutual information[1] is an

[1]In this thesis, mutual information is always considered to be between the random variable

information-theoretic quantity (Cover and Thomas, 1991) that predicts how much future observations will decrease the robots' uncertainties associated with their inferred beliefs of the environment state. By moving in the direction of increasing mutual information reward, the robots provably increase the information gain resulting from their next joint observation. State-of-the-art mutual information-based controllers often evaluate the change in mutual information reward for a discrete set of control actions, then select the action resulting in the largest reward. Although this approach is shown to be optimal for greatly simplified cases (Castanon, 1993), the discretization of the control space can be somewhat arbitrary for real robots moving in continuous configuration spaces.

This thesis distinguishes itself from previous works in that the controllers are derived from the *analytical gradient* of the mutual information reward function. By distributively employing a gradient-ascent controller, the control actions are of vector form dictating the desired velocities within the robots' continuous configuration spaces. These control actions can drive any type of sensing robot platform employing probabilistic sensor models and accepting velocity command inputs. In other words, the *implementation* of these mutual information-based gradient-ascent controllers does not require the robot to have particular dynamics (e.g., integrator, second order) or sensors (e.g., passive power sensors, active range sensors).

Not surprisingly, the *performance* of these controllers do depend on these properties, and thus we focus on robots that have integrator dynamics and sensors with sufficiently smooth probabilistic models. With respect to integrator dynamics, one can often employ lower-level closed-loop controllers to emulate such dynamics, e.g., on quadrotors that are traditionally modeled with second ordered dynamics (Michael et al., 2010). With respect to sufficiently smooth sensor models, there exists many simple sensors (e.g., magnetometers, light sensors) that inherently have this property, especially in the absence of sensor obscuration. For example, an underwater robot may more accurately sense the distance to an acoustic source the closer the robot

representing the environment state and the random variable representing the robots' next joint observation.

is to that source, and the *change* in the corresponding measurement probabilities[2] is often continuous with respect to the robot's motion. When tasked to track these acoustic sources, a team of such underwater robots employing the controllers in this thesis would continuously balance the tendencies to move closer to the sources while managing environmental uncertainties (e.g., yet to be discovered sources) elsewhere within a bounded volume of water. Again, the robots' behaviors are driven by the goal of maximizing the informativeness of their next joint observation such that the *system* can best locate all sources. More complex sensors (e.g., cameras, laser rangefinders) may not inherently have sensor models that are sufficiently smooth, however, one can often derive different forms of the sensor models to induce smoothness. We employ such a technique to derive a smooth sensor model for laser rangefinders in this thesis.

Given these robot dynamics and probabilistic assumptions, the mutual information-based gradient-ascent controllers are shown to be convergent between sensor observations and, in their generalized form, locally optimal. Thus, the robots' *locally* calculated control actions are provably favorable from a *system-wide* perspective, which for our problem formulation encompasses stronger control-theoretic statements than can be made for the state-of-the-art mutual information-based controllers. For example, the performance of controllers that evaluate mutual information reward for a discrete set of control actions is often correlated to the size of the control set. On the other hand, directly calculating the gradient of the mutual information reward function avoids the need to discretize the robots' continuous control or configuration spaces, which weakens the controllers' optimality guarantees. In addition, since calculating the gradient of the mutual information reward function is equivalent in computational complexity to calculating one reward value, the ability to calculate the gradient versus a set of rewards for robot control results in a constant factor improvement in computational complexity.

To overcome the requirement of global knowledge required for mutual information-based calculations, we introduce sample-based methods to support the distributive

[2]We define the measurement probability as the probability of particular observation given a realized environment state.

approximation of system-wide probability distributions. These methods enable each robot to form non-parametric representations for measurement probability distributions[3] describing the next joint sensor observation, which can be realized from probability distributions of arbitrary form. For example, we do not assume that the statistics of the involved random variables are exactly described by particular distributions defined using a finite number of parameters. Such assumptions are commonplace in decentralized mutual information-based control, e.g., Kalman filter-based controllers employing the Decentralized Data Fusion architecture (Manyika and Durrant-Whyte, 1994). The use of non-parametric representations of the joint measurement distributions preserve aspects of the mutual information reward function and its gradient that are discarded by Gaussian and other parametric-based assumptions. In addition, this non-parametric formulation yields controllers that are compatible with any type of Bayesian filter, where state-of-the-art controllers are often designed for specialized filters assuming particular distributions.

Given the resulting set of likely joint observations, we show that the measurement likelihoods[4] required for the controller calculations can be approximated using a variety of distributed algorithms. In fact, the joint measurement likelihood calculations are decoupled from the controller calculations. We show that this "modular" approach to the likelihood and controller calculations allows for scalable implementations with respect to the number of robots. We then give two example consensus-based algorithms that run on an undirected network graph. The first consensus-based algorithm approximates discrete measurement probabilities for discrete environment state random variables, while the second approximates continuous likelihood distributions for continuous environment state random variables. We show that these anytime approximations provably converge to the correct values on a static and connected graph without knowledge of the number of robots in the network or the corresponding graph's topology.

[3]We define the measurement probability distribution as the probability distribution with respect to all observations given a realized environment state.

[4]We define the measurement likelihood as the probability of an observation with respect to all possible environment states.

Lastly, we incorporate the resulting consensus-based algorithms into both a hardware system and a simulation environment to allow for controller evaluation under non-ideal network settings. For the hardware experiments, the task is to infer the state of a bounded, planar environment by deploying five quadrotor flying robots with simulated sensors in both indoor and outdoor settings. For all experiments, the controllers are ran in distributed fashion either on a single ground computer or on multiple onboard computers, with the latter resulting in a completely decentralized hardware system. For the numerical simulations, Monte Carlo-based analyses are performed for 100 robots, where each robot is simulated on an independent computer node within a computer cluster system. By incorporating a message passing protocol between simulated robots on independent nodes, the controllers again ran in completely distributed fashion. Simulations are also performed for 1000 robots using a single workstation computer equipped with a multicore GPU-enabled graphics card. The results from both the hardware experiments and numerical simulations validate our theoretical and computational claims throughout the thesis.

1.0.1 Aerial surveillance use case

Let us revisit the task of monitoring a national park susceptible to forest fires using flying robots. For such aerial surveillance, it is common to employ downward facing infrared cameras to detect elevated heat signatures from active fires. Prior to deployment, these cameras are calibrated to characterize their noise characteristics in support of building a belief of the national park's state, i.e., the Bayesian inference. It is the resulting sensor models that link the positions of the robots to the accuracy associated with any heat signature reading.

As shown in Figure 1-3, the controllers in this thesis are applicable to a single robot, continuously driving it along the positive gradient of the mutual information between the state of the national park and the next heat signature reading. If the robot is restricted to move translationally in two-dimensions, it would move towards areas of high uncertainty with respect to the Bayesian inference. In other words, the more the robot is unsure of the existence of fire at a particular location within

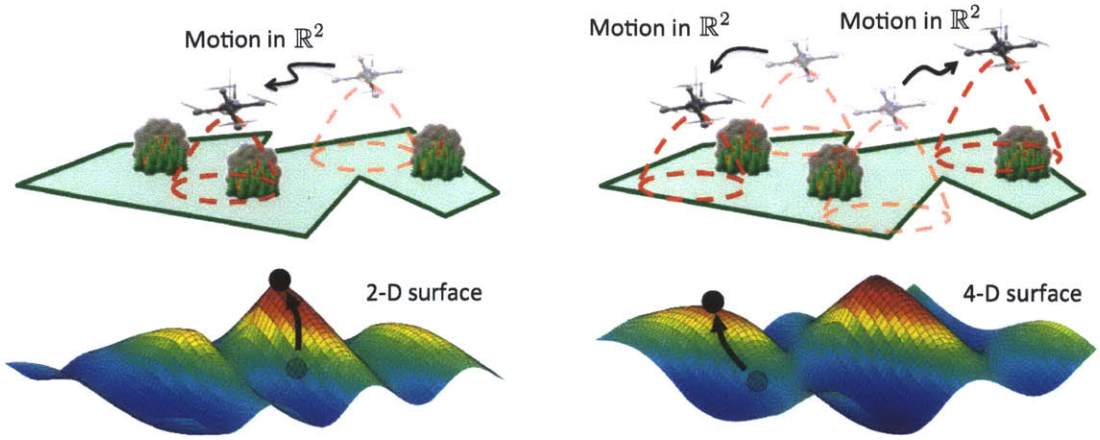


Figure 1-3: Single- and multi-robot systems using mutual information-based control to monitor the state of a national park. *Left:* By following the positive gradient of the mutual information between the state of the national park and the next sensor observation, a single robot positions its sensor in a globally optimal fashion, although such a controller only guarantees local optimality. *Right:* By navigating on a higher dimensional mutual information reward surface, two robots position their sensors in a similar manner.

the national park, the more attractive the overhead airspace becomes. The resulting motion in physical space can be visualized in information space as a single point ascending a two-dimensional mutual information reward surface.

To increase the pervasiveness of this aerial surveillance system, additional robots of varying sensors can be used, e.g., downward facing optical cameras. The decentralized mutual information-based gradient ascent controllers remain applicable even though every robot's sensor model may differ. Much like the single robot system, the joint motion of the robot team can again be visualized in information space as a single point ascending a mutual information reward surface, however, the surface is now higher dimensional. Regardless, the robots are continuously attracted to areas of high uncertainty, maximizing the informativeness of their next joint observation.

1.1 Technical approach

Our technical approach focuses on three research pillars that support the novel class of decentralized mutual information-based gradient-ascent controllers (Figure 1-4).

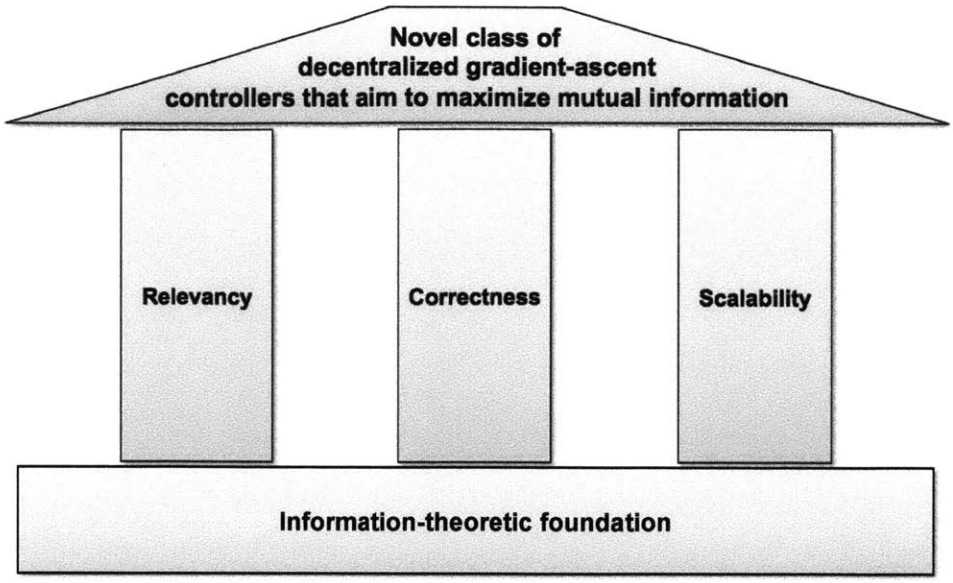


Figure 1-4: The research pillars and foundation that support the contributions of this thesis.

We now describe each of these pillars in greater detail.

i) **Relevancy:** We show that maximizing the mutual information between the environment state and the robots' next joint observation is *relevant* to robot inference tasks, e.g., occupancy grid mapping (Chapter 4). We also show that multi-robot systems employing sensors whose noise characteristics smoothly depend on the robots' positions can be cast in terms of an optimization problem to maximize a mutual information reward function (Chapter 5). A decentralized gradient-ascent solution approach to this optimization problem yields the novel controller class.

ii) **Correctness:** We prove that any member of the novel controller class is convergent between sensor observations, i.e., the velocities of all robots converge to zero given enough time between sensor observations (Chapter 5). We considered this property to be *correct* from a control-theoretic perspective since it rigorously characterizes how the robots improve their positions prior to the next update of the sequential Bayesian filter. In the generalized case when no sample-based approximations are made in support of the controller, we prove that the resulting goal positions are locally optimal.

iii) **Scalability:** We implement several members of the novel controller class and show that these algorithms *scale* to large multi-robot teams (Chapter 5). More specifically, the computational complexities of these decentralized controllers are independent of the number of robots. To induce scalability on non-ideal communication networks, distributed algorithms are used to approximate sample sets of measurement likelihoods describing system-wide observations (Chapter 6). These anytime approximations provably converge to the correct values as more computational resources are allocated for the multi-robot system.

These three research pillars are built upon an information-theoretic foundation that supports many other robot controller classes (Chapter 2). Throughout this thesis, we validate our theoretical and computational claims through numerical simulations and hardware experiments concerning traditional robot applications such as mapping, exploration, and surveillance.

1.1.1 Scope and limitations

Our technical approach is built on several assumptions that focus the scope of this thesis. In the following list, we detail the assumptions made in each chapter, highlighting the limitations they may introduce. Note that assumptions do not carry over to other chapters unless explicitly stated.

i) **Mutual information-based control for mapping applications (Chapter 4):** We assume that a single robot equipped with a range sensor employs the occupancy grid mapping algorithm. For simplicity, we focus on constructing a map in two dimensions using a narrow beam-based sensor model, but hypothesize that the results are valid for higher dimensions and divergent beam-based sensor models. We also assume that the sensor model is unbiased and the measurement's prior is clamped. Intuitively speaking, an unbiased sensor model ensures that a correct map will be built in expectation, while a clamped measurement's prior prevents overconfidence with respect to the sensor calibration process.

ii) **Gradient-ascent control for multi-robot inference tasks (Chapter 5):** We assume that each robot equipped with a decentralized Bayesian filter has simple dynamics where the control input is the robot's velocity, i.e., the robots have integrator dynamics. We also assume that each robot knows its own position with certainty, as well as the physical extent (i.e. boundary) of the environment and the set of possible values (i.e., alphabet) for the random variable representing the environment state. Note that if probabilistic localization is employed, each robot can choose one hypothesized position as input to its controller, e.g., the maximum a-posteriori position. In addition, we assume that the sensor is calibrated and of finite range, with the former implying that each robot knows its local sensor model, which is also assumed to be sufficiently smooth. A robot's position and sensor model do not need to be known by any other robot.

iii) **Distributed algorithms enabling decentralized inference and coordination (Chapter 6):** We again assume that the local sensor observations are synchronous and conditionally independent between all robots. The former implies that all robots simultaneously perform Bayesian filter calculations at a constant rate. The latter implies that given the environment state, the errors on the local observations are uncorrelated between robots. To formalize the convergence of consensus-based algorithms, we assume that the robots' communication is governed by an unweighted and undirected network graph that is static and time-invariant between consecutive joint observations. These last assumptions combined with the existence of an upper bound on each robot's in/out degree form the communication model employed in Chapter 6.

1.2 Research contributions

The main research contributions of this thesis are as follows.

i) We prove that any occupancy grid mapping controller tasked to maximize a mutual information reward function is eventually attracted to unexplored space.

This proof relies on a commonly employed beam-based sensor model versus a more abstract additive white Gaussian noise model.

ii) We develop an analytical approach that allows for the identification of mutual information-based behaviors for occupancy grid mapping. This approach allows us to identify the effect of changing the belief of one or more cells while keeping all other aspects (e.g., robots' positions, realized map) the same between two theoretical situations. We believe this approach can be used to prove many future claims concerning control for more general inference tasks.

iii) We provide a computationally tractable algorithmic implementation of the mutual information reward function. More specifically, the time and space complexities of this algorithm are at worst quadratic and linear, respectively, with respect to the occupancy grid map's spatial resolution. In contrast, a naïve implementation ignoring the assumption of sensor obscuration has a time complexity that is exponential, i.e., is computationally intractable.

iv) We apply the gradient of the mutual information reward function to control robots to make informative measurements for such tasks such as exploration and surveillance. The computation engine for this controller is generalizable to many other Bayesian inference tasks.

v) We provide non-parametric sample-based methods for representing the robots' likely observations to enable decentralized coordination. The resulting approximation for the gradient of the mutual information reward function forms the basis for the novel class of decentralized gradient-ascent controllers.

vi) We present novel consensus-based algorithms for approximating the robots' joint measurement likelihoods in both a discrete and continuous setting. We prove that these approximations converge to the true values as more system resources are allocated or as the network graph becomes complete. We also formalize rates of convergence and information-based bounds for the continuous case.

vii) We provide the results of numerous quadrotor hardware experiments in both indoor and outdoor environments, one of which is conducted on a completely decentralized hardware system. We also give the results of numerical simulations using as many as 1000 robots.

1.2.1 Publications

This thesis is a culmination of the following journal articles, conference proceedings, and workshop abstracts.

i) B. J. Julian, S. Karaman, and D. Rus. On mutual information-based control of range sensing robots for mapping applications. *The Int. J. of Robotics Research*, 2013. Submitted.

ii) B. J. Julian, S. Karaman, and D. Rus. On mutual information-based control of range sensing robots for mapping applications. In *Proc. IEEE/RSJ Int. Conference on Intelligent Robots and Syst.*, Tokyo, Japan, Nov 2013. Accepted.

iii) B. J. Julian, M. Angermann, and D. Rus. Nonparametric inference and coordination for distributed robotics. In *Proc. IEEE Conference on Decision and Control*, Grand Wailea, HI, USA, Dec 2012.

iv) B. J. Julian, M. Angermann, M. Schwager, and D. Rus. Distributed robotic sensor networks: an information-theoretic approach. *The Int. J. of Robotics Research*, 31(10):1134–1154, Sep 2012.

v) B. J. Julian, S. L. Smith, and D. Rus. Distributed approximation of joint measurement distributions using mixtures of Gaussians. In *Proc. Robotics: Sci. and Syst. Conference*, Sydney, Australia, Jul 2012.

vi) B. J. Julian, M. Angermann, M. Schwager, and D. Rus. A scalable information theoretic approach to distributed robot coordination. In *Proc. IEEE/RSJ Int. Conference on Intelligent Robots and Syst.*, San Francisco, CA, USA, Sep 2011.

vii) B. J. Julian, M. Angermann, M. Frassl, M. Lichtenstern, and D. Rus. Towards a unifying information theoretic framework for multi-robot exploration and surveillance. In *RSS Workshop on 3D Exploration, Mapping, and Surveillance with Aerial Robots*, Los Angeles, CA, USA, Jun 2011.

1.3 Thesis outline

This thesis is organized as follows. The next two chapters provide background material essential to the writing of this thesis, starting in Chapter 2 with a discussion of prior works from three influential research areas. We follow with a discussion of mathematical notation and definitions in Chapter 3.

In Chapter 4, we introduce the well-known occupancy grid mapping algorithm, providing in Section 4.1 the formal statement that any controller tasked to maximize a mutual information reward function is eventually attracted to unexplored space. In Section 4.2, we rigorously prove this result using numerous lemmas and corollaries that are interesting in their own right. The proof for attraction hinges on the construction of an analytical approach employing two independent theoretical experiments. In Section 4.3, we formally derive the analytical and discretized expressions for the mutual information reward function and its gradient, then discuss the computational complexity of their algorithmic implementations. We conclude the chapter by discussing the results from two separate hardware experiments in Section 4.4.

In Chapter 5, we formalize the general multi-robot inference and coordination problem, deriving in Section 5.1 an expression for the analytical gradient of mutual information between the environment state and the robots' next joint sensor observation. In Section 5.2, we formulate a mutual information-based gradient-ascent controller, then apply non-parametric sampling methods to form the novel class of decentralized controllers. Several algorithmic implementations of these controllers are given in Section 5.3, from which two are realized in hardware experiments in both an indoor and outdoor setting. We conclude the chapter by discussing the results from these experiments in Section 5.4.

In Chapter 6, we discuss distributed algorithms that enable decentralized inference and coordination. We formalize our distributed robot network in Section 6.1, specifically the employed communication model and properties associated with the underlying network graph. We then provide consensus-based approaches that enable distributed approximations for discrete measurement probabilities (Section 6.2) and continuous measurement likelihood distributions (Section 6.3). Algorithmic implementations of these two approaches are also provided, allowing for numerical simulations to be distributively run on a computer cluster system. We conclude the chapter by discussing the results from these parallelized numerical simulations in Section 6.4.

Lastly, we conclude the thesis in Chapter 7, giving final thoughts, lessons learned, and future work.

Chapter 2

Prior works

This thesis builds on prior works from the research areas of i) occupancy grid mapping and simultaneous localization and mapping (SLAM); ii) decentralized inference and coordination; and iii) distributed algorithms for multi-robot systems. We discuss in this chapter the prior works from all three research areas, noting that there is considerable overlap between the areas. The material in this chapter does not constitute novel research contributions by this thesis.

2.1 Occupancy grid mapping and SLAM

The seminal work of Moravec and Elfes (1985) introduced the concept of occupancy grid maps as a space representation model. Since then, much research has focused on constructing maps in higher dimensions, fusing measurements from multiple sensors/robots, and simultaneously/concurrently localizing. With respect to constructing maps in higher dimensions, Moravec and Martin (1994) used stereo vision to extend occupancy grid maps to three-dimensions, with Johnson and Kang (1999) improving on this extension by using a modified iterative closest point algorithm to merge multiple textured data sets. Inspirational to our focus on multi-robot systems, Thrun et al. (1998a) first introduced multi-sensor fusion for occupancy grid mapping, which lead to multiple approaches for multi-robot exploration, e.g., (Simmons et al., 2000; Burgard et al., 2000). Also introduced by Thrun et al. (1998c,b)

was the concept of incorporating topological mapping approaches (Smith and Cheeseman, 1986; Moutarlier and Chatila, 1989; Smith et al., 1990; Kuipers and Byun, 1991; Leonard and Durrant-Whyte, 1991) alongside occupancy grid mapping for SLAM applications.

Unlike the works cited in the previous paragraph, the purpose of Chapter 4 is not to propose new occupancy grid-based mapping algorithms. Instead, we focus on conducting information-theoretic analyses using these preexisting algorithms for the underlying Bayesian inference framework. For example, we evaluate our theoretical results in Section 4.4.3 using a Monte Carlo-based SLAM algorithm (Montemerlo et al., 2002) that employs the occupancy grid mapping algorithm. We believe that our analyses can be applied to many other Bayesian derived mapping algorithms where sensor noise depends on the robot's position in the environment, such as those relying solely on object-based (Chatila and Laumond, 1985) or landmark-based (Smith and Cheeseman, 1986) maps.

The task of maximizing a mutual information reward function has recently emerged as a powerful objective for controlling a robot to improve the quality of the occupancy grid map. The early work of Elfes (1995) proposed employing mutual information between sensor observations and the map as a reward function for general information acquisition tasks. Specific to SLAM exploration, Bourgault et al. (2002) combined this reward function with the mutual information between sensor observations and robot localization. Many following works focused on mutual information-based control for robot localization (Kröse et al., 2001) and joint robot/target localization (Grocholsky, 2002; Stump et al., 2009). Specific to frontier identification, Rocha et al. (2005) used a decentralized gradient-based approach for selecting boundaries between explored and unexplored space. The explorative behaviors reported by this work are well described by the results in Chapter 4, as are the behaviors from many other mutual information-based approaches (Stachniss, 2006; Visser and Slamet, 2008; Amigoni and Caglioti, 2010; Bhattacharya et al., 2010).

Chapter 4 also gives insight into a larger class of uncertainty-driven controllers enabling occupancy grid map building. Many early works used entropy-based methods

to again improve target and/or robot localization, see (Burgard et al., 1997; Whaite and Ferrie, 1997; Roy et al., 1999). With respect to uncertainty associated with a map, Elfes (1995) suggested using entropy to quantify the uncertainty associated with all grid cells. At the same time, Cassandra et al. (1996) proposed using information maximization methods to drive robot navigation. These contributions were followed by Moorehead et al. (2001) who computed the entropy of all cells to determine the utility of future measurements. Not surprisingly, these (if not all) exploration approaches for mapping are benchmarked against the traditional geometric approach of frontier exploration first proposed by Yamauchi (1998).

2.2 Decentralized inference and coordination

Concerning state estimation of an environment, the body of work addressing Bayesian estimation methods based on a family of Kalman filters, which are Bayesian filters for linear Gaussian systems, have commonly been used. For example, Lynch et al. (2008) proposed a distributed Kalman filtering approach in which the robots use consensus-based algorithms to share information while controlling their positions to decrease the error variance of the state estimate. In addition, Cortés (2009) developed a distributed filtering algorithm based on the Kriged Kalman filter for estimating environmental fields. The algorithm also estimated the gradient of the field, which is then used for multi-robot control. There have been similar Kalman filter approaches for tracking multiple targets, such as in (Chung et al., 2004). Although research in Kalman-based filtering has provided many advances concerning Bayesian inference, we emphasize that this thesis does not restrict itself to linear Gaussian systems. Moreover, the non-parametric representations we employ are compatible with any Bayesian filter approach, including Kalman-based ones.

Concerning non-parametric representations, the use of sample-based filters have become popular in robotics as the platforms become more computationally capable. In an early work, Engelson and McDermott (1992) used a sequential Monte Carlo method to construct an algorithm robust enough to address the kidnapped robot

problem. Since then, non-parametric methods have become commonplace in localization (Borenstein et al., 1997), target tracking (Schulz et al., 2001), and SLAM (Montemerlo et al., 2002). Fox et al. (2000) applied these methods to multiple collaborating robots using a sample-based version of Markov localization. Similar to the work in Chapter 5 are the recent efforts of Hoffmann and Tomlin (2010) who proposed a sequential Monte Carlo method to propagate a Bayesian estimate of the environment, then used greedy and pair-wise approximations to calculate mutual information. In addition, Belief Propagation (Pearl, 1988) has seen non-parametric extensions (Ihler et al., 2005), which use Gaussian mixtures to solve graphical inference problems.

Concerning mutual information-based control, the early work of Cameron and Durrant-Whyte (1990) used mutual information as a reward function for sensor placement without explicitly considering the mobility of the sensors. Later Grocholsky et al. (2003) proposed controlling multiple robot platforms near an object of interest so as to increase mutual information in tracking applications. Bourgault et al. (2002) used similar methods for exploring and mapping uncertain environments, which relates to our work in Chapter 4. In addition, the difficult problem of planning paths through an environment to optimize mutual information has been recently investigated, most notably the work by Singh et al. (2007, 2009), Ny and Pappas (2009), Choi and How (2010), and Ryan and Hedrick (2010). Once again, all works mentioned in this paragraph assume some form of a Gaussian process, a control assumption we do not make in this thesis.

Regarding the formulation of the novel class of controllers in this thesis, our work (Julian et al., 2011b) in collaboration with Schwager et al. (2011a) demonstrates the first results in robotics on using the analytically derived expression for the gradient of mutual information.[1] Given the long lineage of information-theoretic control approaches, we believe that the inherent mathematical complexity of the mutual information expression deterred the robotics research community from analytically evaluating its gradient. Therefore, we believe the closest related works are ones that

[1]Note that similar efforts have previously been discussed for applications such as image alignment (Viola, 1995) and channel coding (Palomar and Verdú, 2007)

evaluate the change in mutual information for a discrete set of control actions, e.g., (Grocholsky, 2002; Bourgault et al., 2002; Rocha et al., 2005; Stachniss, 2006; Hoffmann and Tomlin, 2010). Since publishing (Julian et al., 2011b), the work by Dames et al. (2012) has explored using mutual information-based gradient-ascent control when considering hazardous environments, as well as methods to improve controller performance when a centralized server and/or cloud computing is accessible (Dames and Kumar, 2013).

Although we do not explicitly discuss the task of robot localization (i.e., each robot inferring its own configuration within a global frame) and map alignment, the problem of decentralized inference is related to the multi-robot SLAM problem. Most relevant is the work of Leung et al. (2012), which proposed a decentralized SLAM approach able to obtain "centralized-equivalent" solutions on non-complete communication graphs. This approach is sufficiently generic to employ a wide range of Bayesian filtering methods, contrary to prior works specifically using extended Kalman filters (Nettleton et al., 2000), sparse extended information filters (Thrun and Liu, 2005), or particle filters (Howard, 2006). For our work, incorporating the robots' configurations and map offsets into the robots' beliefs has direct benefits for multi-robot SLAM exploration. For a more in-depth discussion concerning SLAM and multi-robot SLAM, please see (Thrun et al., 2005).

2.3 Distributed algorithms for multi-robot systems

In a multi-robot context, the main challenges in using mutual information for control are computational complexity and network communication constraints. As previously stated, the complexity of computing mutual information and its gradient is exponential with respect to the number of robots, and thus is intractable in realistic applications using a large multi-robot team. Furthermore, the computation of mutual information requires that every robot has current knowledge of every other robot's position and sensor measurements. Thus, many of the prior mutual information methods are restricted to small groups of robots with all-to-all communication

infrastructure. To relax this communication requirement with respect to Bayesian distributed hypothesis testing, Olfati-Saber et al. (2005) developed a consensus algorithm with provable convergence guarantees when run on a static sensor network. In Chapter 6, we develop a consensus-based algorithm inspired by Olfati-Saber et al. to compute the joint measurement probabilities needed for the mutual information based controller and the Bayesian filter calculations.

We have also been inspired by over two decades worth of advancements in distributed estimation algorithms. An approach to compute locally optimal estimators from many independent sensor measurements at a central fusion center was described in detail by Gubner (1993). Concerning decentralization, the early work of Durrant-Whyte et al. (1990) with decentralized Kalman filters laid the basis for the Decentralized Data Fusion architecture (Manyika and Durrant-Whyte, 1994). Extensions incorporating consensus-based algorithms (Cortés, 2008; Xiao et al., 2007) have been used for maximum-likelihood parameter estimation (Xiao et al., 2005), maximum a-posteriori estimation (Olfati-Saber et al., 2005), and distributed Kalman filtering (Alighanbari and How, 2006; Yang et al., 2007).

One of the most relevant works in distributed Kalman filtering is by Ren et al. (2005), who showed the convergence of a filter incorporating information-based states. The proof of convergence for the Gaussian parameters for the joint distribution approximation closely follows in the work in Chapter 6, even though our algorithms apply to a larger class of Bayesian filters, such as the work by Aragués et al. (2010) supporting map merging. This generality is shared by the approach of Fraser et al. (2012) using hyper-parameters. However, our work enables the early termination of the consensus-based algorithm without the risk of 'double-counting' any single observation, even when the maximum in/out degree and the number of robots are unknown.

Lastly, the positioning of the robots influences the communication properties of the system. Hence, Krause et al. (2008) formalized the task of balancing the informativeness of sensor placement with the need to communicate efficiently. This task was shown to be an "NP-hard tradeoff," which motivated the development of

a polynomial-time, data-driven approximate algorithm for choosing sensor positions. Closely related is the work of Zavlanos and Pappas (2008), which describes a connectivity controller that enables the robots to remove communication links with network neighbors while maintaining global connectedness. The controller uses local knowledge of the network to estimate its topology, rendering the algorithm distributed among robots.

Although in our work we analyze the performance of our algorithms assuming global connectedness (but not completeness), we note that our approach naturally accommodates the splitting and merging of network subgraphs by correctly fusing the beliefs of the involved robots. This property also facilitates the use of communication schemes that lack the notion of active connectivity maintenance. For example, we have previously presented how state estimates can be efficiently exchanged within a robot team by broadcasting estimates of nearby robots more frequently than distant ones (Julian et al., 2009).

2.4 Summary

In this chapter we discussed prior works relevant to the contributions of this thesis. In the coming chapters, we i) analyze the relevance of mutual information-based control to occupancy grid mapping and SLAM; ii) derive a novel decentralized gradient-ascent controller class; and iii) present distributed algorithms to induce controller scalability. However, we first introduce in Chapter 3 the mathematical notation and definitions used throughout the thesis.

Chapter 3

Background

This chapter gives the mathematical notation and definitions used throughout the thesis. The material in this chapter does not constitute novel research contributions by this thesis.

3.1 Mathematical notation

Throughout this thesis, we make use of the following notation.

Notation	Description
t	Time (continuous)
k	Time (discrete step)
κ	Time (communication round)
$p_k^{[i,\ell]}$	Approximation for the ith robot's joint ℓth measurement likelihood at time step k
$p_k^{[i]}$	Approximation for the ith robot's joint measurement likelihoods at time step k
$\check{p}_k^{[i,\ell]}$	Approximation for the ith robot's sampled joint ℓth measurement likelihood at time step k
$\check{p}_k^{[i]}$	Approximation for the ith robot's sampled joint measurement likelihoods at time step k

Notation	Description
$r_k^{[i]}$	Belief of the ith occupancy grid cell at time step k
$O(\cdot)$	Big O notation
$\mathcal{A}_K$	Conditioning event for the independent theoretical experiments at time step k
$\mathcal{B}_{K,k}$	Conditioning event for the independent theoretical experiments at time steps k and k
$\psi_\kappa^{[i]}$	Consensus state for the ith robot at communication round κ
$\psi^{[i]}$	Consensus state for the ith robot at communication round n_π
$\pi_\kappa^{[i]}$	Consensus state (exponential) for the ith robot at communication round κ
$\pi^{[i]}$	Consensus state (exponential) for the ith robot at communication round n_π
$u_t^{[i]}$	Control action of ith robot at time t
u_t	Control action of multi-robot system at time t
$\gamma^{[i]}$	Control gain of ith robot
$\rightarrow$	Converges to
$\mathrm{diam}(\cdot)$	Diameter of a network graph
$H(\cdot)$	Entropy of a random variable
$[\cdot]_j$	Entry in jth position of a tuple or vector
$[\cdot]_{jk}$	Entry in (j,k) position of a matrix
$\mathbb{E}(\cdot)$	Expectation of a random variable
$\beta^{[i]}$	Exponential factor for the ith robot's approximated joint measurement likelihoods
$\vartheta_\kappa^{[i]}$	FloodMax state for the ith robot at communication round κ
$\vartheta^{[i]}$	FloodMax state for the ith robot at communication round n_π
$\mathcal{N}_c(\cdot,\cdot)$	Gaussian function defined using canonical parameters
$\mathcal{N}(\cdot,\cdot)$	Gaussian function defined using normal parameters
E	Independent theoretical experiment (first)
$\tilde{E}$	Independent theoretical experiment (second)

Notation	Description
$\mathbf{\Omega}^{[i]}$	Information matrix for the ith robot's measurement likelihood distribution of scaled Gaussian form
$\mathbf{\Omega}$	Information matrix for the joint measurement likelihood distribution of scaled Gaussian form
$\xi^{[i]}$	Information vector for the ith robot's measurement likelihood distribution of scaled Gaussian form
ξ	Information vector for the joint measurement likelihood distribution of scaled Gaussian form
$\delta^{[i,j]}(\cdot)$	Inverse sensor model for the jth range measurement at the ith occupancy grid cell
$\delta^{[i]}(\cdot)$	Inverse sensor model for the range sensor observation at the ith occupancy grid cell
$c_t^{[i]}$	Known position of ith robot at time t
c_t	Known position of multi-robot system at time t
$c_k^{[i]}$	Known position of ith robot at time step t
c_k	Known position of multi-robot system at time step k
$\int(\cdot)$	Lebesgue integral
$\max(\cdot)$	Maximum valued element of a set
$\min(\cdot)$	Minimum valued element of a set
$I(\cdot\,;\cdot)$	Mutual information between two random variables
$\widehat{I}(\cdot\,;\cdot)$	Mutual information numerical approximation between two random variables
$\log(\cdot)$	Natural logarithm
$\mathcal{G}$	Network graph of the multi-robot system
$\mathbf{A}$	Network's adjacency matrix
Δ	Network's maximum in/out degree
$\mathbf{W}$	Network's Metropolis-Hastings matrix
n_m	Number of cells in the occupancy grid map
n_π	Number of communication rounds in a single time step

Notation	Description
n_z	Number of measurements per range sensor observation
$\Delta^{[i]}$	Number of neighbors for the ith robot
n_r	Number of robots in the multi-robot system
$\mathcal{I}_k^{[j]}$	Perceptual range of the jth measurement at time step k
$\mathcal{I}_k$	Perceptual range of the sensor observation at time step k
$\mathbb{P}(\cdot)$	Probability measure
$R_k^{[i]}$	Random variable for the belief of the ith occupancy grid cell at time step k
X_k	Random variable for the environment state at time step k
$Y_k^{[i]}$	Random variable for the ith robot's sensor observation at time step k
Y_k	Random variable for the joint sensor observation at time step k
$\check{Y}_k^{[i]}$	Random variable for the sampled ith robot's sensor observation at time step k
$\check{Y}_k$	Random variable for the sampled joint sensor observation at time step k
$M^{[i]}$	Random variable for the ith occupancy grid cell
M	Random variable for the occupancy grid map
$Z_k^{[j]}$	Random variable for the jth range measurement at time step k
Z_k	Random variable for the range sensor observation at time step k
U_k	Reward function for robot control at time step k
$\partial\check{U}_k^{[i]}$	Reward function approximation for the ith robot at time step k
$\widehat{\partial U}_k^{[i]}$	Reward function numerical approximation for the ith robot at time step k
$\rho^{[i]}$	Scaling factor for the ith robot's measurement likelihood distribution of scaled Gaussian form
ρ	Scaling factor for the joint measurement likelihood distribution of scaled Gaussian form
$\mathbb{S}$	Set denoting a sphere

Notation	Description
$N^{[i]}$	Set of neighbors for the ith robot
$\mathcal{U}^{[i]}$	Set of possible control actions for the ith robot
$\mathcal{U}$	Set of possible control actions for the multi-robot system
$\mathcal{X}$	Set of possible environment states
$\mathcal{Y}^{[i]}$	Set of possible ith robot's sensor observations
$\mathcal{Y}$	Set of possible joint sensor observations
$\check{\mathcal{Y}}_k^{[i]}$	Set of sampled ith robot's sensor observations at time step k
$\check{\mathcal{Y}}_k$	Set of sampled joint sensor observations at time step k
$\mathcal{M}$	Set of possible occupancy grid maps
$\mathcal{Z}$	Set of possible range measurements
$\mathbb{R}$	Set of real numbers
$\mathbb{Z}$	Set of integers
$\prod(\cdot)$	Set product
$\sum(\cdot)$	Set summation
n_x	Size of sample set of environment states
n_y	Size of sample set of likely observations
λ_z	Spatial resolution of the numerical approximation of mutual information
λ_x	Spatial resolution of the numerical approximation of mutual information for a continuous-valued environment state
λ_m	Spatial resolution of the occupancy grid map
$\text{int}(\cdot)$	Subset of interior points
T_s	Time interval between consecutive observations for the multi-robot system
$\mathbf{e}_i$	Tuple for ith Euclidean basis
$\mathbf{1}$	Tuple of all ones
$\mathbf{0}$	Tuple of all zeros

Table 3.1: Notation used in this thesis.

3.2 Mathematical preliminaries

Let $x \in \mathcal{S}$ denote an element x belonging to a set $\mathcal{S}$, while $\{x \in \mathcal{S} : P(x)\}$ denotes the set of elements belonging to $\mathcal{S}$ with the assertion $P(x)$ being true. If $\mathcal{S}$ is finite, let $|\mathcal{S}|$ denote the number of its elements. Let $\text{int}(\mathcal{S})$ denote the interior of the set $\mathcal{S}$. If $\mathcal{R}$ is a subset of or equal to $\mathcal{S}$, then we write $\mathcal{R} \subset \mathcal{S}$. Given two sets $\mathcal{S}^{[1]}$ and $\mathcal{S}^{[2]}$, let $\mathcal{S}^{[1]} \cup \mathcal{S}^{[2]}$, $\mathcal{S}^{[1]} \cap \mathcal{S}^{[2]}$, and $\mathcal{S}^{[1]} \times \mathcal{S}^{[2]}$ denote their union, intersection, and Cartesian product, respectively. Note that indices are written as superscript within square brackets (multiple indices will be separated with commas). Let superscripts without square brackets denote Cartesian powers, e.g., $\mathcal{S}^3 = \mathcal{S} \times \mathcal{S} \times \mathcal{S}$. Given a collection of sets $\{\mathcal{S}^{[i]}\}_{i \in \mathcal{I}}$ indexed by a set $\mathcal{I}$, let $\bigcup_{i \in \mathcal{I}} \mathcal{S}^{[i]}$, $\bigcap_{i \in \mathcal{I}} \mathcal{S}^{[i]}$, and $\prod_{i \in \mathcal{I}} \mathcal{S}^{[i]}$ denote their indexed union, intersection, and Cartesian product, respectively.

Let $>$ and $\geq$ denote greater than and greater than or equal to, respectively, for two real numbers. We mean for these symbols to be applied element-wise for two vectors, while for two matrices we mean positive-definiteness and nonnegative-definiteness, respectively. Let $\mathbb{Z}$ and $\mathbb{R}$ denote the set of all integers and real numbers, respectively, with $\mathbb{Z}_{\geq 0}$, $\mathbb{Z}_{>0}$, $\mathbb{R}_{\geq 0}$, and $\mathbb{R}_{>0}$ being the sets of nonnegative integers, positive integers, nonnegative real numbers, and positive real numbers, respectively. Given $d \in \mathbb{Z}_{>0}$, let $\mathbb{R}^d$ and $\mathbb{S}^d \subset \mathbb{R}^{d+1}$ denote the d-dimensional Euclidean space and the d-dimensional sphere, respectively. Given $x, y \in \mathbb{R}$, let (x, y) and $[x, y]$ denote the open and closed interval, respectively, between x and y.

3.2.1 Probability theory

Given a sample space Ω, a σ-algebra $\mathcal{F} \subset 2^\Omega$, and a probability measure $\mathbb{P}$, let the triple $(\Omega, \mathcal{F}, \mathbb{P})$ denote a probability space. Given a random variable X that maps Ω to an alphabet $\mathcal{X}$, i.e., $X : \Omega \to \mathcal{X}$, let $\mathbb{P}(X = x)$ denote the probability or probability density when X takes the value $x \in \mathcal{X}$. For convenience, let $\mathbb{P}(X)$ be shorthand for $\mathbb{P}(X = x)$ with $x \in \mathcal{X}$. Given a second random variable $Y : \Omega \to \mathcal{Y}$, let $\mathbb{P}(X, Y)$

denote the joint probability, i.e., $\mathbb{P}(\{X = x\} \cap \{Y = y\})$ with $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. From Kolmogorov's definition, the conditional probabilities of X given Y is defined as $\mathbb{P}(X|Y) := \mathbb{P}(X,Y)/\mathbb{P}(Y)$. Given a real random variable $Z : \Omega \to \mathbb{R}$, let $\mathbb{E}(Z)$ denote the expected value, i.e., $\int_\Omega Z \, d\mathbb{P}$.

Given two independent theoretical experiments $E^{[1]}$ and $E^{[2]}$ that are statistically identical, let $(\Omega^{[1]}, \mathcal{F}^{[1]}, \mathbb{P}^{[1]})$ denote the probability space for the first experiment $E^{[1]}$. Similarly, define $(\Omega^{[2]}, \mathcal{F}^{[2]}, \mathbb{P}^{[2]})$ for the second experiment $E^{[2]}$. Since the two experiments are statistically identical, these two probability spaces are identical. Let $X^{[1]}$ and $X^{[2]}$ be random variables defined on $(\Omega^{[1]}, \mathcal{F}^{[1]}, \mathbb{P}^{[1]})$ and $(\Omega^{[2]}, \mathcal{F}^{[2]}, \mathbb{P}^{[2]})$, respectively. The two random variables are statistically identical, i.e., they are identical as a function from the σ-algebra to the appropriate range space. Let $(\Omega^{[3]}, \mathcal{F}^{[3]}, \mathbb{P}^{[3]})$ denote the product probability space of $(\Omega^{[1]}, \mathcal{F}^{[1]}, \mathbb{P}^{[1]})$ and $(\Omega^{[2]}, \mathcal{F}^{[2]}, \mathbb{P}^{[2]})$, see (Halmos, 1974). More formally, the sample space is $\Omega^{[3]} = \Omega^{[1]} \times \Omega^{[2]}$, the sigma algebra $\mathcal{F}^{[3]} \subset \Omega^{[3]}$ is the smallest σ-algebra such that $\mathcal{F}^{[3]} \supset \mathcal{F}^{[1]} \times \mathcal{F}^{[2]}$, and the probability measure $\mathbb{P}^{[3]}$ is such that, for any event $\mathcal{C}^{[3]} = \mathcal{C}^{[1]} \times \mathcal{C}^{[2]} \subset \mathcal{F}^{[3]}$ with $\mathcal{C}^{[1]} \subset \mathcal{F}^{[1]}$ and $\mathcal{C}^{[2]} \subset \mathcal{F}^{[2]}$, we have $\mathbb{P}^{[3]}(\mathcal{C}^{[3]}) = \mathbb{P}^{[1]}(\mathcal{C}^{[1]}) \times \mathbb{P}^{[2]}(\mathcal{C}^{[2]})$.

3.2.2 Information theory

Given a random variable X, let $H(X)$ denote the entropy, i.e.,

$$H(X) := -\mathbb{E}\left[\log\left(\mathbb{P}(X)\right)\right]$$

with $\log(\cdot)$ being the natural logarithm function. Given a finite collection of random variables $X^{[1]}, \ldots, X^{[n]}$, let $H(X^{[1]}, \ldots, X^{[n]})$ denote the joint entropy, i.e.,

$$H(X^{[1]}, \ldots, X^{[n]}) = -\mathbb{E}\left[\log\left(\mathbb{P}(X^{[1]}, \ldots, X^{[n]})\right)\right].$$

We have that

$$H(X^{[1]}, \ldots, X^{[n]}) \leq \sum_{i=1}^{n} H(X^{[i]})$$

with equality if and only if the $X^{[i]}$ are independent, i.e.,

$$\mathbb{P}(X^{[1]},\ldots,X^{[n]}) = \prod_{i=1}^{n} \mathbb{P}(X^{[i]}).$$

In addition, the conditional entropy

$$H(X^{[i]}|X^{[j]}) := -\mathbb{E}\left[\log\left(\mathbb{P}(X^{[i]}|X^{[j]})\right)\right]$$

is less that or equal to $H(X^{[i]})$, for any $i,j \in \{1,\ldots,n\}$, with equality if and only if $X^{[i]}$ and $X^{[j]}$ are independent.

The mutual information between two random variables X and Y is defined as

$$I(X;Y) := \int_{y\in\mathcal{Y}} \int_{x\in\mathcal{X}} \mathbb{P}(Y=y|X=x)\mathbb{P}(X=x) \log\left(\frac{\mathbb{P}(X=x|Y=y)}{\mathbb{P}(X=x)}\right) dx\, dy,$$

and is endowed with the following properties:

i) $I(X;Y) = I(Y;X) \geq 0$ with equality if and only if X and Y are independent;

ii) $I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) = H(X) + H(Y) - H(X,Y)$;

iii) $I(X^{[1]},\ldots,X^{[n]};Y) = \sum_{i=1}^{n} I(X^{[i]};Y|X^{[1]},\ldots,X^{[i-1]})$, where

$$I(X^{[i]};Y|X^{[1]},\ldots,X^{[i-1]}) = H(X^{[i]}|X^{[1]},\ldots,X^{[i-1]}) - H(X^{[i]}|Y,X^{[1]},\ldots,X^{[i-1]})$$

is the conditional mutual information of $X^{[i]}$ and Y given the finite collection of random variables $X^{[1]},\ldots,X^{[i-1]}$.

Note that the last property implies that

$$I(X^{[1]},\ldots,X^{[n]};Y) \leq \sum_{i=1}^{n} I(X^{[i]};Y)$$

with equality if and only if the $X^{[i]}$ are independent.

3.2.3 Complexity theory

Let $f, g : \mathbb{R} \to \mathbb{R}$ be two functions that map the set of real numbers to itself. The function f is said to be order at most g, denoted $f \in O(g)$, if there exists a constant $c \in \mathbb{R}$ such that $\lim_{n\to\infty} f(n)/g(n) \leq c$. In other words, asymptotically the function f grows much slower than the function g (Russell et al., 1995).

3.3 Summary

Although used throughout the thesis, the mathematical notation and preliminaries presented in this chapter are heavily relied on within the proof-based Sections 4.1.4, 4.6, 5.6, and 6.6. The material within these upcoming chapters constitutes novel research contributions, beginning in Chapter 4 with a rigorous proof that mutual information-based mapping controllers are eventually attracted to unexplored space.

Chapter 4

Mutual information-based control for mapping applications

For occupancy grid mapping, the most successful robot control algorithms are grounded in geometric-based principles. Frontier detection, or the identification of boundaries between unknown and unoccupied space (Yamauchi, 1998), drives many exploration algorithms, while path planning algorithms considering these frontiers often employ spatial metrics and heuristics (LaValle, 2006). This approach is not surprising given that the mapping algorithm's output is a representation of shapes, sizes, and positions of obstacles within the environment. However, underneath the mask of maximum likelihood, a map is fundamentally a field of binary random variables (Moravec and Elfes, 1985; Elfes, 1987), and a sensor is a probabilistic channel that links robot motion in the physical world to information gain for the Bayesian inference. One should expect that geometric-based intuition agrees with information-based reasoning concerning robot control, however, this agreement up till now is at most a conjecture.

Thus, our goal in this chapter is to rigorously characterize how the robot's motion in the physical world affects the relevance of its sensor measurements with respect to occupancy grid mapping. This goal goes beyond identifying which of the map's grid cells correspond to the largest amount of uncertainty, i.e., *entropy* (Elfes, 1995). We instead consider *mutual information* between the map and future sensor measurements to be the main reward for information-based control. Mutual information

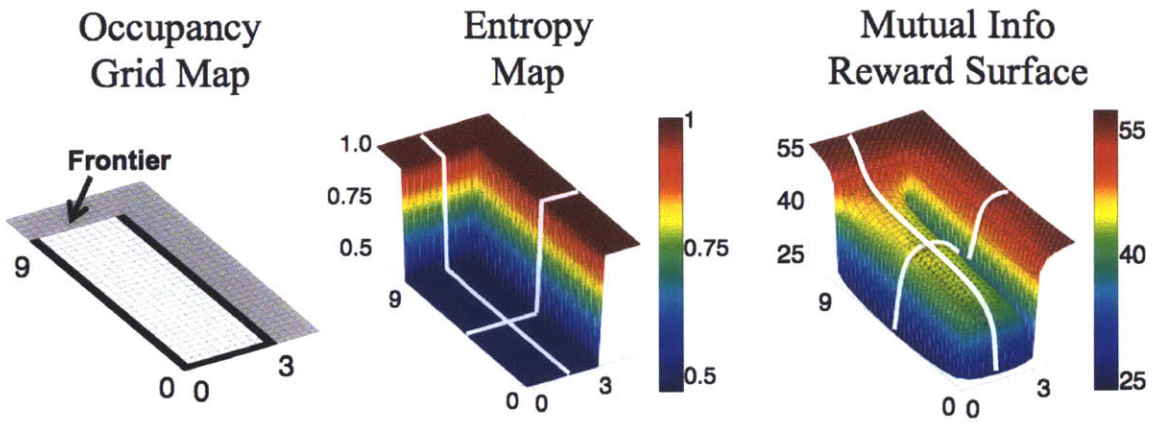


Figure 4-1: A 3 m by 9 m hallway "dead end" with three walls and one open frontier. *Left:* The occupancy grid map with occupied (black), unoccupied (white), and unknown (gray) cells. Surrounding cells not shown are assumed to be unknown. *Middle:* A drawing of the the resulting entropy map where the open space is indistinguishable from the walls and, more importantly, the open frontier is indistinguishable from boundaries between occupied and unknown cells. *Right:* A more geometrically relevant mutual information reward surface that guides a robot away from the three walls and towards the open frontier at the top left. Note that the white curves highlighting the entropy map's and mutual information reward surface's geometries are for illustrative purposes, and that the two linear color scales represent different bit ranges.

is an information-theoretic quantity (Cover and Thomas, 1991) that, in occupancy grid mapping applications, predicts how much future measurements will decrease the robot's uncertainty associated with all grid cells (Elfes, 1995). Since it is a function of both the robot's position and the uncertainty of the surrounding cells, mutual information encodes geometric relationships that are fundamental to robot control, thus yielding geometrically relevant reward surfaces on which the robot can navigate.

Consider the hallway mapping situation illustrated in Figure 4-1. With respect to the shown occupancy grid map, the mutual information reward surface within the unoccupied space rises to higher values towards the center of the hallway, reaching its largest value at the open frontier. A mutual information-based gradient-ascent controller acting on this surface would drive a robot away from the three walls and out past the open frontier. In other words, this controller would naturally exhibit safe navigation (i.e., stay away from obstacles) and frontier exploration (i.e., go towards

open frontiers) without geometrically interpreting the map (e.g., label space as either occupied or unoccupied). An entropy-based gradient-ascent controller acting on the adjacent entropy map would not exhibit such behavior and instead result in no robot motion within the unoccupied space. In other words, the entropy map lacks the geometrical relevance that is inherent to the mutual information reward surface.

To this effect, the main result of this chapter provides a geometric interpretation for the mutual information-based reasoning of a robot using the occupancy grid mapping algorithm and a narrow beam-based range sensor. More specifically, we prove that any controller tasked to maximize a mutual information reward function is eventually[1] attracted to unexplored space.[2] To prove this result, we develop a novel analytical approach that allows for the identification of mutual information-based behaviors. We consider simultaneously running two independent theoretical experiments defined under the same probability space. At a given time instance, we have the ability to induce slight differences in the conditioning of these experiments, e.g., the ability to condition both experiments with identical map posteriors except for a single cell of interest. We then employ expected value calculations (e.g., future map posteriors) to illuminate qualitative robot behaviors (e.g., attraction to unexplored space) that result from the quantitative behaviors of mutual information (e.g., increasing entropy overlap between random variables). Remarkably, the analytic approach and thus our theoretical findings are independent of the implemented control algorithm. Sampling-based, gradient-based, and potential field methods all exhibit attractive behavior as long as they aim to maximize a mutual information reward function.

To enable such control algorithms, we provide an algorithmic implementation for this reward function that is at worst quadratic and linear in time and space complexities, respectively, with respect to the map's spatial resolution. This computational tractability is achieved by exploiting a sensor obscuration assumption that is inher-

[1]By *eventually* we mean that this attraction does not need to be initially true, but instead it will be true at and after some time in the future.

[2]By *attracted to unexplored space* we mean that unknown cells within the map yield a higher mutual information reward compared to the hypothetical situation where these cells are correctly known.

ent in the occupancy grid mapping algorithm. Using this algorithm, we present experimental results that support our theoretical and computational findings. We consider two different hardware experiments, each using an omnidirectional ground robot equipped with a laser rangefinder. The first experiment involves mapping a laboratory equipped with a motion capture system, and thus we provide the occupancy grid mapping algorithm with accurate robot position information. This experiment shows the evolution of open frontiers as the robot expands the map, as well as suggests that obstacles are repulsive. The second experiment involves mapping a single floor of a large university building using a simultaneous localization and mapping (SLAM) algorithm (Leonard and Durrant-Whyte, 1991; Montemerlo et al., 2002). Even when calculating mutual information using the robot path of maximum a posteriori (Kretzschmar and Stachniss, 2012), the resulting reward surface highlights areas of the map where there exists open frontiers, obscured space, and potential SLAM issues (e.g., failed loop closures). These powerful experimental results support our theoretical findings while simultaneously validating the algorithm's time and space complexities.

4.1 Eventual attraction to unexplored space

In this section we review the occupancy grid mapping algorithm then formally state the main theoretical result of this chapter. A detailed proof of this theorem is given in Section 4.2.3, and experimental results that support our findings are provided in Section 4.4.3.

4.1.1 The occupancy grid mapping algorithm

Consider the problem of constructing a map of an environment using a single robot equipped with a sensor that provides range measurements, i.e., the robot's distance to nearby obstacles. Suppose the robot employs the well-known occupancy grid mapping algorithm (Moravec and Elfes, 1985; Elfes, 1987) for constructing the map from the noisy observations. For simplicity, we focus on the two-dimensional mapping problem, meaning that a robot of known position $c_k \in \mathbb{R}^2 \times \mathbb{S}$ at the discrete time

step[3] $k \in \mathbb{Z}_{\geq 0}$ moves in the two-dimensional Euclidean plane. The occupancy grid mapping algorithm models the static map of the robot's environment as an n_m-tuple random variable $M = (M^{[1]}, \ldots, M^{[n_m]})$, where each independent binary random variable $M^{[i]}$ takes the value 0 or 1 when an obstacle is absent or present, respectively, within a uniquely identified cell $i \in \{1, \ldots, n_m\}$. Inference of this map is enabled by observations the robot receives originating from its sensor. These observations are modeled as random variables Z_k that take values z_k, for all times k.

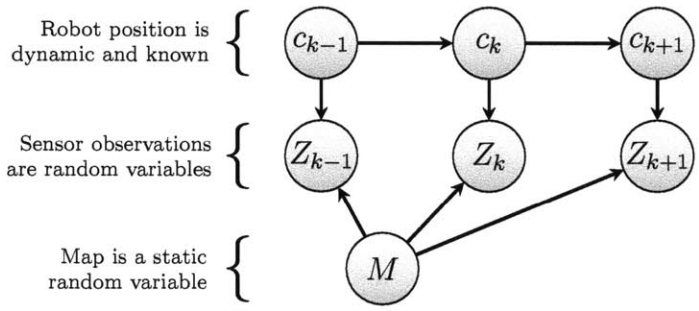


Figure 4-2: Dynamic Bayesian network for occupancy grid mapping. Note that since the robot's position is known, we show in the top row the realized values c_k instead of corresponding random variables.

As illustrated in Figure 4-2, both the unknown map and the robot's known positions influence the noisy observations within this dynamic Bayesian network. Our goal of building a map is equivalent to estimating the posterior probability over all possible maps given the previous observations, which leads to the binary Bayesian filter

$$r^{[i]}(z_{1:k}) := \frac{\mathbb{P}(M^{[i]} = 1 | z_{1:k})}{\mathbb{P}(M^{[i]} = 0 | z_{1:k})} = \delta^{[i]}(z_k) r^{[i]}(z_{1:k-1}), \tag{4.1}$$

where $r^{[i]}$ is the odds ratio[4] of the posterior of cell i being occupied, $\delta^{[i]}$ is the odds ratio of the inverse sensor model, and $z_{1:k}$ is shorthand for $\bigcap_{k'=1}^{k} \{Z_{k'} = z_{k'}\}$. Note that for all cells $i \in \{1, \ldots, n_m\}$, we assume an uninformative map prior $\mathbb{P}(M^{[i]} = 0) = \mathbb{P}(M^{[i]} = 1) = 0.5$. We also use $r_k^{[i]}$ as shorthand for $r^{[i]}(z_{1:k})$ and refer to it

[3] In this chapter, we only consider discrete time steps and refer to these as simply time k.

[4] The reader may be more familiar with the log odds representation of the binary Bayesian filter (4.1). Although we use log odds ratios in the implementation for our experiments, the mathematical conclusions in this chapter are more compactly expressed using odds ratios.

as the robot's *belief* for cell i at time k. When considering this belief as a random variable, we use the notation $R_k^{[i]} := r^{[i]}(Z_{1:k})$, where $Z_{1:k}$ is shorthand for $\bigcap_{k'=1}^{k}\{Z_{k'}\}$.

We consider in this chapter a beam-based range sensor that produces finite radial measurements of known bearing to nearby obstacles, even though our analysis is applicable to a wide range of sensor models. In addition, we assume that the beam is narrow, i.e., the width of the sensor's beam is negligible when compared to the size of our occupancy grid cells. This is the case for many laser-based sensors, e.g., laser rangefinders. Multiple measurements per time step are accommodated by modeling the observation Z_k as an n_z-tuple random variable $(Z_k^{[1]}, \ldots, Z_k^{[n_z]})$ whose elements take values $z_k^{[j]}$ from a common set of range measurements $\mathcal{Z} \subset \mathbb{R}_{\geq 0}$. The occupancy grid mapping algorithm assumes independence among the measurements $j \in \{1, \ldots, n_z\}$, and computes the odds ratio of the inverse sensor model by $\delta^{[i]}(z_k) = \prod_{j=1}^{n_z} \delta^{[i,j]}(z_k^{[j]})$, where

$$\delta^{[i,j]}(z_k^{[j]}) = \begin{cases} r_{occ}, & \text{if } z_k^{[j]} \text{ and } c_k \text{ imply a measurement} \\ & \text{location within cell } i, \\ r_{emp}, & \text{if } z_k^{[j]} \text{ and } c_k \text{ imply a sensor beam} \\ & \text{passed through cell } i, \\ 1, & \text{otherwise,} \end{cases}$$

with $r_{occ} \in (1, \infty)$ and $r_{emp} \in (0, 1)$ being constants. Note that for all functions, we do not include the robot's position c_k as an argument since it is known at all times k.

It is useful to consider the ordered set of cell indices $\mathcal{I}_k^{[j]}$ located along the jth measurement's beam in the absence of obstacles. In other words, the jth measurement's *perceptual range* contains all cells that align with this beam and are within the maximum range of the sensor. As illustrated for the example system in Figure 4-3, by 'ordered' we mean that set operators will consider the cell indices sequentially based on proximity to the robot, with the nearest cells being considered first. Also note that for all cells ℓ not in the observation's perceptual range $\mathcal{I}_k := \bigcup_{j=1}^{n_z} \mathcal{I}_k^{[j]}$, any realized observation z_k results in an odds ratio of the inverse sensor model of $\delta^{[\ell]}(z_k) = 1$. This

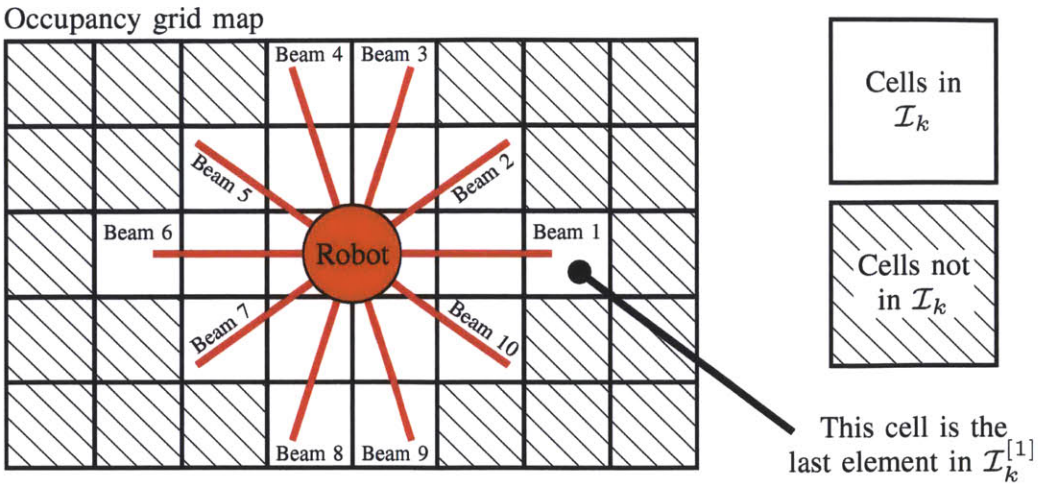


Figure 4-3: The perceptual ranges for an example robot with a ten beam range sensor. Note that the assumption of measurement independence within a single observation may overlook the property that one cell is traversed by multiple measurements' beams.

implies that the belief $r_k^{[\ell]}$ of this cell will not change over time if the robot remains stationary. On the other hand, for a cell i in the observation's perceptual range $\mathcal{I}_k$, any realized observation z_k with $\delta^{[i]}(z_k) \neq 1$ will change the corresponding belief, i.e., $r_k^{[i]} \neq r_{k-1}^{[i]}$. We call such an observation that influences the belief an *informative observation* for cell i. Clearly an informative observation $z_k = (z_k^{[1]}, \ldots, z_k^{[n_z]})$ must be composed of at least one *informative measurement* $z_k^{[j]}$ for cell i. For simplicity, we sometimes say the jth measurement is informative for cell i when $\delta^{[i,j]}(z_k^{[j]}) \neq 1$.

4.1.2 Main technical assumptions

Throughout the chapter, we assume that the sensor model is unbiased and the measurement's prior is clamped. Given enough number of informative observations for each of the cells, an unbiased sensor model assumes that the robot is able to correctly infer the realized states of all cells in the occupancy grid map. For all practical purposes, an unbiased sensor model is the desired result of proper sensor calibration since it implies the following intuitive property. Given that an observation Z_k is informative for a cell i, the expected value of the inverse sensor model $\delta^{[i]}(Z_k^{[i]})$ is less than 1 for an unoccupied cell (i.e., $M^{[i]} = 0$) and greater than 1 for an occupied one (i.e., $M^{[i]} = 1$). Otherwise, incorrect inferences in expectation would be possible.

The assumption of a clamped measurement's prior $\mathbb{P}(Z_{k+1}^{[j]}|z_{1:k})$ is a consequence of the calibration process that generates the sensor model, which commonly involves measuring a reference marker (e.g., a white kent sheet for laser rangefinders) within assumed free space. In other words, the calibration process considers the marker to be an ideal obstacle and does not account for the non-zero probability of other unexpected obstacles (e.g., airborne particulates, electromagnetic interference). Thus, we make the technical assumption that the resulting measurement's prior $\mathbb{P}(Z_{k+1}^{[j]}|z_{1:k})$ is clamp with respect to the belief $r_k^{[i]}$ of a cell i within that measurement's perceptual range $\mathcal{I}_{k+1}^{[j]}$. For example, a cell i strongly believed to be unoccupied (i.e., $r_k^{[i]} \ll 1$) results in a measurement's prior equal to $\mathbb{P}(Z_{k+1}^{[j]}|z_{1:k}, M^{[i]} = 0)$. We now formally define these two assumptions.

Assumption 1 (Unbiased sensor model) *For all times $k \in \mathbb{Z}_{\geq 0}$, given that the next observation Z_{k+1} is informative $\delta^{[i]}(Z_{k+1}) \neq 1$ for a cell $i \in \mathcal{I}_{k+1}$, the expected value of $\delta^{[i]}(Z_{k+1})$ is less than 1 if the cell is unoccupied and greater than 1 if it is occupied.*

Assumption 2 (Clamped measurement's prior) *For all times $k \in \mathbb{Z}_{\geq 0}$, cells $i \in \mathcal{I}_{k+1}^{[j]}$, and measurements $j \in \{1, \ldots, n_z\}$, there exists an $\epsilon_0 > 0$ such that if $r_k^{[i]} < \epsilon_0$, then*

$$\mathbb{P}(Z_{k+1}^{[j]}|z_{1:k}) = \mathbb{P}(Z_{k+1}^{[j]}|z_{1:k}, M^{[i]} = 0).$$

Similarly, there exists an $\epsilon_1 < \infty$ such that if $r_k^{[i]} > \epsilon_1$, then

$$\mathbb{P}(Z_{k+1}^{[j]}|M) = \mathbb{P}(Z_{k+1}^{[j]}|z_{1:k}, M^{[i]} = 1).$$

Both of these assumptions are arguably satisfied by any range sensing robot running the occupancy grid mapping algorithm in a reasonable manner. For example, consider a stationary robot using a biased sensor to continuously measure an obstacle. If the magnitude of the bias is much larger than the size of the occupancy grid cells, one can easily show that the mapping algorithm will almost surely yield an

incorrect map. This example also highlights why the proper sizing of the occupancy grid cells with respect to sensor performance is important. Conversely, given that the sensor is unbiased, almost surely there will be a time after which the uncertainty associated with the map will be lower than with the calibration environment (see Corollary 11 in Section 4.2.2). Our assumption about a clamped measurement's prior insures that the resulting probabilities remain conservative when the robot becomes overly confident about the map, i.e., more confident that during the calibration process. Interestingly, this assumption is usually satisfied in a real system not because the system designer clamps the measurement's prior, but instead he/she clamps the robot's beliefs to avoid numerical instabilities associated with the Bayesian inference and mutual information calculations (Yguel et al., 2007). Note that the latter is a more strict assumption that we do not enforce.

4.1.3 Identification of mutual information-based behaviors

Given the previous observations $z_{1:k}$, we are interested in minimizing the expected uncertainty (i.e., conditional entropy[5]) of the map M after receiving the robot's next observation Z_{k+1}, i.e., we wish to minimize

$$H(M|Z_{k+1}, z_{1:k}) = H(M|z_{1:k}) - I(M; Z_{k+1}|z_{1:k}).$$

Since the entropy $H(M|z_{1:k})$ of the map is independent of the robot's future position c_{k+1}, the control objective of minimizing the conditional entropy $H(M|Z_{k+1}, z_{1:k})$ is equivalent to maximizing the mutual information $I(M; Z_{k+1}|z_{1:k})$ between the map and the next observation. In addition, since both M and Z_{k+1} are tuple random variables with independence among elements, we express total mutual information as the summation of $I(M^{[i]}; Z_{k+1}^{[j]}|z_{1:k})$ between $M^{[i]}$ and $Z_{k+1}^{[j]}$ over all cells $i \in \{1, \ldots, n_m\}$

[5]The *conditional* entropy given the next observation should not be confused with the current entropy of the map M. The latter is independent of the next robot position c_{k+1} and the next observation Z_{k+1}.

and measurements $j \in \{1, \ldots, n_z\}$. We further simplify this expression to

$$I(M; Z_{k+1}|z_{1:k}) = \sum_{j=1}^{n_z} \sum_{i \in \mathcal{I}_{k+1}^{[j]}} I(M^{[i]}; Z_{k+1}^{[j]}|z_{1:k}) \tag{4.2}$$

since the mutual information between a measurement and any cell not in the measurement's perceptual range is equal to 0, i.e., $I(M^{[i]}; Z_{k+1}^{[j]}|z_{1:k}) = 0$ for all $j \in \{1, \ldots, n_z\}$ with $i \notin \mathcal{I}_{k+1}^{[j]}$ (see Lemma 6 for a rigorous treatment). We refer to $I(M^{[i]}; Z_{k+1}^{[j]}|z_{1:k})$ as the *mutual information contribution* of cell i, but note that it is also specific to the particular measurement j.

We refer to expression given in (4.2) as the *mutual information reward function*, a function that drives many information-based control approaches that are cast in terms of an optimization problem. In this chapter, we exploit certain algebraic properties of this reward function to show that maximizing (4.2) implies eventual attraction to unknown space. Intuitively speaking, we show that unknown cells (i.e., those cells i for which the belief $r_k^{[i]}$ is equal to 1) at and after some time are expected to yield a higher mutual information reward compared to the hypothetical situation where these cells are correctly known (i.e., the belief $r_k^{[i]}$ is less than or greater than 1 depending on if cell i was realized to be unoccupied or occupied, respectively).

Remark 3 (Cell-wise effect on mutual information-based behavior) *Ideally we would define attraction to be the behavior that, given a map where one cell i is unknown (i.e., $r_k^{[i]} = 1$) and another cell ℓ is known (i.e., $r_k^{[\ell]} \neq 1$), the robot moves in the direction of cell i. This definition is unfortunately not useful since many other factors influence robot motion, including but not limited to the sensor model, the position of these two cells, the realization of the rest of the map, and the implemented control algorithm. In other words, the effect the cells i and ℓ have on mutual information-based behavior cannot be reasoned about using this definition without additional assumptions.*

To formalize the simultaneous existence of an actual situation and a hypothetical one, we construct a novel analytical approach that we call *two independent theoretical*

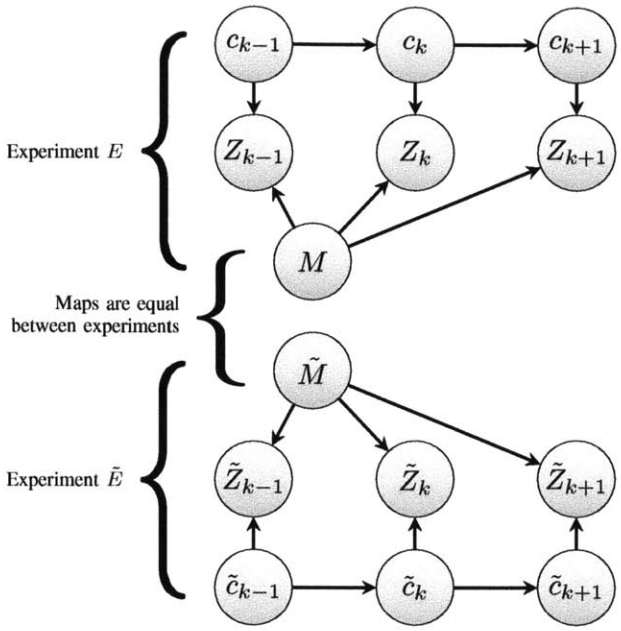


Figure 4-4: Dynamic Bayesian network for the two independent theoretical experiments, E and $\tilde{E}$, for which all random variables are generated by a single probability space. Note that the positions c_k and $\tilde{c}_k$ are equal, for all times $k \in \mathbb{Z}_{\geq 0}$.

experiments. This approach allows us to identify the effect of changing the belief of one or more cells while keeping all other aspects (e.g., robot's positions, realized map) the same between the two situations. These situations are, in fact, modeled as two independent theoretical experiments running simultaneously under the same probability space and conditioned appropriately. Both experiments start at time 0, run until time $k \geq 0$, and have identical robots with identical sensors moving over identical paths, all of which are known. It is also known that these experiments are carried out on maps conditioned to be equal, however, the value of the maps themselves are not given to the two robots. As a result, each robot can only infer the state of its map based on conditions given at some fixed time in the past. The result of these experiments is to conclude expected characteristics of the mutual information reward function given subtle yet relevant differences in the conditioning of both robots' beliefs while avoiding the issues discussed in Remark 3. Informally speaking, the two independent theoretical experiments approach is a proof technique that allows us to consider these differences in beliefs with "all else being equal." Defined rigorously

below, we believe that this approach is important on its own right for proving the validity of widely-used heuristics in information-based control.

Definition 4 (Two independent theoretical experiments) *Let E and $\tilde{E}$ denote two independent theoretical experiments with identical robots, paths, sensor models, and inverse sensors models, i.e., all deterministic variables and functions are the same. Suppose the first experiment's map M is of equal value to the second experiment's map $\tilde{M}$, although the value itself is unknown to the robots. In addition, suppose the observations $Z_{1:k} := \bigcap_{k'=1}^{k}\{Z_{k'}\}$ and $\tilde{Z}_{1:k} := \bigcap_{k'=1}^{k}\{\tilde{Z}_{k'}\}$ are independently realized in E and $\tilde{E}$, respectively. At some time $K \leq k$ in the past, the robots have known beliefs $R_K^{[i]}$ and $\tilde{R}_K^{[i]} := r^{[i]}(\tilde{Z}_{1:K})$ for all cells, which can be used to condition their expected beliefs $\mathbb{E}(R_k^{[i]})$ and $\mathbb{E}(\tilde{R}_k^{[i]})$, respectively, at time k.*

The intuition behind the experiments E and $\tilde{E}$ comes from thinking of time k as being in the future versus thinking of time K as being in the past. Doing so immediately raises the following question: *Given information about the map at time K, can we reason about the expected behavior of mutual information-based control at time k?* Even for arbitrary maps, the expected beliefs $\mathbb{E}\big(R_k^{[i]}\big)$ and $\mathbb{E}\big(\tilde{R}_k^{[i]}\big)$ give insight into how past observations can affect future uncertainty - note that future beliefs are indeed random variables. It is this insight that provides a basis for qualitatively evaluating the behaviors of controllers employing the mutual information reward function.

4.1.4 A theorem for attractive behavior

The main theoretical result of this chapter implies that any controller seeking to maximize the mutual information reward function is eventually attracted to unexplored regions of the map. Below we formally state this result with the use of the analytical approach described in the previous section.

For the first experiment E, suppose that a subset of cells $\mathbb{I} \subset \{1, \ldots, n_m\}$ is composed of *unknown* cells at time K, meaning that the beliefs $R_K^{[i]}$ are all equal to 1 for all $i \in \mathbb{I}$. In the second experiment $\tilde{E}$, however, suppose that the same subset of cells are *correctly known* at time K. By correctly known we mean that for any

cell $i \in \mathbb{I}$ at time K, the belief $\tilde{R}_K^{[i]}$ is less than 1 if the cell is unoccupied, otherwise, it is greater than 1. We again note that the maps M and $\tilde{M}$ are of equal value yet unknown to the robots. Lastly, suppose that for all other cells not in the subset $\mathbb{I}$, the beliefs between experiments E and $\tilde{E}$ are equal but otherwise arbitrary in value. More formally, we have that $R_K^{[\ell]} = \tilde{R}_K^{[\ell]}$ for all cells $\ell \in \mathbb{I}^c = \{1, \ldots, n_m\}/\mathbb{I}$.

Let the event $\mathcal{A}_K$ represent the intersection of all aforementioned events between the two experiments at time K, i.e.,

$$\mathcal{A}_K := \{M = \tilde{M}\} \cap \bigcap_{\ell \in \mathbb{I}^c} \{R_K^{[\ell]} = \tilde{R}_K^{[\ell]}\} \cap \bigcap_{i \in \mathbb{I}} \{R_K^{[i]} = 1\}$$
$$\cap \bigcap_{i \in \mathbb{I}} \left\{ \left\{ \{\tilde{M}^{[i]} = 0\} \cap \{\tilde{R}_K^{[i]} < 1\} \right\} \cup \left\{ \{\tilde{M}^{[i]} = 1\} \cap \{\tilde{R}_K^{[i]} > 1\} \right\} \right\}.$$

When reasoning about the mutual information contributions $I(M^{[i]}; Z_{k+1}^{[j]} | z_{1:k})$ for all measurements $j \in \{1, \ldots n_z\}$ and cells $i \in \{1, \ldots n_m\}$ at time $k \geq K$, we wish to evaluate these values based on the expected beliefs $\mathbb{E}(R_k^{[i]})$ and $\mathbb{E}(\tilde{R}_k^{[i]})$ given the event $\mathcal{A}_K$. In other words, we are evaluating the resulting mutual information reward function $I(M; Z_{k+1} | z_{1:k})$ employing the expected beliefs for the robots in both experiments E and $\tilde{E}$. This event is formally described by

$$\mathcal{B}_{K,k} := \bigcap_{i=1}^{n_m} \left\{ \left\{ R_k^{[i]} = \mathbb{E}(R_k^{[i]} | \mathcal{A}_K) \right\} \cap \left\{ \tilde{R}_k^{[i]} = \mathbb{E}(\tilde{R}_k^{[i]} | \mathcal{A}_K) \right\} \right\}.$$

Note that this event does not explicitly include the event $\mathcal{A}_K$ since this information (e.g., the state of the map) is unknown to the robots. We now state the main theoretical result of the chapter.

Theorem 5 (Attraction to unexplored space) *Consider the two independent experiments E and $\tilde{E}$. Given enough number of informative observations with $\delta^{[i]}(z_k) \neq 1$ for the subset of cells $i \in \mathbb{I}$, there exists a time $\varkappa \geq K$ such that for all times $k \geq \varkappa$, we have that*

$$I(M; Z_{k+1} | \mathcal{B}_{K,k}) \; > \; I(\tilde{M}; \tilde{Z}_{k+1} | \mathcal{B}_{K,k}) \tag{4.3}$$

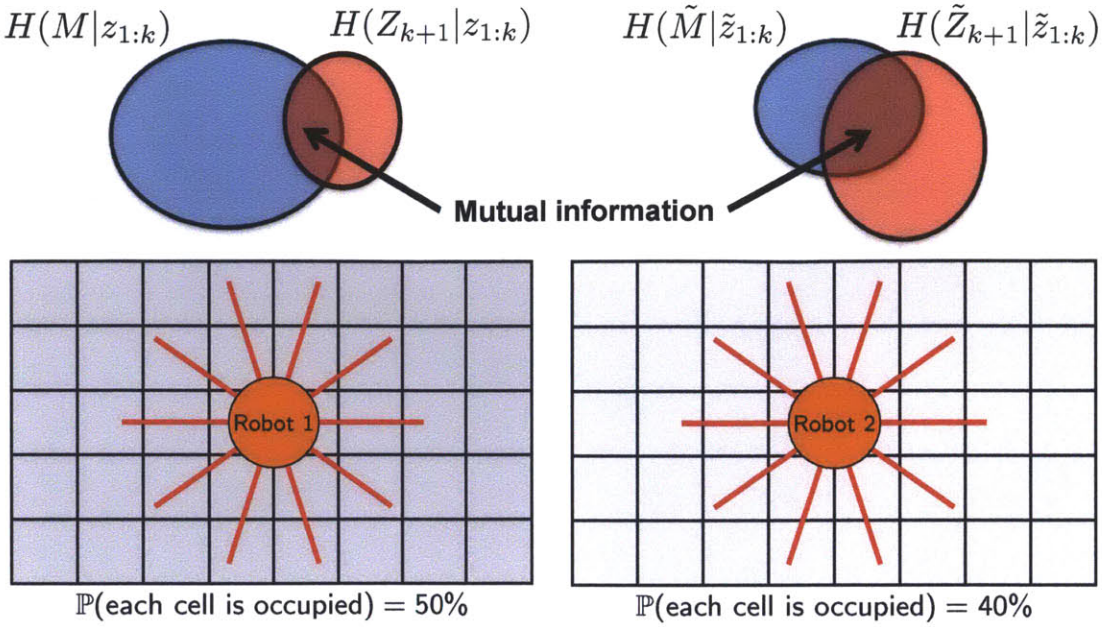


Figure 4-5: An illustration of when, given identical robots, a larger amount of entropy at a given time does not imply a greater mutual information reward, i.e., a greater overlap of the shown entropies. Consider the case when the path of the two robots are identical and their maps are composed of the shown cells realized to be unoccupied. Theorem 5 states that, in expectation, there is a future time at and after which mutual information will be greater for robot 1 than robot 2.

if any cell $i \in \mathbb{I}$ *is in the observation's perceptual range* $\mathcal{I}_{k+1}$.

Note that although an unknown cell always has the maximum amount of entropy at time K, more certain beliefs (i.e., $r_k^{[i]} \neq 1$) of that cell may result in larger mutual information contributions at time K, e.g., Figure 4-5. Conditioning on the expected beliefs for all cells, Theorem 5 states that, in expectation, there is a future time at and after which the mutual information reward function is greater if these cells are unknown at time K than if they were correctly known. This statement is stronger than one stating that the cell's mutual information contribution will be greater for a particular measurement. This statement also holds for any observation that perceives at least one of these cells, regardless of the sensor model. Thus, we conclude that mutual information-based controllers are eventually attracted to unexplored space.

4.2 On informative behaviors

In this section, we prove this chapter's main theoretical result stated in Theorem 5. Along the way, we provide a number of other results that may be interesting in their own right. The proofs for the later results are given in Section 4.6 so that we can focus primarily on the proof for the main result.

4.2.1 Algebraic properties of mutual information

Concerning an arbitrary cell $i \in \{1, \ldots, n_m\}$ and measurement $j \in \{1, \ldots, n_z\}$, the following lemma expresses the corresponding mutual information contribution as an integral of two functions, one of which has useful algebraic properties.

Lemma 6 (Mutual information contribution) *The mutual information at time k between a single measurement j and a single cell i given the observations $z_{1:k}$ can be written as*

$$I(M^{[i]}; Z_{k+1}^{[j]} | z_{1:k}) = \int_{z \in \mathcal{Z}} \mathbb{P}(Z_{k+1}^{[j]} = z | z_{1:k}) \; f\left(\delta^{[i,j]}(z), r_k^{[i]}\right) \; dz \tag{4.4}$$

where the nonnegative function $f(\delta, r)$ is endowed with the following two algebraic properties:

1. *For $\delta = 1$ we have that $f(\delta, r) = 0$ for any $r \in (0, \infty)$;*
2. *For any value of $\delta \in (0, \infty)/\{1\}$, there exists some $\tau(\delta) \in (0, \infty)$ such that $f(\delta, r)$ is monotonically increasing in r for all $r < \tau(\delta)$ and monotonically decreasing in r for all $r > \tau(\delta)$.*

Remark 7 (Belief thresholds) *For an informative measurement, we have that δ is equal to either r_{occ} or r_{emp}, implying that the belief threshold $\tau(\delta)$ is either less than 1 or greater than 1, respectively. Suppose cell i is unknown, i.e. $r_k^{[i]} = 1$. Then measurement locations within cell i result in $\tau(r_{occ}) < 1$, implying that $f(\delta, r_k^{[i]})$ is monotonically decreasing in the belief $r_k^{[i]}$. For measurements of greater value, $\tau(r_{emp}) > 1$ and $f(\delta, r_k^{[i]})$ is instead monotonically increasing.*

One can interpret $f(\delta, r)$ in Lemma 6 as an information gain function for potential measurement values $z \in \mathcal{Z}$. For example, the second algebraic property states that no information about a cell is received from a measurement that does not intersect that cell, whether because the measurement "stops short" of the cell or simply because the cell is not in the measurement's perceptual range. This geometric dependency relates back to the construction of the index sets $\mathcal{I}_k$ in Section 4.1.1, and likewise allows for the mutual information reward function to be calculated over a reduced domain size. In other words, no matter how the jth measurement's prior $\mathbb{P}(Z_{k+1}^{[j]}|z_{1:k})$ behaves, the integral of (4.4) does not need to evaluate measurement values $z \in \mathcal{Z}$ that imply $\delta^{[i,j]}(z) = 1$ for any cell i in the measurement's perceptual range $\mathcal{I}_{k+1}^{[j]}$.

For many robot exploration tasks, one can use mutual information to create reward surfaces for realtime motion planning algorithms, e.g., grid-based and sample-based searches. On the other hand, the gradient of mutual information with respect to the robot's position is needed for potential field planners and greedy reactive controllers. Previous works often approximate this gradient by calculating the difference in expected information gain associated with moving the robot to neighboring cells, see (Grocholsky, 2002; Rocha et al., 2005; Hoffmann and Tomlin, 2010). Here we show that the *analytical gradient* of mutual information can be calculated for this general class of range sensors, although we will discuss in Section 5.2.2 a variant that is more appropriate for gradient-based control.

Corollary 8 (Gradient of mutual information contribution) *Let the sensor model $\mathbb{P}(Z_k|M)$ be differentiable almost everywhere with respect to the robot's position. Then the gradient of mutual information at time k between a single measurement j and a single cell i given the observations $z_{1:k}$ exists and can be written as*

$$\frac{\partial I(M^{[i]}; Z_{k+1}^{[j]}|z_{1:k})}{\partial c_{k+1}} = \int_{z \in \mathcal{Z}} \frac{\partial \mathbb{P}(Z_{k+1}^{[j]} = z|z_{1:k})}{\partial c_{k+1}} \, f\left(\delta^{[i,j]}(z), r_k^{[i]}\right) \, dz. \tag{4.5}$$

Remark 9 (Similarity of forms) *It is interesting that the mutual information contribution and its gradient have the same form, except that in the latter the measure-*

ment's prior is replaced by its gradient with respect to the robot's position. Schwager et al. (2011a) observed a similar property for general mutual information expressions being summed (or integrated) over all possible states. We will show in Section 4.3.2 that the expectation over all possible map states is inherently incorporated into the total mutual information expression and its gradient without a summation operation, thus further simplifying these expressions.

4.2.2 Expected values of future beliefs

For the main theoretical result in Theorem 5, we are interested in the geometric interpretations of the monotonic properties of the information gain function $f(\delta, r)$, i.e., the second algebraic property of Lemma 6. For example, consider an unoccupied cell i with $M^{[i]} = 0$ that is unknown (i.e., $r_k^{[i]} = 1$) to a stationary robot. Suppose this robot receives a sequence of measurements all extending past the aforementioned cell. Lemma 6 says that the information gain function $f(\delta, r_k^{[i]})$ itself provably *decreases* with each observation, implying that the potential to receive informative observations was highest when the cell was originally unknown. This behavior suggests that unknown cells have stronger attractive properties for measurements implying the absence of an obstacle.

We would like to have the same result for measurement locations falling within a cell. This would imply that for all measurements large enough to "reach" the unoccupied cell, the information gain function provably *decreases* as the robot becomes more certain of the cell's unoccupied state. From Remark 7, this is not the case for $r_k^{[i]} = 1$. Fortunately, $\tau(r_{occ})$ in Lemma 6 gives the necessary belief threshold below which the monotonicity result holds for any informative measurement. Given the expectation to infer the correct state for observable cells, we have that all unoccupied cells would eventually exhibit the desired behavior, i.e., the expected information gain decreases with increased certainty of the state of a cell. The same interpretation holds true for occupied cells, which motivates the following lemma and corollaries that explore the robot's beliefs and measurements' priors in expectation.

Lemma 10 (Relationship between expected beliefs) *Suppose a numerical relationship (i.e., less than, greater than, or equal to) is given at time $K \geq 0$ between the beliefs $R_K^{[i]}$ and $\tilde{R}_K^{[i]}$, for any cell i. Then for all times $k \geq K$, the numerical relationship between the expected beliefs $\mathbb{E}(R_k^{[i]}|\mathcal{A}_K^{[i]})$ and $\mathbb{E}(\tilde{R}_k^{[i]}|\mathcal{A}_K^{[i]})$ will remain the same.*

Corollary 11 (Convergence of expected beliefs) *The expected beliefs $\mathbb{E}(R_k^{[i]}|\mathcal{A}_K^{[i]})$ and $\mathbb{E}(\tilde{R}_k^{[i]}|\mathcal{A}_K^{[i]})$ of any unoccupied (occupied) cell i monotonically converge to 0 (∞, respectively) as the number of informative observations for this cell tends to ∞.*

Corollary 12 (Equality of measurements' priors) *Consider the set $\mathbb{I}$ from Section 4.1.4 as a set consisting of a single cell i. Given enough number of informative observations for this cell, there exists a time $\varkappa \geq K$ such that for all times $k \geq \varkappa$, we have that the jth measurements' priors $\mathbb{P}(Z_{k+1}^{[j]}|\mathcal{B}_{K,k}^{[i]})$ and $\mathbb{P}(\tilde{Z}_{k+1}^{[j]}|\mathcal{B}_{K,k}^{[i]})$ are equal for all measurements j.*

Consider the example given in Figure 4-5, and suppose the shown state of the two experiments are that of time K. Lemma 10 implies that for all time $k \geq K$, the expected beliefs of robot 1 will be greater than those of robot 2 for all shown cells. Corollary 11 implies that the beliefs of both robots will converge to 0 as the number of informative observations for the corresponding cell tends to ∞. Lastly, given enough number of these informative observations, Corollary 12 implies that there exists a time $\varkappa \geq K$ such that for all times $k \geq \varkappa$, any measurement priors between the experiments is equal.

4.2.3 A proof for attractive behavior

We now give the proof for the main theoretical result in Theorem 5.

Proof 13 (Theorem 5) *Suppose the subset $\mathbb{I} \subset \{1, \ldots, n_m\}$ for the two experiments E and $\tilde{E}$ consists of a single cell i of arbitrary state. Denote this subset as $\mathbb{I}_1$. We first show that* (4.3) *holds for this subset regardless of the state of cell i.*

Suppose this cell is realized to be unoccupied, i.e., $M^{[i]} = 0$. *From Corollary 11, the two expected beliefs* $\mathbb{E}(R_k^{[i]}|\mathcal{A}_K)$ *and* $\mathbb{E}(\tilde{R}_k^{[i]}|\mathcal{A}_K)$ *monotonically converge to* 0 *as the number of informative observations for cell* i *tends to* ∞. *Thus, there exists a time* $\varkappa_1^{[1]} \geq K$ *such that for all* $k \geq \varkappa_1^{[1]}$, *both these expected beliefs are smaller than the belief thresholds* $\tau(r_{occ})$ *and* $\tau(r_{emp})$. *In addition, the second algebraic property of Lemma 6 guarantees that for all times* $k \geq \varkappa_1^{[1]}$, *the information gain function* $f(\delta, r)$ *employing* $\mathbb{E}(R_k^{[i]}|\mathcal{A}_K)$ *and* $\mathbb{E}(\tilde{R}_k^{[i]}|\mathcal{A}_K)$ *is monotonically decreasing in* r *for both* $\delta = r_{occ}$ *and* $\delta = r_{emp}$. *Since at time* K *cell* i *was unknown (i.e.,* $R_K^{[i]} = 1$*) in the first experiment* E *and correctly known (i.e.,* $\tilde{R}_K^{[i]} < 1$*) in the second experiment* $\tilde{E}$, *we have from Lemma 10 that* $\mathbb{E}(R_k^{[i]}|\mathcal{A}_K)$ *is larger than* $\mathbb{E}(\tilde{R}_k^{[i]}|\mathcal{A}_K)$, *hence the corresponding* $f(\delta, r)$ *is larger. Also due to Lemma 10, we have for any time* k *that the expected beliefs* $\mathbb{E}(R_k^{[\ell]}|\mathcal{A}_K)$ *and* $\mathbb{E}(\tilde{R}_k^{[\ell]}|\mathcal{A}_K)$ *of all other cells* $\ell \in \mathbb{I}_1^c$ *are equal. This implies that the information gain function* $f(\delta, r)$ *for these expected beliefs are also equal between experiments.*

Now consider the expected behavior of the measurements' priors $\mathbb{P}(Z_{k+1}^{[j]}|\mathcal{B}_{K,k})$ *and* $\mathbb{P}(\tilde{Z}_{k+1}^{[j]}|\mathcal{B}_{K,k})$ *for the two experiments* E *and* $\tilde{E}$, *respectively. From Corollary 12, there exists another time* $\varkappa_1^{[2]} \geq K$ *such that for all times* $k \geq \varkappa_1^{[2]}$, *these measurements' priors are equal. From Lemma 6, we have for all time* $k \geq \max\{\varkappa_1^{[1]}, \varkappa_1^{[2]}\}$ *that* $I(M^{[i]}; Z_{k+1}^{[j]}|\mathcal{B}_{K,k})$ *is greater than* $I(\tilde{M}^{[i]}; \tilde{Z}_{k+1}^{[j]}|\mathcal{B}_{K,k})$ *for any measurement* j *with* $i \in \mathcal{I}_{k+1}^{[j]}$. *For all other measurements or cells in* $\mathbb{I}_1^c$, *these mutual information contributions are equal. This proves* (4.3) *holds for the subset* $\mathbb{I}_1$ *consisting of a single unoccupied cell* i *with a threshold time of* $\max\{\varkappa_1^{[1]}, \varkappa_1^{[2]}\}$.

The proof that (4.3) *holds for the subset* $\mathbb{I}_1$ *consisting of a single occupied cell* i *follows verbatim; however, this proof results in two subtle difference. Firstly, there exists a time* $\varkappa_1^{[3]} \geq K$ *such that for all times* $k \geq \varkappa_1^{[3]}$, *the two expected beliefs* $\mathbb{E}(R_k^{[i]}|\mathcal{A}_K)$ *and* $\mathbb{E}(\tilde{R}_k^{[i]}|\mathcal{A}_K)$ *are greater than both belief thresholds* $\tau(r_{occ})$ *and* $\tau(r_{emp})$. *Secondly from Corollary 12, there exists a time* $\varkappa_1^{[4]} \geq K$ *such that for all times* $k \geq \varkappa_1^{[4]}$, *the measurements' priors* $\mathbb{P}(Z_{k+1}^{[j]}|\mathcal{B}_{K,k})$ *and* $\mathbb{P}(\tilde{Z}_{k+1}^{[j]}|\mathcal{B}_{K,k})$ *are equal. Thus,* (4.3) *holds for the subset* $\mathbb{I}_1$ *consisting of a single cell* i *of arbitrary state with a threshold time of* $\varkappa_1 := \max\{\varkappa_1^{[1]}, \varkappa_1^{[2]}, \varkappa_1^{[3]}, \varkappa_1^{[4]}\}$.

We conclude by showing that (4.3) *holds for subsets* $\mathbb{I}_s$ *of arbitrary size* s. *The proof proceeds by induction on* s, *with* $s = 1$ *already proved. Let us assume that the result is true for some* $s \geq 1$. *For* $s+1$ *with the subset* $\mathbb{I}_{s+1}$, *define a one cell subset* $\mathbb{J} := \mathbb{I}_{s+1}/\mathbb{I}_s$ *containing the one cell not originally in* $\mathbb{I}_s$, *then let* $\mathbb{J}^c := \{1, \ldots, n_m\}/\mathbb{J}$. *The proof to show that the mutual information reward function* $I\big(M; Z_{k+1}|\mathcal{B}_{K,k}\big)$ *in the first experiment* E *is strictly greater that* $I\big(\tilde{M}; \tilde{Z}_{k+1}|\mathcal{B}_{K,k}\big)$ *in the second experiment* $\tilde{E}$ *follows verbatim as the proof for* $\mathbb{I}_1$, *with the exception that the analysis yields a threshold time of* $\varkappa_s$. *Thus, Theorem 5 holds in general for an arbitrary subset* $\mathbb{I}_s$ *with a threshold time of* $\varkappa = \max \bigcup_s \{\varkappa_s\}$.

4.3 Computing mutual information for mapping

In this section we give the complete derivation of a numerical expression used to compute the mutual information reward function for a range sensing robot. We also discuss the resulting computational complexity and highlight the benefits of employing a narrow beam-based sensor model predisposed to sensor obscuration.

4.3.1 Computing the measurement's prior

The inverse sensor model $\delta^{[i,j]}$ for the occupancy grid mapping algorithm inherently involves the assumption of sensor obscuration. Roughly speaking, the closest occupied cell within a given measurement's perceptual range obscures all cells behind it along the beam, i.e., the sensor model is conditionally independent of the state of all cells located behind an obstacle. This assumption leads to significant computational benefits when numerically evaluating the mutual information reward function. However, we first rigorously state the assumption of sensor obscuration, using the fact that the index set $\mathcal{I}_k^{[j]}$ representing the jth measurement's perceptual range is ordered for any measurement j at any time $k \in \mathbb{Z}_{>0}$.

Assumption 14 (Sensor obscuration) *For any cell* ℓ *appearing after an occupied*

cell i in the index set $\mathcal{I}_{k+1}^{[j]}$, we have that

$$\mathbb{P}\big(Z_{k+1}^{[j]} | M^{[i]} = 1, M^{[\ell]} = 0\big) = \mathbb{P}\big(Z_{k+1}^{[j]} | M^{[i]} = 1, M^{[\ell]} = 1\big) \tag{4.6}$$

for any time k and measurement j.

Remark 15 (Finite range sensors) *Suppose a cell ℓ is not within the jth measurement's perceptual range, i.e., $\ell \notin \mathcal{I}_{k+1}^{[j]}$. For finite range sensors, (4.6) provides a rigorous statement for probabilistic independence with respect to any cell i in the map.*

We now express the measurement's prior $\mathbb{P}(Z_{k+1}^{[j]} | z_{1:k})$ and its gradient $\partial \mathbb{P}(Z_{k+1}^{[j]} | z_{1:k}) / \partial c_{k+1}$ compactly as functions of the robot's sensor model and beliefs.

Lemma 16 (Measurement's prior) *Suppose the assumption of sensor obscuration holds (Assumption 14). We have for any time k and measurement j that*

$$\mathbb{P}(Z_{k+1}^{[j]} | z_{1:k}) = \frac{1}{\prod\limits_{\ell \in \mathcal{I}_{k+1}^{[j]}} (r_k^{[\ell]} + 1)} \mathbb{P}(Z_{k+1}^{[j]} | M = \mathbf{0}) + \sum_{i \in \mathcal{I}_{k+1}^{[j]}} \frac{r_k^{[i]}}{\prod\limits_{\ell \leq i} (r_k^{[\ell]} + 1)} \mathbb{P}(Z_{k+1}^{[j]} | M = \mathbf{e}_i), \tag{4.7}$$

where $\ell \leq i$ denotes the sequence of indices in $\mathcal{I}_{k+1}^{[j]}$ appearing before or equal to i.

Corollary 17 (Measurement's prior gradient) *Suppose the assumption of sensor obscuration holds (Assumption 14). We have for any time k and measurement j that*

$$\frac{\partial \mathbb{P}(Z_{k+1}^{[j]} | z_{1:k})}{\partial c_{k+1}} = \sum_{i \in \mathcal{I}_{k+1}^{[j]}} \frac{r_k^{[i]}}{\prod\limits_{\ell \leq i} (r_k^{[\ell]} + 1)} \frac{\partial \mathbb{P}(Z_{k+1}^{[j]} | M = \mathbf{e}_i)}{\partial c_{k+1}}. \tag{4.8}$$

Remark 18 (Sensor obscuration) *We immediately see the effects of sensor obscuration on a cell inferred to be occupied, i.e., $r_k^{[i]}$ is greater than 1. As the robot becomes more certain that this cell is occupied, (4.7) and (4.8) imply that all cells in*

$\mathcal{I}_{k+1}^{[j]}$ *located farther from the robot have a decreasing influence on the measurement's prior and its gradient.*

By inherently accounting for sensor obscuration, the measurement's prior compensate for the otherwise strong attractive properties of unknown cells located behind obstacles. Without this compensation, the robot may believe that additional information about these unknown cells can be obtained by simply moving towards the obstacle. This hypothetical and undesirable behavior best highlights the dangers associated with assuming conditional independence of the sensor model among cells, i.e., $\mathbb{P}(Z_{k+1}^{[j]}|M) = \mathbb{P}(Z_{k+1}^{[j]}|M^{[i]})$ for a given $i \in \mathcal{I}_{k+1}^{[j]}$.

4.3.2 Computing the mutual information reward function

We now finalize the expressions for analytically evaluating the mutual information reward function $I(M; Z_{k+1}|z_{1:k})$ and its gradient $\partial I(M; Z_{k+1}|z_{1:k})/\partial c_{k+1}$.

Lemma 19 (Analytically evaluating reward) *For any time k, we have that*

$$\begin{aligned} I(M; Z_{k+1}|z_{1:k}) = &\sum_{j=1}^{n_z} \sum_{i \in \mathcal{I}_{k+1}^{[j]}} \int_{z \in \mathcal{Z}} f\big(\delta^{[i,j]}(z), r_k^{[i]}\big) \\ &\times \Bigg(\sum_{\ell \in \mathcal{I}_{k+1}^{[j]}} \frac{r_k^{[\ell]}}{\prod_{\zeta \leq \ell} (r_k^{[\zeta]} + 1)} \mathbb{P}(Z_{k+1}^{[j]} = z | M = \boldsymbol{e}_\ell) \\ &+ \frac{1}{\prod_{\zeta \in \mathcal{I}_{k+1}^{[j]}} (r_k^{[\zeta]} + 1)} \mathbb{P}(Z_{k+1}^{[j]} = z | M = \boldsymbol{0}) \Bigg) dz. \end{aligned} \tag{4.9}$$

Corollary 20 (Analytically evaluating the gradient of reward) *For any time k, we have that*

$$\begin{aligned} \frac{\partial I(M; Z_{k+1}|z_{1:k})}{\partial c_{k+1}} = &\sum_{j=1}^{n_z} \sum_{i \in \mathcal{I}_{k+1}^{[j]}} \int_{z \in \mathcal{Z}} f\big(\delta^{[i,j]}(z), r_k^{[i]}\big) \\ &\times \sum_{\ell \in \mathcal{I}_{k+1}^{[j]}} \frac{r_k^{[\ell]}}{\prod_{\zeta \leq \ell} (r_k^{[\zeta]} + 1)} \frac{\partial \mathbb{P}(Z_{k+1}^{[j]} = z | M = \boldsymbol{e}_\ell)}{\partial c_{k+1}} dz. \end{aligned} \tag{4.10}$$

From an implementation standpoint, the mutual information reward function (4.9) and its gradient (4.10) are numerically calculated in a discretized fashion. More specifically, the integral over all possible range measurements is computed using a discretized approximation. Let $\widehat{\mathcal{Z}}$ be the discretized set of range measurements representing the set $\mathcal{Z}$. In addition, let λ_z be the numerical integration's resolution for this discretized set. The expression to approximate $I(M; Z_{k+1}|z_{1:k})$ and $\partial I(M; Z_{k+1}|z_{1:k})/\partial c_{k+1}$, denoted $\widehat{I}(M; Z_{k+1}|z_{1:k})$ and $\partial \widehat{I}(M; Z_{k+1}|z_{1:k})/\partial c_{k+1}$, respectively, are given in the following corollaries.

Corollary 21 (Numerically evaluating reward) *For any time k, we have that*

$$\begin{aligned}\widehat{I}(M; Z_{k+1}|z_{1:k}) = &\sum_{j=1}^{n_z} \sum_{i \in \mathcal{I}_{k+1}^{[j]}} \sum_{z \in \widehat{\mathcal{Z}}} f\big(\delta^{[i,j]}(z), r_k^{[i]}\big) \\ &\times \Bigg(\sum_{\ell \in \mathcal{I}_{k+1}^{[j]}} \frac{r_k^{[\ell]}}{\prod_{\zeta \leq \ell} (r_k^{[\zeta]} + 1)} \mathbb{P}(Z_{k+1}^{[j]} = z | M = \boldsymbol{e}_\ell) \\ &+ \frac{1}{\prod_{\zeta \in \mathcal{I}_{k+1}^{[j]}} (r_k^{[\zeta]} + 1)} \mathbb{P}(Z_{k+1}^{[j]} = z | M = \boldsymbol{0}) \Bigg) \lambda_z^{-1}. \end{aligned} \tag{4.11}$$

Corollary 22 (Numerically evaluating the gradient of reward) *For any time k, we have that*

$$\begin{aligned}\frac{\partial \widehat{I}(M; Z_{k+1}|z_{1:k})}{\partial c_{k+1}} = &\sum_{j=1}^{n_z} \sum_{i \in \mathcal{I}_{k+1}^{[j]}} \sum_{z \in \widehat{\mathcal{Z}}} f\big(\delta^{[i,j]}(z), r_k^{[i]}\big) \\ &\times \sum_{\ell \in \mathcal{I}_{k+1}^{[j]}} \frac{r_k^{[\ell]}}{\prod_{\zeta \leq \ell} (r_k^{[\zeta]} + 1)} \frac{\partial \mathbb{P}(Z_{k+1}^{[j]} = z | M = \boldsymbol{e}_\ell)}{\partial c_{k+1}} \lambda_z^{-1}. \end{aligned} \tag{4.12}$$

4.3.3 Computational complexity

Consider the implementations of (4.11) and (4.12) given in Algorithm 1 and Algorithm 2, respectively. To determine their computational complexities, let us assume that the occupancy grid map composed of square cells has a spatial resolution of λ_m, i.e., λ_m^{-2} is the area of any cell. We now state the time and space complexity for Algo-

Algorithm 1 MappingMutualInformationReward()

Require: Beliefs $r_k^{[i]}$ for all cells $i \in \mathcal{I}_{k+1}$;
1: // initialize reward function
   $\widehat{I} \leftarrow 0$;
2: **for** $j = 1, \ldots, n_z$ **do**
3: // employ favorite ray casting algorithm
    $\mathcal{I}_{k+1}^{[j]} \leftarrow \text{RayCast}(j)$;
4: // precompute denominators found in (4.8)
    $\eta \leftarrow \text{BeliefCumulationProduct}(\mathcal{I}_{k+1}^{[j]})$;
5: **for all** $i \in \mathcal{I}_{k+1}^{[j]}$ **do**
6: $P_{occ}, P_{emp} \leftarrow 0$;
7: **for all** $z \in \widehat{\mathcal{Z}}$ **do**
8: **if** $\delta^{[i,j]}(z) = r_{occ}$ **then**
9: // measurement suggests occupied cell
    $P_{occ} \leftarrow P_{occ} + \text{MeasurementPrior}(j, z, \eta)$;
10: **else if** $\delta^{[i,j]}(z) = r_{emp}$ **then**
11: // measurement suggests unoccupied cell
    $P_{emp} \leftarrow P_{emp} + \text{MeasurementPrior}(j, z, \eta)$;
12: **end if**
13: **end for**
14: // update mutual information accumulation
    $\widehat{I} \leftarrow \widehat{I} + f(r_{occ}, r_k^{[i]})P_{occ} + f(r_{emp}, r_k^{[i]})P_{emp}$;
15: **end for**
16: **end for**
17: **return** $\widehat{I}\lambda_z^{-1}$;

rithm 1 and Algorithm 2 in terms of Big-O notation $O(\cdot)$ using the map's resolution λ_m, the numerical integration's resolution λ_z, and the number of range measurements n_z per observation.

Proposition 23 (Computational complexity) *For both Algorithm 1 and Algorithm 2, the time and space complexities are $O(\lambda_m^2 \lambda_z n_z)$ and $O(\lambda_m)$, respectively. In words, the time (space) complexity of calculating mutual information reward and its gradient is at worst quadratic (linear, respectively) in the map's spatial resolution and linear (independent, respectively) in both the number of range measurements per observation and the numerical integration's resolution.*

We have found that the serial implementation of Algorithm 1 and Algorithm 2 is appropriate for most post processing tasks concerning mutual information reward (see

Algorithm 2 MappingMutualInformationRewardGradient()

Require: Beliefs $r_k^{[i]}$ for all cells $i \in \mathcal{I}_{k+1}$;
1: // initialize reward function's gradient
$d\widehat{I} \leftarrow \mathbf{0}$;
2: **for** $j = 1, \ldots, n_z$ **do**
3: // employ favorite ray casting algorithm
$\mathcal{I}_{k+1}^{[j]} \leftarrow \text{RayCast}(j)$;
4: // precompute denominators found in (4.8)
$\eta \leftarrow \text{BeliefCumulationProduct}(\mathcal{I}_{k+1}^{[j]})$;
5: **for all** $i \in \mathcal{I}_{k+1}^{[j]}$ **do**
6: $dP_{occ}, dP_{emp} \leftarrow \mathbf{0}$;
7: **for all** $z \in \widehat{\mathcal{Z}}$ **do**
8: **if** $\delta^{[i,j]}(z) = r_{occ}$ **then**
9: // measurement suggests occupied cell
$dP_{occ} \leftarrow dP_{occ} + \text{MeasurementPriorGradient}(j, z, \eta)$;
10: **else if** $\delta^{[i,j]}(z) = r_{emp}$ **then**
11: // measurement suggests unoccupied cell
$dP_{emp} \leftarrow dP_{emp} + \text{MeasurementPriorGradient}(j, z, \eta)$;
12: **end if**
13: **end for**
14: // update mutual information accumulation
$d\widehat{I} \leftarrow d\widehat{I} + f(r_{occ}, r_k^{[i]})dP_{occ} + f(r_{emp}, r_k^{[i]})dP_{emp}$;
15: **end for**
16: **end for**
17: **return** $d\widehat{I}\,\lambda_z^{-1}$;

Sections 4.4). Further speedups leading to realtime computation can be obtained by the parallelization of these algorithms at line 2, i.e., the for loop with respect to the independent measurements. With respect to computing mutual information reward surfaces, additional parallelizations can be used among the desired robot locations. These implementation details are relevant due to the computational tractability implied by Proposition 23, which is the direct consequence of assuming sensor obscuration (Assumption 14). The following proposition highlights this property.

Proposition 24 (Intractability when not assuming sensor obscuration) *Suppose Assumption 14 does not hold. The necessary modifications to Algorithm 1 and Algorithm 2 result in a time complexity of* $O(2^{\lambda_m}\lambda_m\lambda_z n_z)$. *In other words, the time complexity of calculating mutual information reward and its gradient without the as-*

Algorithm 3 BeliefCumulationProduct($\mathcal{I}_{k+1}^{[j]}$)

for all $i \in \mathcal{I}_{k+1}^{[j]}$ **do**
  if $i = \text{First}(\mathcal{I}_{k+1}^{[j]})$ **then**
    // first entry of tuple
    $\eta \leftarrow r_k^{[i]} + 1$;
  else
    // tuple augmentation
    $\eta \leftarrow \big(\eta, \text{Last}(\eta)(r_k^{[i]} + 1)\big)$;
  end if
end for
return η

Algorithm 4 MeasurementPrior($j, \mathcal{I}_{k+1}^{[j]}, z, \eta$)

// first line of (4.7)
$P \leftarrow \mathbb{P}(Z_{k+1}^{[j]} = z | M = \mathbf{0}) / \text{Last}(\eta)$;
for all $i \in \mathcal{I}_{k+1}^{[j]}$ **do**
  // second line of (4.7)
  $P \leftarrow P + r_k^{[i]} \mathbb{P}(Z_{k+1}^{[j]} = z | M = \mathbf{e}_i) / \eta^{[i]}$;
end for
return P

sumption of sensor obscuration is exponential in the map's spatial resolution and thus computationally intractable.

4.4 Experiments in mapping

Using Algorithm 1 presented in Section 4.3.3, we present experimental results that support our theoretical and computational findings. We consider two different hardware experiments, each using an omnidirectional ground robot equipped with a laser rangefinder. The first experiment employs a traditional occupancy grid mapping algorithm to show the progression of open frontiers as the robot expands the map, as well as suggests that obstacles are repulsive. The second experiment employs a simultaneous localization and mapping (SLAM) algorithm to show that the resulting mutual information reward surface highlights areas of the occupancy grid map where there exists open frontiers, obscured space, and potential SLAM issues (e.g.,

Algorithm 5 MeasurementPriorGradient$(j, \mathcal{I}_{k+1}^{[j]}, z, \eta)$

1: // initialize summation
$dP \leftarrow 0$;
2: **for all** $i \in \mathcal{I}_{k+1}^{[j]}$ **do**
3: // second line of (4.8)
$dP \leftarrow dP + r_k^{[i]} \frac{\partial \mathbb{P}(Z_{k+1}^{[j]} = z | M = \mathbf{e}_i)}{\partial c_{k+1}} / \eta^{[i]}$;
4: **end for**
5: **return** dP

failed loop closures). We first, however, discuss the technical aspects of our hardware mapping system.

Parameter	Symbol	Value
Hit weighting	z_{hit}	0.7
Range set	$\mathcal{Z}$	$(0.0, 5.6)$
Short weighting	z_{short}	0.1
Max weighting	z_{max}	0.1
Random weighting	z_{rand}	0.1
Occupy constant	r_{occ}	1.50
Unoccupy constant	r_{emp}	0.66
Hit standard deviation	σ_{hit}	$\max\{0.03 \text{ m}, 0.03 z_{exp}\}$
Short decay	λ_{short}	5.0 m^{-1}

Table 4.1: Parameters for the beam-based proximity mixture sensor model.

4.4.1 Hardware mapping system

Using the Robot Operating System (Quigley et al., 2009), we employed an omnidirectional ground robot custom built at MIT Lincoln Laboratory (Figure 4-6). We equipped this robot with a Hokuyo laser rangefinder module that produces $n_z = 1024$ horizontal measurements over 360 degrees at a rate of 10 Hz. For both hardware experiments, we selected measurement range values of $\mathcal{Z} = (0.0, 5.6)$ m, mapping algorithm parameters of $(r_{occ}, r_{emp}) = (1.50, 0.66)$, and a beam-based proximity mixture sensor model (Thrun et al., 2005) with parameters found in Table 4.1. Figures 4-7 and 4-8 show the corresponding measurement distribution $\mathbb{P}(Z_{k+1}^{[j]} | M = \mathbf{e}_i)$ and gradient $\partial \mathbb{P}(Z_{k+1}^{[j]} | M = \mathbf{e}_i) / \partial c_{k+1}$, respectively, for various expected measurement values.

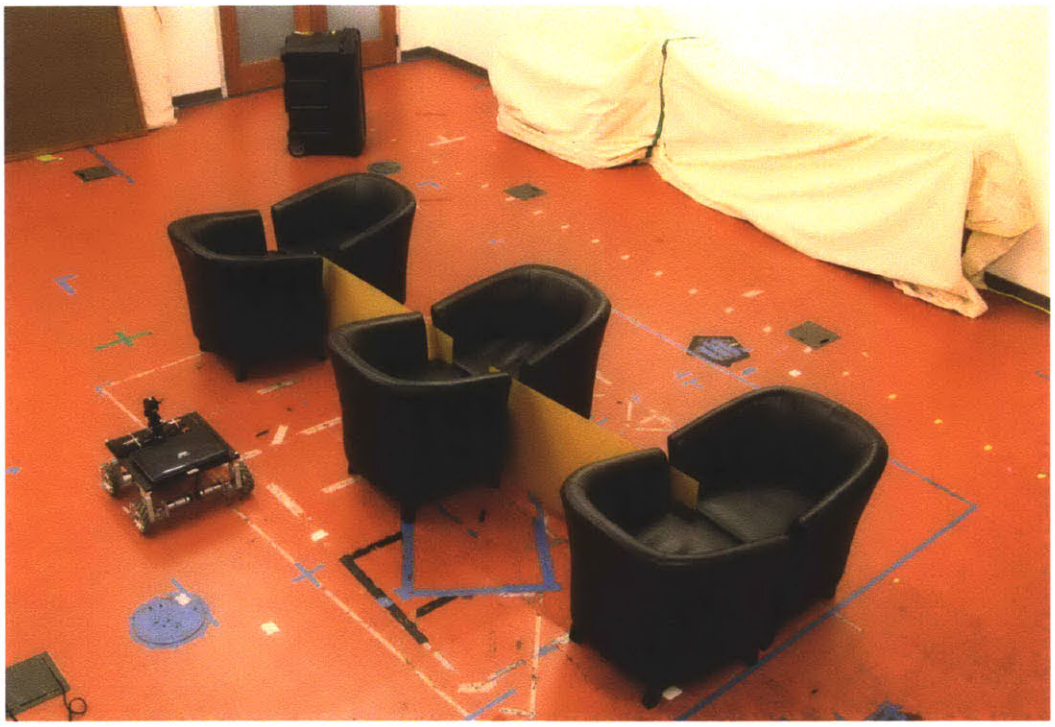


Figure 4-6: An omnidirectional ground robot with 360 degree laser rangefinder is driven clockwise around a center opaque barrier for the occupancy grid mapping experiment. The hardware experiment in Section 4.4.3 employing a SLAM algorithm uses the same robot system within a different environment.

4.4.2 Mutual information for occupancy grid mapping

We now discuss the resulting mutual information reward surfaces describing the overlap between map uncertainty and information gain from the robot's laser rangefinder. To simplify these surfaces, we prohibited the omnidirectional robot from rotating, i.e., for a particular translational position, the robot is always assumed to be facing in one direction. This constraint is not overly restrictive due to the rotational symmetry of the sensor, yet it allows us to ignore the, in general, nontrivial effect rotational motion has on mutual information.

Figure 4-6 shows the setup for the occupancy grid mapping experiment. The environment to be mapped was set up with a center barrier and several scattered obstacles. A motion capture system tracked the position of the robot in realtime, negating the need to probabilistically infer the location of the robot during the mapping process. In order to build a mostly complete map using a simplistic path, we manually

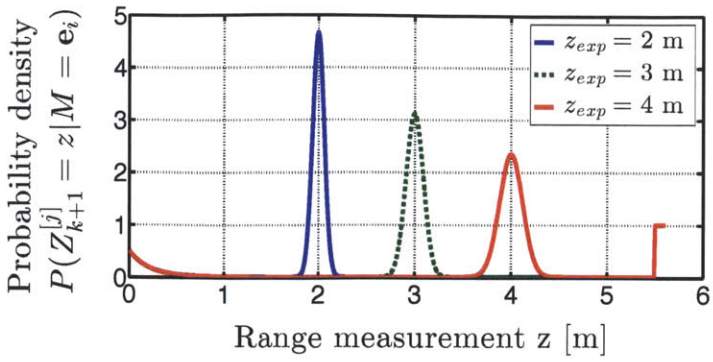


Figure 4-7: The measurement distribution for the sensor model for expected measurement values of 2, 3, and 4 m.

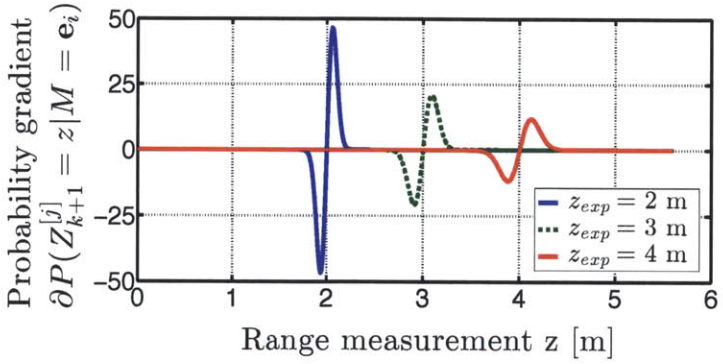


Figure 4-8: The measurement distribution gradient with respect to the robot's position for the sensor model for expected measurement values of 2, 3, and 4 m.

drove the robot around the center barrier in a single loop. The data recorded from this 60 second run was then used to generate a time series of occupancy grid maps with spatial resolution $\lambda_m = 5$ m^{-1}. Figure 4-9 shows five time instances for which we calculated the resulting mutual information reward surfaces. Due to the strong propagating frontiers (red regions) leading the robot, the evolution of these surfaces suggest that mutual information-based controllers would have drawn the robot down a similar path for exploration. In addition, we noted that the center barrier and exterior walls exhibited repulsive tendencies (i.e., low mutual information reward) due to the likelihood of sensor obscuration. These conclusions are not only consisted with the theoretical findings of this chapter, but are consisted among 10 additional runs of varying time durations that we conducted.

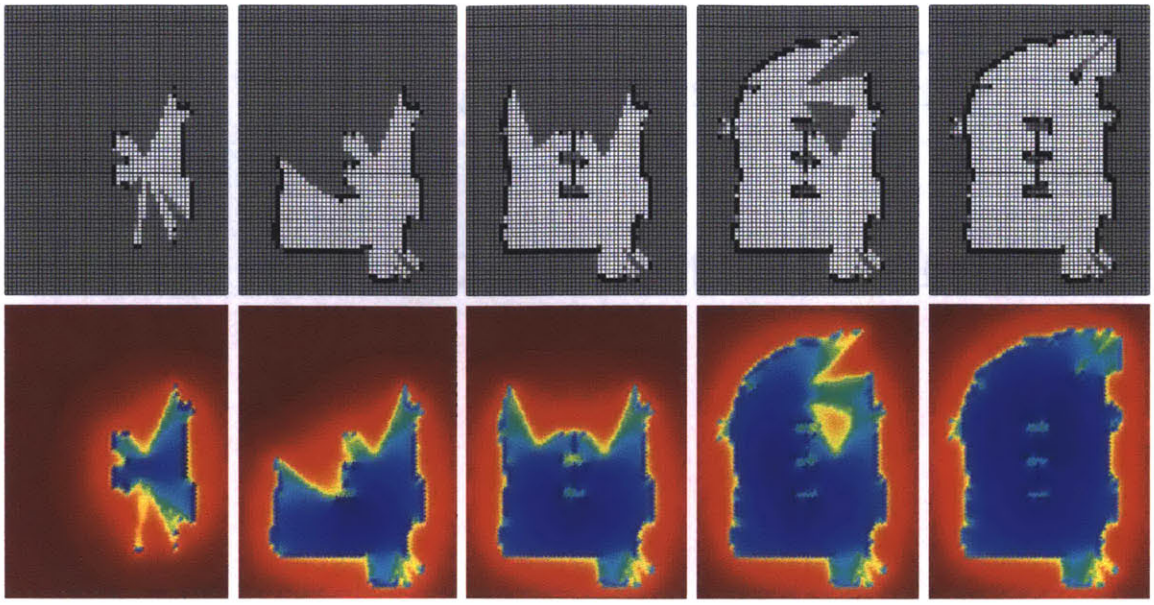


Figure 4-9: The evolution of the occupancy grid map and mutual information reward surface for occupancy grid mapping. *Left to right:* The map (top) and corresponding surface (bottom) at times of 0, 15, 30, 45 and 60 s. As open frontiers move within the occupancy grid map, the mutual information reward evolves to reflect new locations for informative exploration. At the end of the experiment, the mutual information reward surface highlights smaller frontiers and areas obscured by obstacles. These surfaces also suggest that known obstacles (e.g., exterior walls) are repulsive due to their low mutual information rewards. Note that the linear color scale from blue to red represents [4, 68] bit for the mutual information reward surfaces.

All calculations were performed in MATLAB R2012a running on a 3.4 Ghz quad-core desktop equipped with a GTX 690 nVidia graphics card. Each mutual information reward surface was generated using a numerical integration resolution of $\lambda_z = 50$ m^{-1}. Note that these surfaces evaluate mutual information reward with respect to robot positions located at the centroid of all occupancy grid map cells. Considering that most mutual information-based controllers consider finite horizons, we can reduce the set of robot positions over which these reward surfaces are calculated.

4.4.3 Mutual information for SLAM

For the second experiment, we employed the GMapping simultaneous localization and mapping (SLAM) package (Grisetti et al., 2007). We again manually drove the robot, however, the environment to be mapped was much larger (1500+ m^2). Over

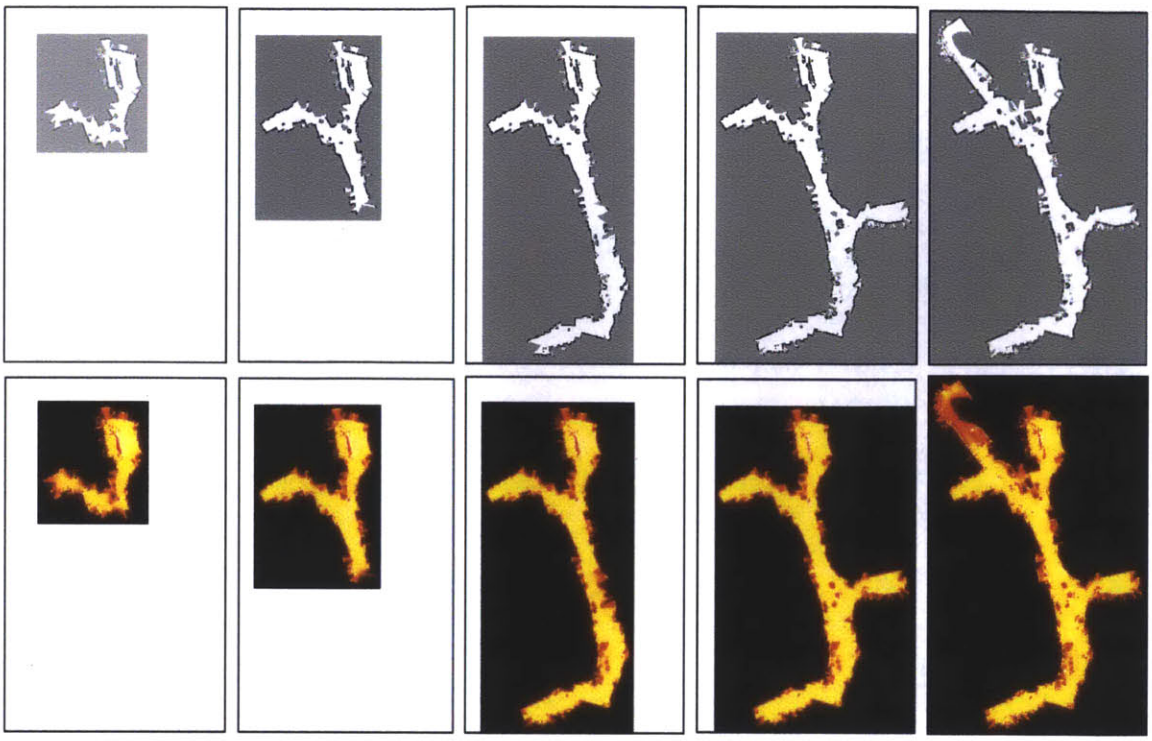


Figure 4-10: The evolution of the occupancy grid map and mutual information reward surface for SLAM. *Left to right:* The map (top) and corresponding surface (bottom) at times of 5, 10, 15, 25 and 50 min. Note that the linear color scale from white to black represents [9, 48] bit for the mutual information reward surfaces.

50 minutes of data was collected, after which SLAM was performed to build an occupancy grid map with spatial resolution $\lambda_m = 3.3$ m^{-1}. Figure 4-11 compares the resulting entropy map against the mutual information reward surface generated with a numerical integration's resolution of $\lambda_z = 33$ m^{-1}. In the entropy map, we see many cells of high uncertainty, however, they tend to be in isolated clusters near objects. On the other hand, the mutual information reward surface gives strong visual indicators throughout the open areas of the map. In particular, locations of interest to a mapping robot (e.g., frontiers, obscured space, and potential SLAM issues) are easily identifiable by either human or algorithm. Thus, we conclude that mutual information can both enable and monitor the SLAM algorithm. Lastly, we note that even for this large map, the mutual information reward surface over all known cells can be calculated from scratch in less than ten seconds on our desktop computer.

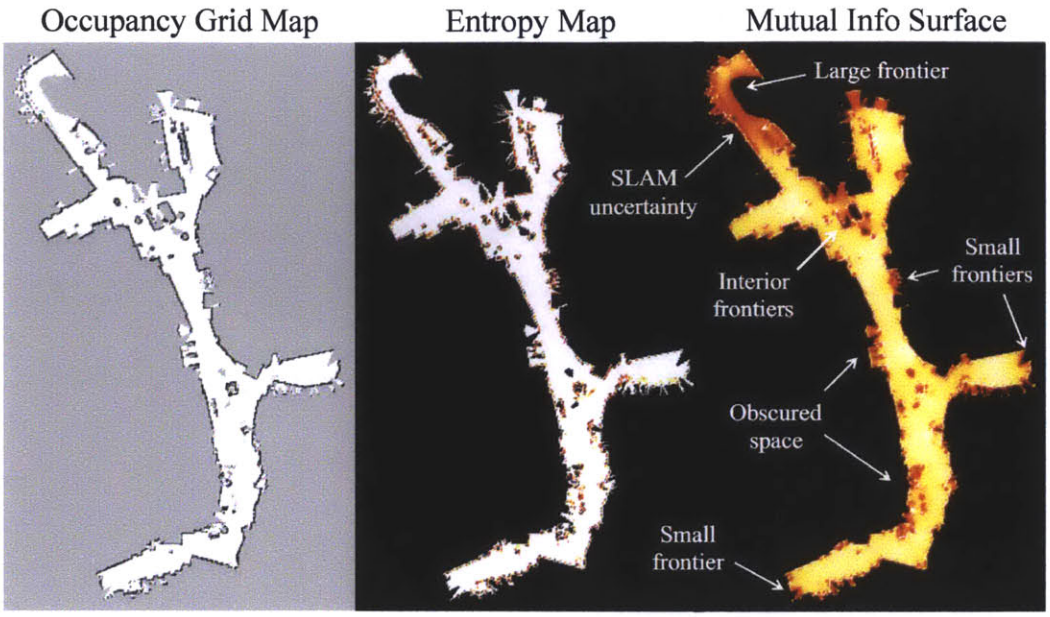


Figure 4-11: The occupancy grid map, entropy map, and mutual information reward surface generated after a 50 minute SLAM experiment performed on the first floor of the MIT Stata Center. *Left:* The resulting occupancy grid map. *Middle:* The entropy map showing cells of high uncertainty. *Right:* A more geometrically relevant mutual information reward surface that highlights locations of frontiers, obscured space, and potential SLAM convergence issues. Note that the linear color scales from white to black represent $[0, 1]$ bit for the entropy map and $[9, 48]$ bit for the mutual information reward surface.

4.5 Summary

Although occupancy grid mapping is a fundamental information acquisition tasks in robotics, it is also one of the simplest from a probabilistic perspective. We next formulate in Chapter 5 the generalized information acquisition problem using many robots equipped with sensors. By applying the gradient of the mutual information reward function to these control robots, we derive the novel class of decentralized mutual information-based gradient-ascent controllers.

4.6 Proofs

This section contains proofs for all lemmas, corollaries, and propositions stated in this chapter.

Proof 25 (Lemma 6) *From information theory (Cover and Thomas, 1991), the mutual information contribution is defined as*

$$I(M^{[i]}; Z_{k+1}^{[j]}|z_{1:k}) = \int_{z\in\mathcal{Z}} \sum_{m^{[i]}\in\{0,1\}} \log\left(\frac{\mathbb{P}(M^{[i]}=m^{[i]}|Z_{k+1}^{[j]}=z, z_{1:k})}{\mathbb{P}(M^{[i]}=m^{[i]}|z_{1:k})}\right)$$
$$\times\, \mathbb{P}(M^{[i]}=m^{[i]}|Z_{k+1}^{[j]}=z, z_{1:k})\mathbb{P}(Z_{k+1}^{[j]}=z|z_{1:k})dz. \tag{4.13}$$

Since the complement of $\mathbb{P}(M^{[i]}|z_{1:k})$ *and* $\mathbb{P}(M^{[i]}|Z_{k+1}^{[j]}, z_{1:k})$ *are equal to* $1-\mathbb{P}(M^{[i]}|z_{1:k})$ *and* $1-\mathbb{P}(M^{[i]}|Z_{k+1}^{[j]}, z_{1:k})$*, respectively, we have from* (4.1) *that*

$$\mathbb{P}(M^{[i]}=m^{[i]}|z_{1:k}) = \frac{(r_k^{[i]}-1)m^{[i]}+1}{r_k^{[i]}+1}, \tag{4.14}$$

and

$$\mathbb{P}(M^{[i]}=m^{[i]}|Z_{k+1}^{[j]}=z, z_{1:k}) = \frac{(r_k^{[i]}\delta^{[i,j]}(z)-1)m^{[i]}+1}{r_k^{[i]}\delta^{[i,j]}(z)+1}. \tag{4.15}$$

Substituting (4.14) *and* (4.15) *into* (4.13) *gives* (4.4)*, where*

$$f(\delta, r) := \log\left(\frac{r+1}{r+\delta^{-1}}\right) - \frac{\log(\delta)}{r\delta+1}. \tag{4.16}$$

Immediately, the first algebraic property of $f(\delta, r)$ *follows directly from* (4.16) *with* $\delta = 1$ *from the inverse sensor model.*

To prove the second algebraic property of $f(\delta, r)$*, consider a measurement value* $z \in \mathcal{Z}$ *that results in* $\delta \neq 1$*. Taking the partial derivative of* $f(\delta, r)$ *with respect to* r*, we get that*

$$\frac{\partial f(\delta, r)}{\partial r} = \frac{1}{r+1} - \frac{1}{r+\delta^{-1}} + \frac{\delta\log(\delta)}{(\delta r+1)^2}. \tag{4.17}$$

Now define the belief threshold $\tau(\delta)$ *as*

$$\tau(\delta) := \frac{\delta\log(\delta)-\delta+1}{\delta(\delta-\log(\delta)-1)}, \tag{4.18}$$

and note that $\tau(\delta) \in (0, \infty)$ *for all* $r \in (0, \infty)$. *Substituting* $(1+\epsilon)\tau(\delta)$ *for* r *in* (4.17) *gives*

$$\left.\frac{\partial f(\delta, r)}{\partial r}\right|_{r=(1+\epsilon)\tau(\delta)} = \frac{-\epsilon\delta\xi\beta^3}{(\epsilon\xi + (\delta-1)^2)(\epsilon\xi+\gamma)^2}, \tag{4.19}$$

where ξ, β, *and* γ *are defined as* $(\delta\log(\delta)-\delta+1)$, $(\delta-\log(\delta)-1)$, *and* $(\log(\delta)-\delta\log(\delta))$, *respectively. By applying the natural logarithm inequalities of* $(\delta-1) \leq \delta\log(\delta) \leq \delta(\delta-1)$, *we have that* ξ, β, *and* γ *are all positive. This implies that* (4.19) *is finite and positive for all* $\epsilon \in (0, \infty)$, *which gives the desired result for negative monotonicity.*

For positive monotonicity, we again apply the previously mentioned natural logarithm inequalities to have that ξ *is less than or equal to both* $(\delta-1)^2$ *and* γ. *This implies that both* $(\epsilon\xi + (\delta-1)^2)$ *and* $(\epsilon\xi+\gamma)$ *are finite and positive for* $\epsilon \in (-1, 0)$, *hence* (4.19) *is finite and negative.*

Proof 26 (Corollary 8) *Given a measurement value* $z \in \mathcal{Z}$, *we have for all cells* i *and measurements* j *that* $f(\delta^{[i,j]}(z), r_k^{[i]})$ *is independent of the robot's position* c_{k+1}. *Thus, the partial derivative of* (4.4) *with respect to the robot position is simply* (4.5). *To show that this gradient exists, we use the total law of probability and the product rule of differentiation to write*

$$\begin{aligned}\frac{\partial \mathbb{P}(Z_{k+1}^{[j]}|z_{1:k})}{\partial c_{k+1}} &= \sum_{m\in\mathcal{M}} \frac{\partial \mathbb{P}(Z_{k+1}^{[j]}|M=m, z_{1:k})}{\partial c_{k+1}} \mathbb{P}(M=m|z_{1:k}) \\ &+ \sum_{m\in\mathcal{M}} \mathbb{P}(Z_{k+1}^{[j]}|M=m, z_{1:k}) \frac{\partial \mathbb{P}(M=m|z_{1:k})}{\partial c_{k+1}},\end{aligned} \tag{4.20}$$

where $\mathcal{M}$ *is the set of all possible occupancy grid maps. Since* $\mathbb{P}(M^{[i]}|z_{1:k})$ *is independent of the robot's position, the second summation on the right hand side of* (4.20) *is equal to zero. The proof follows by evaluating* (4.5) *as an improper integral.*

Proof 27 (Lemma 10) *The proof proceeds by induction on* k. *If* $k = K$, *then the beliefs are equal to their realized values. Thus, this lemma trivially holds for* $k = K$. *Let us now assume that the result is true for some* $k \geq K$. *Note that if* $i \notin \mathcal{I}_{k+1}$, *then the lemma trivially holds for* $k+1$. *Now suppose* $i \in \mathcal{I}_{k+1}$, *and define the conditioning*

event $\mathcal{S}_K$ *to be*

$$\mathcal{S}_K = \{M = m\} \cap \{\tilde{M} = m\} \cap \{R_K^{[i]} < \tilde{R}_K^{[i]}\}$$

for some $m \in \mathcal{M}$. *For the first experiment* E, *we have that*

$$\begin{aligned}
\mathbb{E}\big(R_{k+1}^{[i]}|\mathcal{S}_K\big) &= \mathbb{E}\big(\delta^{[i]}(Z_{k+1})R_k^{[i]}|\mathcal{S}_K\big) \\
&= \mathbb{E}\big(\delta^{[i]}(Z_{k+1})|\mathcal{S}_K\big)\mathbb{E}\big(R_k^{[i]}|\mathcal{S}_K\big) \\
&= \prod_{j=1}^{n_z} \mathbb{E}\big(\delta^{[i,j]}(Z_{k+1}^{[j]})|\mathcal{S}_K\big)\mathbb{E}\big(R_k^{[i]}|\mathcal{S}_K\big), \qquad (4.21)
\end{aligned}$$

where the second and third equality come from the property of independence between observations and measurements, respectively. In addition, (4.21) *is identical to the expression for the second experiment* $\tilde{E}$ *except with* $R_k^{[i]}$, Z_k, *and* $Z_k^{[j]}$ *being replaced by* $\tilde{R}_k^{[i]}$, $\tilde{Z}_k$, *and* $\tilde{Z}_k^{[j]}$, *respectively, for all times* k. *From the inverse sensor model and the law of total probability, we have for the first experiment* E *that*

$$\begin{aligned}
\mathbb{E}\big(\delta^{[i,j]}(Z_{k+1})|\mathcal{S}_K\big) &= \sum_{m' \in \mathcal{M}} \mathbb{P}\big(M = m'|\mathcal{S}_K\big) \int_{z \in \mathcal{Z}} \mathbb{P}\big(Z_{k+1}^{[j]} = z|\mathcal{S}_K, M = m'\big)\delta^{[i,j]}(z)dz \\
&= \int_{z \in \mathcal{Z}} \mathbb{P}\big(Z_{k+1}^{[j]} = z|\mathcal{S}_K\big)\delta^{[i,j]}(z)dz \qquad (4.22)
\end{aligned}$$

where the last equally comes from the property that the event $\mathcal{S}_K$ *contains the condition* $M = m$. *Since* (4.22) *also contains the condition* $\tilde{M} = m$ *for the second experiment* $\tilde{E}$, *and both experiments have identical sensor and inverse sensor models, we have that* $\mathbb{E}\big(\delta^{[i,j]}(\tilde{Z}_{k+1}^{[j]})|\mathcal{S}_K\big)$ *is equal to* $\mathbb{E}\big(\delta^{[i,j]}(Z_{k+1}^{[j]})|\mathcal{S}_K\big)$. *Thus from* (4.21), *we have that*

$$\frac{\mathbb{E}\big(R_{k+1}^{[i]}|\mathcal{S}_k\big)}{\mathbb{E}\big(R_k^{[i]}|\mathcal{S}_k\big)} = \frac{\mathbb{E}\big(\tilde{R}_{k+1}^{[i]}|\mathcal{S}_k\big)}{\mathbb{E}\big(\tilde{R}_k^{[i]}|\mathcal{S}_k\big)}, \qquad (4.23)$$

and thus this proof holds for $R_K^{[i]}$ *being less than* $\tilde{R}_K^{[i]}$ *given that* M *and* $\tilde{M}$ *are both equal to the map value* m. *To complete the proof, we note that the derivation of* (4.23) *is not specific to the map value* $m \in \mathcal{M}$ *nor the numerical relationship between* $R_K^{[i]}$

and $\tilde{R}_K^{[i]}$.

Proof 28 (Corollary 11) *Considering* (4.21) *in Lemma 10, the proof follows from Assumption 1.*

Proof 29 (Corollary 12) *From Corollary 11, there exists a time* $\varkappa \geq K$ *such that for all times* $k \geq \varkappa$*, the two expected beliefs* $\mathbb{E}(R_k^{[i]}|\mathcal{A}_K)$ *and* $\mathbb{E}(\tilde{R}_k^{[i]}|\mathcal{A}_K)$ *are both smaller than some* $\epsilon_0 > 0$ *or larger than some* $\epsilon_1 < \infty$*, depending on whether* $M^{[i]}$ *and* $\tilde{M}^{[i]}$ *are both of value* 0 *or* 1*, respectively. From Assumption 2, this implies that*

$$\begin{aligned}\mathbb{P}(Z_{k+1}^{[j]}|\mathcal{B}_{K,k}) &= \mathbb{P}\Big(Z_{k+1}^{[j]}|\mathcal{B}_{K,k} \cap \bigcap_{i\in\mathbb{I}} \big\{M^{[i]} = \tilde{M}^{[i]}\big\}\Big) \\ &= \mathbb{P}\Big(\tilde{Z}_{k+1}^{[j]}|\mathcal{B}_{K,k} \cap \bigcap_{i\in\mathbb{I}} \big\{\tilde{M}^{[i]} = M^{[i]}\big\}\Big) = \mathbb{P}(\tilde{Z}_{k+1}^{[j]}|\mathcal{B}_{K,k}),\end{aligned}$$

where the third equality is due to all other expected beliefs $\mathbb{E}(R_k^{[\ell]}|\mathcal{A}_K)$ *and* $\mathbb{E}(\tilde{R}_k^{[\ell]}|\mathcal{A}_K)$ *being equal between experiments for any cell* $\ell \in \mathbb{I}^c$.

Proof 30 (Lemma 16) *From the law to total probability, we have for the measurement's prior* $\mathbb{P}(Z_{k+1}^{[j]}|z_{1:k})$ *that*

$$\mathbb{P}(Z_{k+1}^{[j]}|z_{1:k}) = \sum_{m\in\mathcal{M}} \mathbb{P}(Z_{k+1}^{[j]}|M=m)\mathbb{P}(M=m|z_{1:k}),$$

noting that the sensor model $\mathbb{P}(Z_{k+1}^{[j]}|M)$ *is independent of the previous observations* $z_{1:k}$*. Since the posterior distribution* $\mathbb{P}(M|z_{1:k})$ *is independent of* $Z_{k+1}^{[j]}$ *and the index set* $\mathcal{I}_{k+1}^{[j]}$ *is ordered, Assumption 14 implies that*

$$\begin{aligned}\mathbb{P}(Z_{k+1}^{[j]}|z_{1:k}) = \mathbb{P}(Z_{k+1}^{[j]}|M=\boldsymbol{0}) \prod_{\ell\in\mathcal{I}_{k+1}^{[j]}} \mathbb{P}(M^{[\ell]}=0|z_{1:k}) + \\ \sum_{i\in\mathcal{I}_{k+1}^{[j]}} \mathbb{P}(Z_{k+1}^{[j]}|M=\boldsymbol{e}_i)\mathbb{P}(M=\boldsymbol{e}_i|z_{1:k}) \prod_{\ell<i} \mathbb{P}(M^{[\ell]}=0|z_{1:k}) \qquad (4.24)\end{aligned}$$

where $\ell < i$ *denotes the sequence of indices in* $\mathcal{I}_{k+1}^{[j]}$ *appearing before* i*. Substituting* (4.15) *into* (4.24) *gives* (4.7).

Proof 31 (Corollary 17) *Noting that* $\mathbb{P}(Z_{k+1}^{[j]} = z|M = \mathbf{0})$ *is independent of robot position* c_{k+1}, *i.e.,* $\partial\mathbb{P}(Z_{k+1}^{[j]} = z|M = \mathbf{0})/\partial c_{k+1} = 0$, *the proof follows directly from Lemma 16.*

Proof 32 (Lemma 19) *By considering the mutual information reward function* (4.2), *i.e.,*

$$I(M; Z_{k+1}|z_{1:k}) = \sum_{j=1}^{n_z} \sum_{i \in \mathcal{I}_{k+1}^{[j]}} I(M^{[i]}; Z_{k+1}^{[j]}|z_{1:k}),$$

the proof follows from Lemma 6 and Lemma 16.

Proof 33 (Corollary 20) *The proof follows directly from Corollary 8 and Lemma 19.*

Proof 34 (Corollary 21) *The proof follows directly from Lemma 19 with integration over possible measurement values being approximated by a sum of step size* λ_z^{-1}.

Proof 35 (Corollary 22) *The proof follows directly from Corollary 20 with integration over possible measurement values being approximated by a sum of step size* λ_z^{-1}.

Proof 36 (Proposition 23) *For any measurement* j, *the perceptual range set* $\mathcal{I}_k^{[j]}$ *has at most* $\lceil 2+\sqrt{2}\max(\mathcal{Z})\lambda_m \rceil$ *elements. Given the set of range measurements* $\mathcal{Z}$, *we have that the space complexity of constructing* $\mathcal{I}_k^{[j]}$ *is* $O(\lambda_m)$, *while its corresponding time complexity using a naïve ray casting algorithm for* RayCast *is also* $O(\lambda_m)$. *Thus, the time and space complexities of Algorithm 3 are both* $O(\lambda_m)$. *In addition, the time and space complexities of both Algorithm 4 and Algorithm 5 are* $O(\lambda_m)$ *and* $O(1)$, *respectively.*

To finish the proof, we note that Algorithm 1 constructs sets in lines 3 *and* 4 *of size* $O(\lambda_m)$. *Both Algorithm 4 and Algorithm 5 are called within three* for *loops of sizes* n_z, $O(\lambda_m)$ *and* $O(\lambda_z)$.

Proof 37 (Proposition 24) *Let* $M^{[\mathcal{I}_{k+1}^{[j]}]}$ *denote the random variable tuple composed of all cells within the* j*th measurement's perceptual range, i.e., we define* $M^{[\mathcal{I}_{k+1}^{[j]}]} :=$

$(M^{[i_1]}, \ldots, M^{[i_N]})$ *with* $\mathcal{I}_{k+1}^{[j]} = \{i_1, \ldots, i_N\}$. *Without the assumption of sensor obscuration, the expression for the* j*th measurement's prior becomes*

$$\mathbb{P}(Z_{k+1}^{[j]}|z_{1:k}) = \sum_{m \in \mathcal{M}^{[\mathcal{I}_{k+1}^{[j]}]}} \mathbb{P}(Z_{k+1}^{[j]}|M^{[\mathcal{I}_{k+1}^{[j]}]} = m)\mathbb{P}(M^{[\mathcal{I}_{k+1}^{[j]}]} = m|z_{1:k}). \tag{4.25}$$

The time complexity of an algorithm used to calculate (4.25) *is exponential with* $O(2^{\lambda_m})$. *To finish the proof, we update the complexity for Algorithm 1 given in Proposition 23.*

Chapter 5

Gradient-ascent control for multi-robot inference tasks

We consider in this chapter the task of using a robotic sensor network to infer the state of an environment, for example to collect military intelligence, gather scientific data, or monitor ecological events of interest. Our goal is to enable the robots to efficiently, robustly, and provably learn their environment and reason where to make future sensor measurements. To this end, we present the derivation, analysis, and implementation of a novel class of decentralized controllers that continuously move sensing robots to better observe their environment. Built on an information-theoretic foundation, these gradient-ascent controllers attempt to maximize the expected information gain of the robots' future sensor observations when paired with a sequential Bayesian filter. They allow for general convergent results, and lead to practical control strategies that account for the limited computational resources of the robots, the decentralized nature of their computation, and the finite bandwidth of their communication network. The controllers are able to dynamically adapt to changing network connectivity and be scalable with respect to the number of robots, enabling robot teams of large size and heterogeneous composition. Lastly, they exhibit anytime performance when augmented with appropriate distributed algorithms to calculate statistical distributions describing the likelihood of system-wide observations.

The key to our solution is in the use of non-parametric[1] sample-based representations of the probability distributions present in the system. This construction leads to scalability with respect to the number of robots, and allows for completely decentralized computation under few probabilistic assumptions. Specifically, we do not assume that any probability distribution can be accurately represented by a Gaussian distribution. Assuming Gaussianity significantly simplifies many aspects of the system. However, Gaussian distributions often do not adequately represent the characteristics of realistic environments and sensors, and may result in misleading inferences and poor controller performance. Instead, we approximate likely local observations using sample sets that are distributively formed and provably unbiased. In addition to enabling the information seeking controller to move the robots and orient their sensors, these non-parametric sampled distributions can be used in conjunction with a decentralized Bayesian filter to update the robots' beliefs.

5.1 Problem Formulation

We motivate our approach with an information-theoretic justification of the mutual information reward function, then develop the problem formally for a single robot followed by the centralized multi-robot case with an ideal network.

5.1.1 Mutual information reward for Bayesian inference

We wish to infer the state of an environment from measurements obtained by a number of robots equipped with sensors, e.g., see Figure 5-1. Ideally, we would represent the potentially time varying state in a continuous manner. However, the robots' inference calculations happen at discrete times, and for this thesis we assume that all robots perform these calculations synchronously at a constant rate of $1/T_s$. Thus at time $t = kT_s$, where k again denotes the discrete time step, we model the

[1] We use the term *non-parametric* to convey that we do not assume that the statistics of the involved random variables can be exactly described by particular distributions defined using a finite number of parameters.

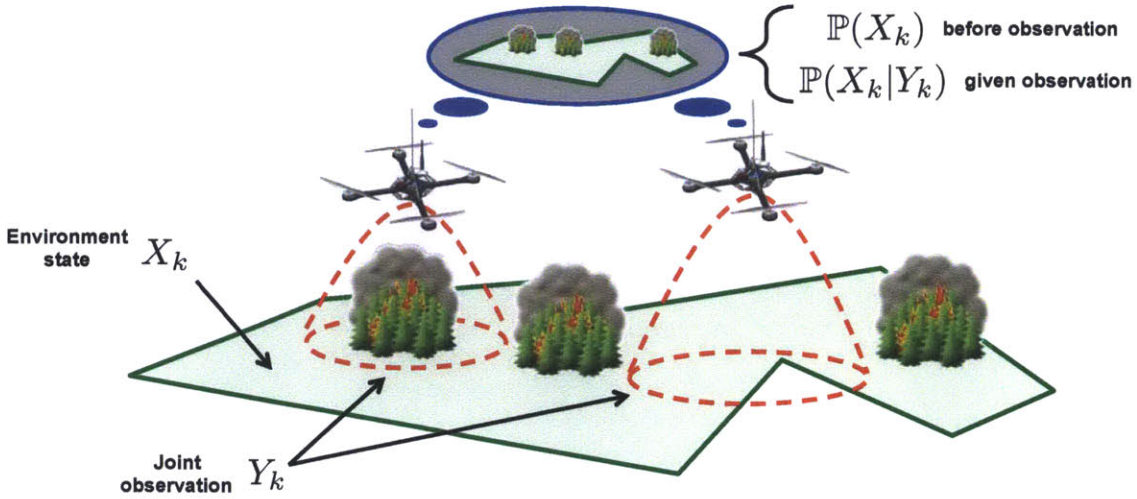


Figure 5-1: A centralized multi-robot system within a probabilistic framework. The robots observe the state of the environment using sensors of finite footprints. The joint measurement distribution describe the accuracy of the joint observations, which through Bayes' Rule provide the relationship between the system's prior distribution $\mathbb{P}(X_k)$ and its posterior distribution $\mathbb{P}(X_k|Y_k)$.

environment state as a discrete-time random variable X_k that takes values from an alphabet $\mathcal{X}$.

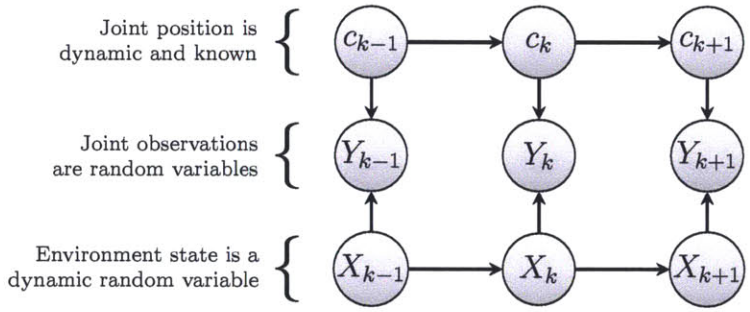


Figure 5-2: Dynamic Bayesian network for multi-robot inference tasks. Note that since the joint position is known, we show in the top row the realized values c_k instead of corresponding random variables.

Our goal is to enable the inference calculations necessary for collectively estimating the environment state and reducing uncertainty in the system. Each robot forms an observation from sensor measurements influenced by noise and other sources of error. We consider the observations of all robots together as a single joint observation, which we model as a discrete-time random variable Y_k that takes values from an alphabet $\mathcal{Y}$. The relationship between the true state and the noisy joint observation is described

by the joint sensor model $\mathbb{P}(Y_k|X_k)$. Sensing may be interpreted as using a noisy channel, and since the sensors are attached to the robots, the joint sensor model is dependent on the position of the robots and the orientation of their sensors. On the dynamic Bayesian network shown in Figure 5-2, we can use a joint observation and Bayes' Rule to compute the system's posterior distribution

$$\mathbb{P}(X_k|Y_k) = \frac{\mathbb{P}(X_k)\mathbb{P}(Y_k|X_k)}{\int\limits_{x\in\mathcal{X}} \mathbb{P}(X_k = x)\mathbb{P}(Y_k|X_k = x)dx}, \tag{5.1}$$

where $\mathbb{P}(X_k)$ is the system's prior distribution.

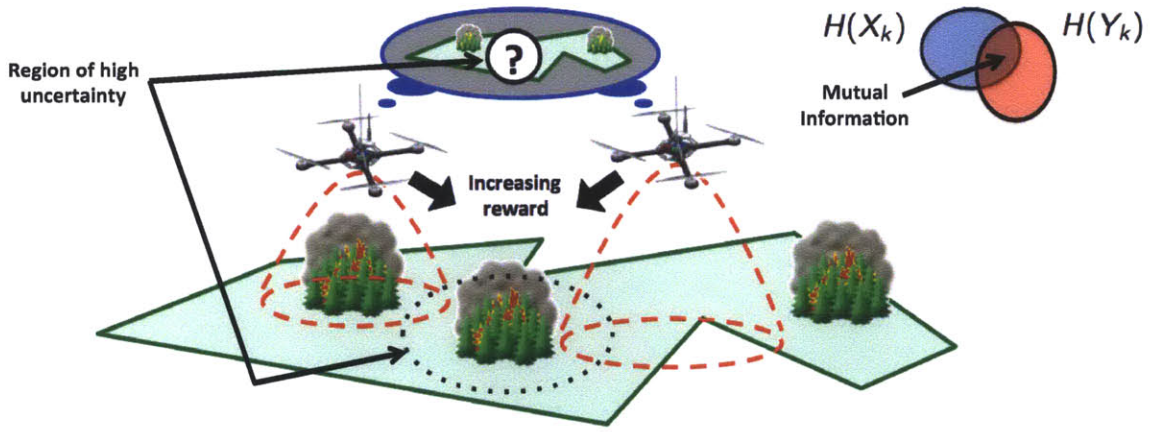


Figure 5-3: Robots moving their sensors' field of view towards a region of the environment that corresponds to high uncertainty. The movement happens in a direction of increasing mutual information (i.e., reward) since this direction corresponds to a decrease in conditional entropy of the environment state given the next joint observation. *Top right:* Mutual information can be visualized in a Venn diagram as the overlap of the entropy of the joint observation, $H(Y_k)$, and the entropy of the environment state, $H(X_k)$.

Since our objective is to best infer the environment state, we are motivated to move the robots and their sensors into a position that minimizes the expected uncertainty of the environment state after receiving the next joint observation. Equivalent to our mapping objective in Chapter 4, our optimization objective is formally stated as minimizing the conditional entropy

$$H(X_k|Y_k) = H(X_k) - I(X_k; Y_k),$$

where $H(X_k)$ is the entropy of the environment state and $I(X_k; Y_k)$ is the mutual information between the environment state and the next joint observation. Since the entropy of the environment state before receiving an observation is independent of the position of the robots, minimizing the conditional entropy is equivalent to maximizing mutual information. Hence, we define the reward function for the system to be

$$\begin{aligned} U_k &:= I(X_k; Y_k) \\ &= \int_{y \in \mathcal{Y}} \int_{x \in \mathcal{X}} \mathbb{P}(Y_k = y | X_k = x) \mathbb{P}(X_k = x) \log \left(\frac{\mathbb{P}(X_k = x | Y_k = y)}{\mathbb{P}(X_k = x)} \right) dx\, dy. \end{aligned} \quad (5.2)$$

Figure 5-3 illustrates the concept of mutual information reward and how it relates to the movement of the robots. We are interested in a class of controllers that use a gradient-ascent approach with respect to the reward function U_k, leading to the following theorem.

Theorem 38 (Gradient of the reward function) *The gradient of the reward function U_k with respect to a single robot's position $c_t^{[i]}$ at continuous time $t \in \big((k-1)T_s, kT_s\big]$ is given by*

$$\frac{\partial U_k}{\partial c_t^{[i]}} = \int_{y \in \mathcal{Y}} \int_{x \in \mathcal{X}} \frac{\partial \mathbb{P}(Y_k = y | X_k = x)}{\partial c_t^{[i]}} \mathbb{P}(X_k = x) \log \left(\frac{\mathbb{P}(X_k = x | Y_k = y)}{\mathbb{P}(X_k = x)} \right) dx\, dy. \quad (5.3)$$

5.1.2 Single robot case

Consider a single robot, denoted i, that moves in a configuration space $\mathcal{C}^{[i]} \subset \mathbb{R}^{r_c^{[i]}} \times \mathbb{S}^{s_c^{[i]}}$. This space describes both the position of the robot platform and the orientation of its sensors, and does not need to be the same space as the environment, denoted $\mathcal{Q} \subset \mathbb{R}^{r_q} \times \mathbb{S}^{s_q}$. For example, if we have a planar environment within $\mathbb{R}^2$, we could have a ground robot with an omnidirectional sensor moving in $\mathbb{R}^2$ or a flying robot with a gimbaled sensor moving in $\mathbb{R}^3 \times \mathbb{S}^3$. Let $c_t^{[i]} \in \mathcal{C}^{[i]}$ denote the robot's position at continuous-time t, with $c_k^{[i]}$ being shorthand for $c_t^{[i]}$ with $t = kT_s$.

Let this robot have a *belief* of the environment state, which is represented by its prior distribution $\mathbb{P}^{[i]}(X_k)$. We model each robot's local observation as a random

variable $Y_k^{[i]}$ that takes values from an alphabet $\mathcal{Y}^{[i]}$ and is characterized by the local sensor model $\mathbb{P}(Y_k^{[i]}|X_k)$. We assume that the local sensor model is known a priori by the robot, or in other words, its sensors are *calibrated.*

Assumption 39 (Calibrated sensors) *For all positions $c_t^{[i]} \in \mathcal{C}^{[i]}$, we have that the ith robot correctly (or conservatively) knows its local sensor model $\mathbb{P}(Y_k^{[i]}|X_k)$.*

Using its sensor model, the robot is able to compute its posterior distribution

$$\mathbb{P}(X_k|Y_k^{[i]}) = \frac{\mathbb{P}^{[i]}(X_k)\mathbb{P}(Y_k^{[i]}|X_k)}{\int\limits_{x\in\mathcal{X}} \mathbb{P}^{[i]}(X_k = x)\mathbb{P}(Y_k^{[i]}|X_k = x)dx}, \tag{5.4}$$

which is used in conjunction with its state transition distribution $\mathbb{P}^{[i]}(X_{k+1}|X_k)$ and its realized local observation $y_k^{[i]} \in \mathcal{Y}^{[i]}$ to form at time $t = (k+1)T_s$ its new prior distribution

$$\mathbb{P}^{[i]}(X_{k+1}) = \frac{\int\limits_{x\in\mathcal{X}} \mathbb{P}^{[i]}(X_{k+1}|X_k = x)\mathbb{P}(X_k = x|Y_k^{[i]} = y_k^{[i]})}{\int\limits_{x'\in\mathcal{X}}\int\limits_{x\in\mathcal{X}} \mathbb{P}^{[i]}(X_{k+1} = x'|X_k = x)\mathbb{P}(X_k = x|Y_k^{[i]} = y_k^{[i]})dx\,dx'}. \tag{5.5}$$

Equations (5.4) and (5.5) form the well-known duet of *update* and *prediction*, respectively, in sequential Bayesian estimation.

5.1.3 Centralized multi-robot case with ideal network

Given a centralized system with an ideal network (i.e., complete with infinite bandwidth and no latency), the multi-robot case with n_r robots is a simple extension of the single robot case with a common prior distribution $\mathbb{P}^{[i]}(X_k) = \mathbb{P}(X_k)$ and state transition distribution $\mathbb{P}^{[i]}(X_{k+1}|X_k) = \mathbb{P}(X_{k+1}|X_k)$ for all robots $i \in \{1, \ldots, n_r\}$. Let the system of joint position $c_t = (c_t^{[1]}, \ldots, c_t^{[n_r]}) \in \mathcal{C} = \prod_{i=1}^{n_r} \mathcal{C}^{[i]}$ be synchronous in that the robots' local observations are simultaneously received at a sampling rate of $1/T_s$. We model the joint observation as an n_r-tuple random variable $Y_k = (Y_k^{[1]}, \ldots, Y_k^{[n_r]})$ that takes values from the Cartesian product of all the robots' local observation alphabets $\mathcal{Y} = \prod_{i=1}^{n_r} \mathcal{Y}^{[i]}$.

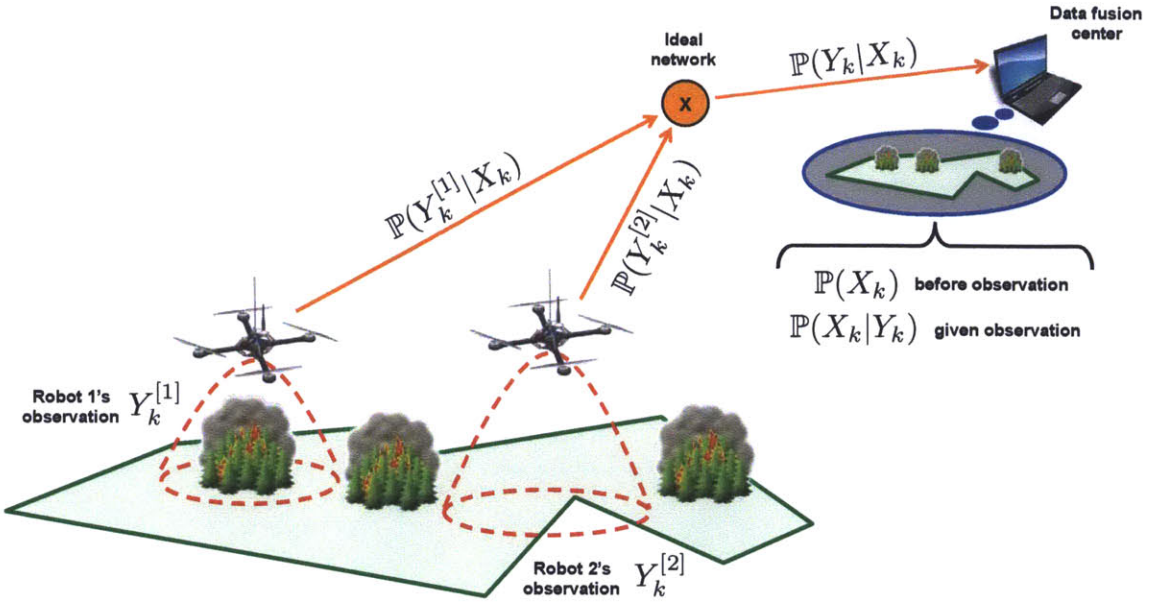


Figure 5-4: A centralized multi-robot system with an ideal network. The robots synchronously receive a local observation then transmit the corresponding likelihoods over an ideal network to a data fusion center. The joint measurement likelihood is then used to update the system's prior distribution to form the posterior distribution.

We assume that the noise on the observations are uncorrelated between robots, or in other words, that the robots' local observations are *conditionally independent.*

Assumption 40 (Conditionally independent observations) *For all joint positions $c_t \in \mathcal{C}$, we have that*

$$\mathbb{P}(Y_k|X_k) = \prod_{i=1}^{n_r} \mathbb{P}(Y_k^{[i]}|X_k). \tag{5.6}$$

Since the sensors of any two robots are physically detached from each other, we can expect that correlated noise is the result of environmental influences. The more these influences are accounted for within the environment state, the more accurate the assumption of conditional independence becomes. Employing Assumption 40, the

system's posterior distribution from (5.4) becomes

$$\mathbb{P}(X_k|Y_k) = \frac{\mathbb{P}(X_k)\prod_{i=1}^{n_r}\mathbb{P}(Y_k^{[i]}|X_k)}{\int_{x\in\mathcal{X}}\mathbb{P}(X_k=x)\prod_{i=1}^{n_r}\mathbb{P}(Y_k^{[i]}|X_k=x)dx}, \tag{5.7}$$

and the system's prior distribution from (5.5) becomes

$$\mathbb{P}(X_{k+1}) = \frac{\int_{x\in\mathcal{X}}\mathbb{P}(X_{k+1}|X_k=x)\mathbb{P}(X_k=x|Y_k=y_k)}{\int_{x'\in\mathcal{X}}\int_{x\in\mathcal{X}}\mathbb{P}(X_{k+1}=x'|X_k=x)\mathbb{P}(X_k=x|Y_k=y_k)dx\,dx'},$$

where $y_k = (y_k^{[1]}, \ldots, y_k^{[n_r]}) \in \mathcal{Y}$ is the value of the realized joint observation. Note that there is one common prior distribution (i.e., belief) and posterior distribution (i.e., update) for the centralized system, as illustrated in Figure 5-4. For the decentralized system, we commonly use the notation $\mathbb{P}^{[i]}$ to represent distributions for a particular robot. The notable exception concerns the local sensor models $\mathbb{P}(Y_k^{[i]}|X_k)$, where writing $\mathbb{P}^{[i]}(Y_k^{[i]}|X_k)$ is unnecessary and thus the extra superscript is omitted for clarity.

5.2 Multi-robot coordination

Consider the system approach for decentralized inference and coordination shown in Figure 5-5. In this section, we discuss the derivation and analysis of the decentralized mutual information-based gradient-ascent controller. We discuss properties of distributed algorithms approximating the soon to be discussed joint measurement likelihoods in Chapter 6.

5.2.1 Gradient-ascent control

At any given time, the robot can choose a control action $u_t^{[i]}$ taken from a control space $\mathcal{U}_t^{[i]} \subset \mathbb{R}^{r_c^{[i]}} \times \mathbb{S}^{s_c^{[i]}}$. For simplicity, we assume that the robots have continuous-time *integrator dynamics.*

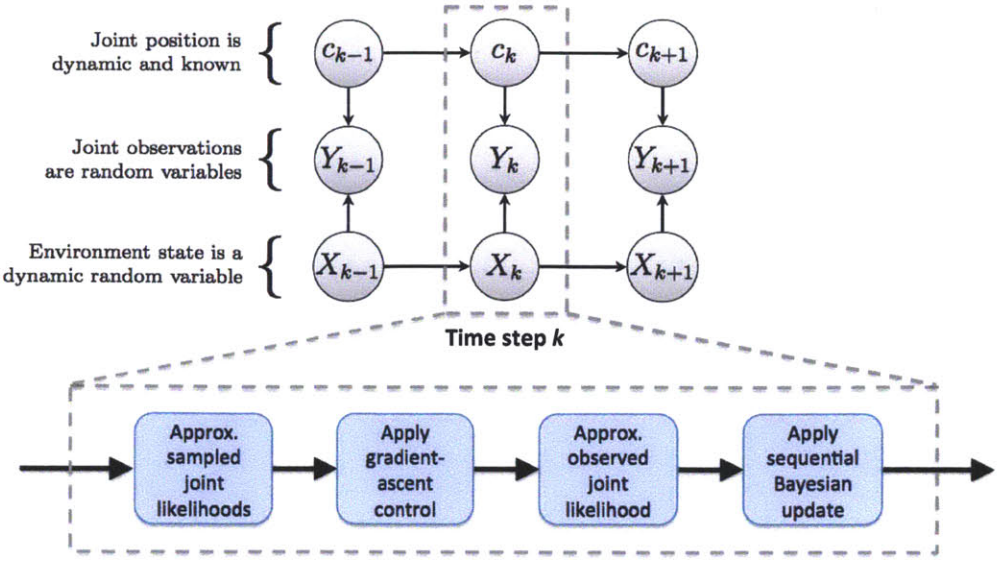


Figure 5-5: Our approach to distributed inference and coordination. Starting with its local belief (i.e., prior distribution), each robot distributively approximates the measurement likelihoods describing a sampled set of joint observations, then applies its decentralized mutual information-based gradient-ascent control. Once the observation is realized, each robot distributively approximates the corresponding joint measurement likelihood to update its belief (i.e., posterior distribution), enabling decentralized control during the next time step.

Assumption 41 (Integrator dynamics) *The ith robot is governed by the dynamics*

$$\frac{dc_t^{[i]}}{dt} = u_t^{[i]}. \tag{5.8}$$

The assumption of integrator dynamics is common in the multi-robot coordination literature (Bullo et al., 2009). In our applications using the quadrotor flying robot platform, we found that generating position commands at a relatively slow rate (e.g., 1 Hz) and feeding these inputs into a relatively fast (e.g., 40 Hz) low-level position controller emulates integrator dynamics (5.8).

Consider the control objective from Section 5.1.1. We wish to move the robot system into a position that minimizes the conditional entropy of the environment state given the next joint observation. With respect to the reward function U_k, our objective is equivalent to solving the constrained optimization problem $\max_{c \in \mathcal{C}} U_k$. One

solution approach is to have the each robot calculate from (5.3) the partial derivative of the reward function with respect to that robot's position

$$\frac{\partial U_k}{\partial c_t^{[i]}} = \int_{y\in\mathcal{Y}} \int_{x\in\mathcal{X}} \frac{\partial \mathbb{P}(Y_k^{[i]} = y^{[i]} | X_k = x)}{\partial c_t^{[i]}} \mathbb{P}(X_k = x) \times \prod_{i'\in\{1,\ldots,n_r\}/\{i\}} \mathbb{P}(Y_k^{[i']} = y^{[i']} | X_k = x) \times \log\left(\frac{\mathbb{P}(X_k = x | Y_k = y)}{\mathbb{P}(X_k = x)}\right) dx\, dy, \tag{5.9}$$

then continuously move in a valid direction of increasing reward. Note that receiving an observation may induce instantaneous changes in the reward function's gradient even if the gradient is continuous on the configuration space during the time step. Because of this property, we use the phrase *convergent between observations* to describe the limit of a state given enough time between Bayesian filter updates. We considered this to be a useful property since it describes how the robots improve their positions based on the information at hand, i.e., prior to receiving the next observation. Considering (5.9) and its dependence on the sensor model gradient $\partial \mathbb{P}(Y_k^{[i]} | X_k)/\partial c_t^{[i]}$, we are interested in *finite range* sensors whose probabilistic models are *sufficiently smooth.*

Assumption 42 (Finite range sensor) *There exists some constant* $d^{[i]} > 0$ *such that for all positions* $c_t^{[i]} \in \mathcal{C}^{[i]}$ *with* $\min_{q\in\mathcal{Q}} \text{dist}(c_t^{[i]}, q) \geq d^{[i]}$, *we have that*

$$\frac{\partial \mathbb{P}(Y_k^{[i]} | X_k)}{\partial c_t^{[i]}} = 0, \tag{5.10}$$

where $\text{dist} : \mathcal{C}^{[i]} \times \mathcal{Q} \to \mathbb{R}_{\geq 0}$ *is a valid distance function.*

Assumption 43 (Sufficiently smooth sensor model) *The local sensor model* $\partial \mathbb{P}(Y_k^{[i]} | X_k)/\partial c_t^{[i]}$ *is differentiable on the configuration space* $\mathcal{C}^{[i]}$, *while the resulting gradient* $\partial \mathbb{P}(Y_k^{[i]} | X_k)/\partial c_t^{[i]}$ *is continuous on* $\mathcal{C}^{[i]}$.

The finite range assumption is arguably satisfied by any realistic robot sensor. For example, the range of an optical sensor on a UAV can be limited by atmospheric

visibility, imaging resolution, or (at the very extreme) the curvature of the Earth. In other words, at some distance away from the environment of interest, the change in local sensing uncertainty is independent of the motion of the robot. The assumption of a sufficiently smooth sensor model is more abstract; however, it is common that a robot's sensor model can be formulated to satisfy this assumption, e.g., see Section 5.2.2. We now state the main theorem for convergence and local optimality for mutual information-based gradient-ascent controllers.

Theorem 44 (Convergence and local optimality) *Suppose Assumptions 39-43 hold. Let the robots move in the same configuration space and sense a bounded environment that is a subset of the configuration space. Then for a positive scalar* $\gamma^{[i]}$, *the controller*

$$u_t^{[i]} = \gamma^{[i]} \frac{\partial U_k}{\partial c_t^{[i]}} \tag{5.11}$$

is convergent to zero between observations for all robots. In addition, an equilibrium joint position $c_* = (c_*^{[1]}, \ldots, c_*^{[n_r]})$ *defined by*

$$\left. \frac{\partial U_k}{\partial c_t^{[i]}} \right|_{c_t^{[i]} = c_*^{[i]}} = 0, \quad \forall i \in \{1, \ldots, n_r\}$$

is Lyapunov stable if and only if it is locally optimal with respect to maximizing the reward function U_k.

Remark 45 (Required local and global knowledge) *For the class of controllers presented in this thesis, we assume that each robot has knowledge of* i) *its local configuration* $c_t^{[i]}$; ii) *its local sensor model* $\mathbb{P}(Y_k^{[i]}|X_k)$; *and* iii) *the extent of the environment* $\mathcal{Q}$. *However, the gradient-ascent controller* (5.11) *also requires that each robot has global knowledge of* i) *the centralized prior* $\mathbb{P}(X_k)$; ii) *the joint position* c_t; *and* iii) *the joint sensor model* $\mathbb{P}(Y_k|X_k)$. *Thus, the controller is not distributed among the robots.*

Remark 46 (Intractability of the general form) *Consider both* X_k *and* Y_k *to be*

discrete-valued random variables that take values from alphabets of size $|\mathcal{X}|$ *and at most* $\max_i |\mathcal{Y}^{[i]}|^{n_r}$*, respectively. To calculate and store all possible instantiations from the posterior calculation* (5.7)*, an algorithm requires* $\max_i O(n_r|\mathcal{X}||\mathcal{Y}^{[i]}|^{n_r})$ *time and* $\max_i O(|\mathcal{X}||\mathcal{Y}^{[i]}|^{n_r})$ *space. In addition, an algorithm requires* $\max_i O(n_r|\mathcal{X}||\mathcal{Y}^{[i]}|^{n_r})$ *time to calculate the reward function's gradient* $\partial U_k/\partial c_t^{[i]}$*. Since the computational complexities of implementing the gradient-ascent controller* (5.11) *are exponential with respect to number of robots* n_r*, such an algorithm is not scalable.*

5.2.2 Occupancy grid mapping case

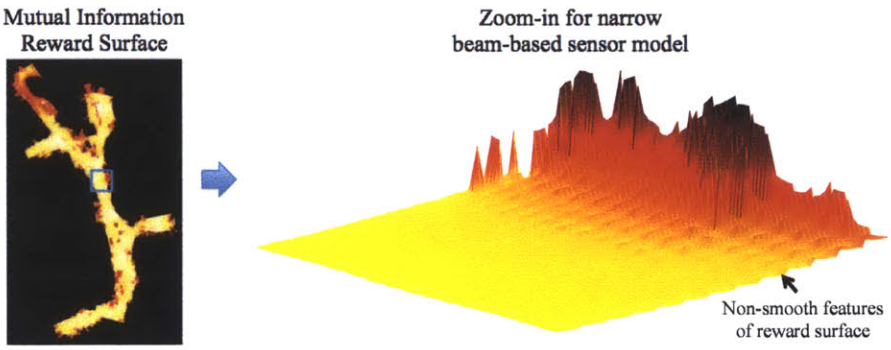


Figure 5-6: Non-smooth features of the mutual information reward surface for the narrow beam-based sensor model. Obtaining sufficient smoothness for the sensor model comes at the cost of computational complexity for our occupancy grid mapping robot.

For a single robot, consider the narrow beam-based sensor model discussed throughout Chapter 4. For control approaches that navigate directly on the mutual information reward surface, we see that "smooth" surfaces can be efficiently generate by only considering robot positions in the center of the occupancy grid cells, e.g., see Figure 4-1. However, as the robot is positioned elsewhere in the cells, "non-smooth" features of the mutual information reward surface become more apparent, e.g., see Figure 5-6. These discontinuities are the result of abrupt narrow beam transitions that occur along the cell boundaries, e.g., see Figure 5-8. The narrow beam-based sensor model is not sufficiently smooth as defined in Assumption 43, which suggests that a gradient-ascent controller employing Algorithm 2 may perform poorly. Note

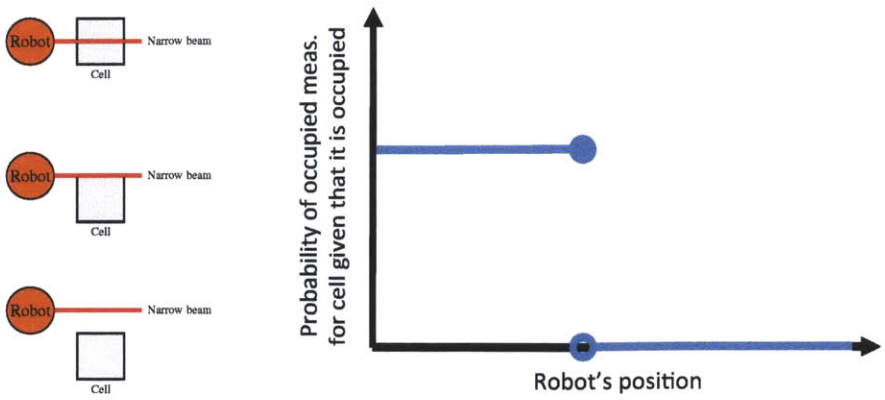


Figure 5-7: Non-smooth characteristics of the narrow beam-based sensor model. *Left:* As the robot moves in an orthogonal direction to a grid cell, one if its narrow beams instantaneously transitions from measuring within the cell to measuring completely outside the cell. *Right:* The measurement probability associated with this cell experiences a discontinuous "jump" with respect to the robot's position. The resulting mutual information reward surface on which the robot navigates will contain discontinuities, thus precluding the use of LaSalle's Invariance Principle to prove convergence between observations for the corresponding controller.

that Assumption 43 is a sufficient condition and not a necessary one, thus our analysis from Theorem 44 does not prove the lack of convergence or local optimality for the corresponding controller.

Now we concentrate on formulating an appropriate beam-based sensor model that is sufficiently smooth for occupancy grid mapping. We do so at the cost of computational complexity but arguably produces a more realistic sensor model for proximity sensing. Consider each measurement beam to have a non-zero angular width Θ, i.e., the beam is divergent instead of narrow. This divergence implies that a beam covers a non-zero area of the cells it passes through, meaning that the sensor model should take into account that increased cell coverage equates to better sensing of that cell. More specifically, let the jth measurement's prior for the divergent beam-based sensor

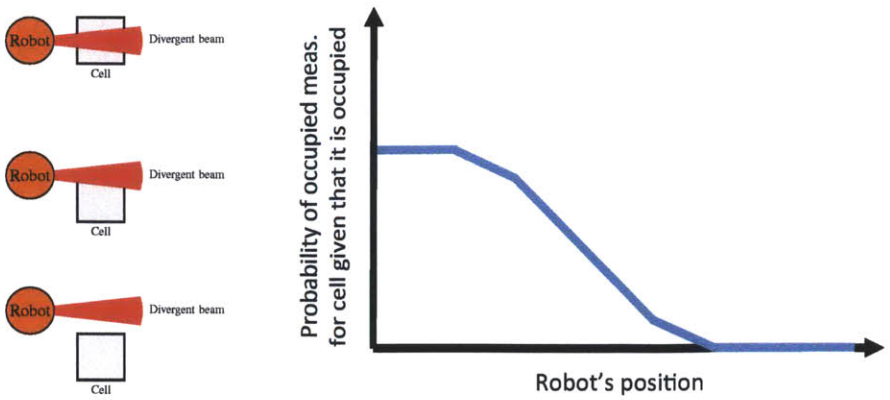


Figure 5-8: Smooth characteristics of the divergent beam-based sensor model. *Left:* As the robot moves in an orthogonal direction to a grid cell, one if its divergent beams gradually transitions from measuring within the cell to measuring completely outside the cell. *Right:* The measurement probability associated with this cell continuously transitions to is smallest value as a larger percentage of the beam leaves the cell. The resulting mutual information reward surface on which the robot navigates does not contain any discontinuities, and thus LaSalle's Invariance Principle can be used to prove convergence between observations for the corresponding controller.

model be

$$\mathbb{P}(Z_{k+1}^{[j]}|z_{1:k}) = \frac{1}{\Theta} \int_{\theta\in\left(-\frac{\Theta}{2},\frac{\Theta}{2}\right)} \left(\frac{1}{\prod_{\zeta\in\mathcal{I}_{k+1}^{[j,\theta]}} (r_k^{[\zeta]}+1)} \mathbb{P}(Z_{k+1}^{[j]} = z|M=\mathbf{0}) \right.$$
$$\left. + \sum_{\ell\in\mathcal{I}_{k+1}^{[j,\theta]}} \frac{r_k^{[\ell]}}{\prod_{\zeta\leq\ell}(r_k^{[\zeta]}+1)} \mathbb{P}(Z_{k+1}^{[j]} = z|M=\mathbf{e}_\ell) \right) d\theta, \tag{5.12}$$

where the measurement's perceptual range $\mathcal{I}_{k+1}^{[j,\theta]}$ now considers the angle within the beam's width. Given square occupancy grid cells and a non-zero minimum for the set of possible range measurements $\mathcal{Z}$, the divergent beam-based sensor model yielding (5.12) and employing the beam-based proximity mixture model from Section 4.4 is sufficiently smooth as defined in Assumption 43, e.g., see Figure **??**. The model also allows for straightforward modifications to Algorithm 1 and Algorithm 2, with the

latter employing the gradient of the measurement's prior

$$\frac{\partial \mathbb{P}(Z_{k+1}^{[j]}|z_{1:k})}{\partial c_{k+1}} = \frac{1}{\Theta} \int_{\theta \in (-\frac{\Theta}{2}, \frac{\Theta}{2})} \sum_{\ell \in \mathcal{I}_{k+1}^{[j,\theta]}} \frac{r_k^{[\ell]}}{\prod_{\zeta \leq \ell} (r_k^{[\zeta]} + 1)} \frac{\partial \mathbb{P}(Z_{k+1}^{[j]} = z | M = \mathbf{e}_\ell)}{\partial c_{k+1}} \, d\theta.$$

5.2.3 Sample sets for likely observations

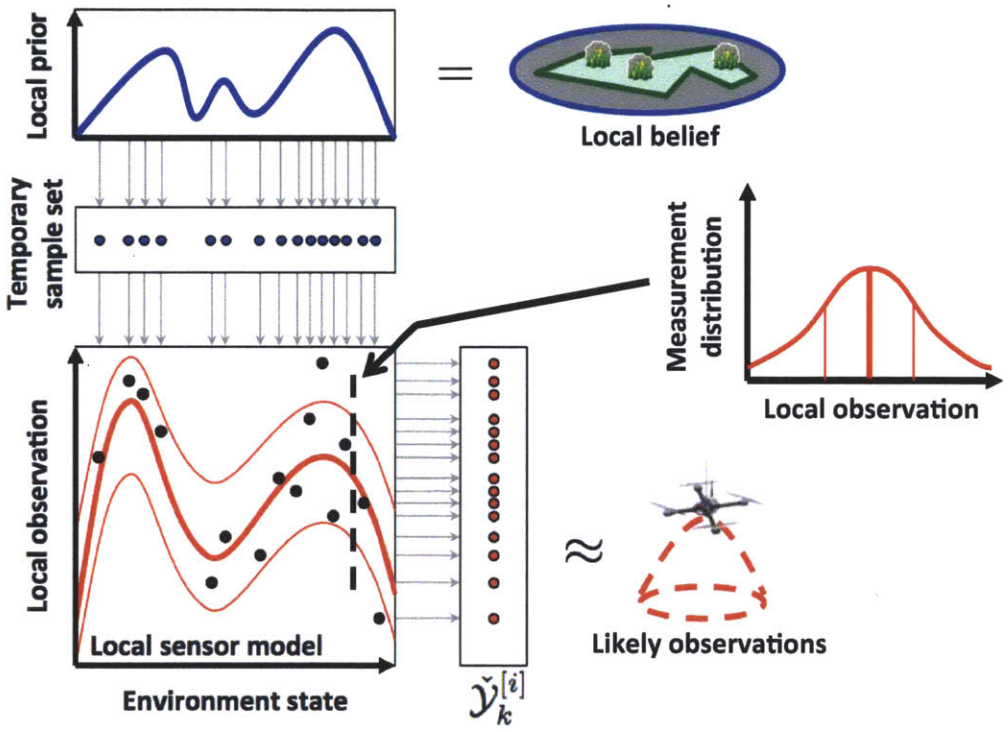


Figure 5-9: The sampling methodology for creating a robot's local observation sample set. Each robot draws samples from its local belief of the environment state to form a temporary unweighted sample set. Using this set, each robot draws from its measurement distribution samples representing likely local observations.

We now discuss the methodology for distributively creating a sample set of likely joint observations needed to approximate the reward function's gradient $\partial U_k / \partial c_i^{[i]}$. Suppose each robot maintains a local belief of the environment state via some decentralized Bayesian filter, e.g., see Section 5.3.2 for an example implementation. Using its local belief, let each robot create a temporary unweighted environment state sample set by drawing n_y samples from its prior distribution. Note that the drawn samples represent equally likely state instantiations as they are formed in a method analogous

to the *importance sampling* step for particle filters (Thrun et al., 2005). Let each robot then form a local observation sample set

$$\check{\mathcal{Y}}_k^{[i]} = \left\{ \check{y}_k^{[i,\ell]} : \ell \in \{1, \ldots, n_y\} \right\},$$

by drawing one observation sample for each entry in the temporary state sample set using its local sensor model. For all $y \in \check{\mathcal{Y}}_k^{[i]}$, the corresponding sampled measurement likelihoods become

$$\mathbb{P}(\check{Y}_k^{[i]} = y | X_k) = \frac{\mathbb{P}(Y_k^{[i]} = y | X_k)}{\sum_{\ell=1}^{n_y} \mathbb{P}(Y_k^{[i]} = \check{y}_k^{[i,\ell]} | X_k)},$$

where $\check{Y}_k^{[i]}$ is a random variable that takes values from $\check{\mathcal{Y}}_k^{[i]}$. This sampling methodology is illustrated in Figure 5-9.

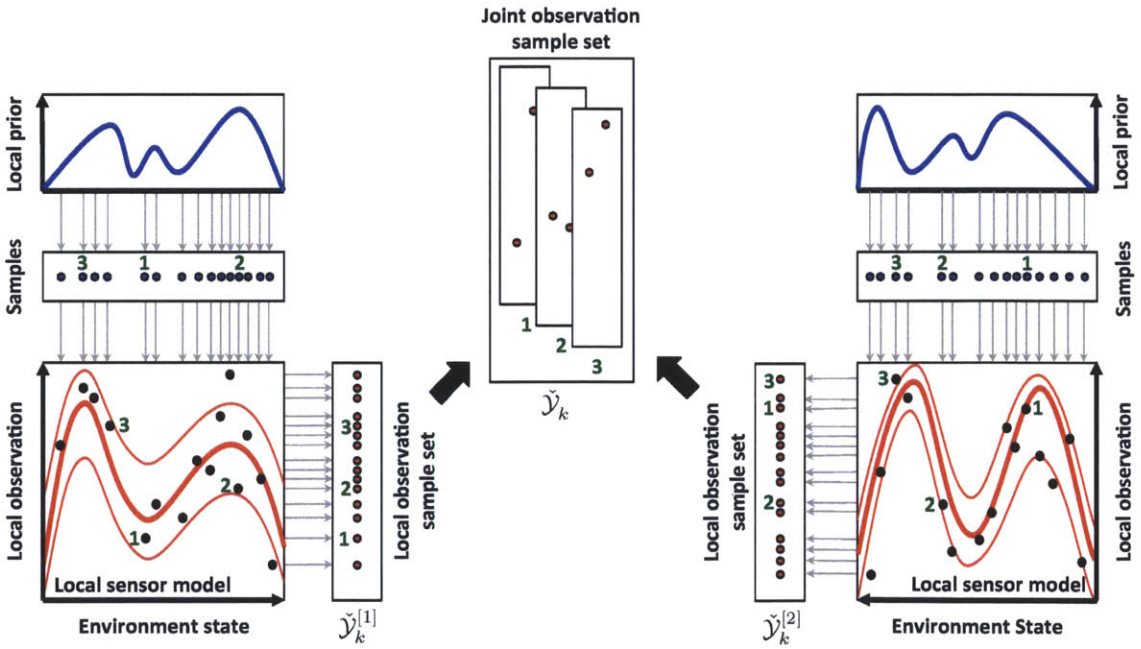


Figure 5-10: The sampling methodology for creating the joint observation sample set. A concatenation among all robots' local observation sample sets is formed over the network. This concatenation process can be interrupted for anytime performance purposes and still yield an joint observation sample set that is unbiased, as shown here with a set consisting of three samples.

We then define the joint observation sample set $\check{\mathcal{Y}}_k$ as the unweighted set of n_r-

tuples formed from the robots' observation samples having equal indices. More formally, we have that

$$\check{\mathcal{Y}}_k = \left\{ \check{y}_k^{[\ell]} = (\check{y}_k^{[1,\ell]}, \ldots, \check{y}_k^{[n_r,\ell]}) : \check{y}_k^{[i,\ell]} \in \check{\mathcal{Y}}_k^{[i]} \right\}.$$

Note that a naïve formulation of a joint observation sample set would be the Cartesian product of all the robots' observation sample sets $\prod_{i=1}^{n_r} \check{\mathcal{Y}}_k^{[i]}$, which scales exponentially in size with respect to the number of robots. Here we use the fact that the local observation sample sets are unweighted (i.e., all samples are equally likely) and conditionally independent to form an unbiased joint observation sample set of constant size with respect to the number of robots. In other words, each robot independently draws its own local observation samples using its local sensor model, and due to conditional independence, the concatenation of these samples across all robots is equivalent to a sample set formed by using the system's joint sensor model. This property also allows each robot to employ its own variant of the Bayesian filter, as long as the alphabet $\mathcal{X}$ of the environment state is known by the robot team. Such a situation is shown in Figure 5-10.

5.2.4 Decentralized control

We will show in Chapter 6 that by using a consensus-based algorithm, each robot can distributively approximate the sampled joint measurement likelihoods $\mathbb{P}(\check{Y}_k = y|X_k)$ for all $y \in \check{\mathcal{Y}}_k$, where $\check{Y}_k$ is a random variable that takes values from the joint observation sample set $\check{\mathcal{Y}}_k$. Let $\check{p}_k^{[i]} = (\check{p}_k^{[i,1]}, \ldots, \check{p}_k^{[i,n_y]})$ denote the n_y-tuple containing these approximations, which gives an approximation to the posterior calculation (5.7) of

$$\mathbb{P}(X_k = x|\check{Y}_k = \check{y}_k^{[\ell]}) \approx \frac{\mathbb{P}^{[i]}(X_k = x)\,\check{p}_k^{[i,\ell]}(x)}{\int_{x' \in \mathcal{X}} \mathbb{P}^{[i]}(X_k = x')\,\check{p}_k^{[i,\ell]}(x')dx'}$$

for all $x \in \mathcal{X}$ and $\ell \in \{1, \ldots, n_y\}$, where $\check{p}_k^{[i,\ell]}(x)$ is the value of $\check{p}_k^{[i,\ell]}$ evaluated at x.

By incorporating the joint observation sample set and their approximated mea-

surement likelihoods into (5.3), we define

$$\begin{aligned}\partial\check{U}_k^{[i]} := \sum_{\ell=1}^{n_y} \int_{x\in\mathcal{X}} & \frac{\partial \mathbb{P}(\check{Y}_k^{[i]} = \check{y}_k^{[i,\ell]} | X_k = x)}{\partial c_t^{[i]}} \\ & \times \frac{\mathbb{P}^{[i]}(X_k = x)\, \check{p}_k^{[i,\ell]}(x)}{\mathbb{P}(\check{Y}_k^{[i]} = \check{y}_k^{[i,\ell]} | X_k = x)\Big|_{t=(k-1)T_s}} \\ & \times \log\left(\frac{\check{p}_k^{[i,\ell]}(x)}{\int_{x'\in\mathcal{X}} \mathbb{P}^{[i]}(X_k = x')\, \check{p}_k^{[i,\ell]}(x')dx'}\right) dx \end{aligned} \tag{5.13}$$

to be the ith robot's approximation of the reward function's gradient given its local measurement likelihoods at time $t = kT_s$. Multiplying this result by the positive scalar control gain $\gamma^{[i]}$ results in a gradient-ascent controller that is distributed among the robots and uses non-parametric sample-based representations of the joint measurement probability distributions. Note that the incorporation of these distributed approximations does not preclude the use of LaSalle's Invariance Principle to prove convergence, leading to the following.

Corollary 47 (Convergence of decentralized controller) *Suppose Assumptions 39-43 hold. Let the robots move in the same configuration space and sense a bounded environment that is a subset of the configuration space. Then for a positive scalar $\gamma^{[i]}$, the controller*

$$u_t^{[i]} = \gamma^{[i]} \partial\check{U}_k^{[i]} \tag{5.14}$$

is convergent to zero between observations for all robots.

Definition 48 (The novel controller class) *The main contribution of this chapter is the derivation, analysis, and implementation of the class of decentralized mutual information-based gradient-ascent controllers of the form* (5.14)*. We will refer to this class as the novel controller class.*

Remark 49 (Distributed among robots) *Compared with its general form* (5.11)*, we note that the gradient-ascent controller* (5.14) *does not require that each robot has*

global knowledge of i) *a centralized prior* $\mathbb{P}(X_k)$*;* ii) *a joint position* c_t*; and* iii) *a joint sensor model* $\mathbb{P}(Y_k|X_k)$*. Thus, the controller is distributed among the robots.*

Remark 50 (Loss of local optimality) *Since the distributed controller* (5.14) *incorporates approximations for the joint observations and corresponding measurement likelihoods, an equilibrium system configuration,* $c_* = (c_*^{[1]}, \ldots, c_*^{[n_r]})$*, defined by*

$$\partial \check{U}_k^{[i]}\Big|_{c_t^{[i]}=c_*^{[i]}} = 0, \quad \forall i \in \{1, \ldots, n_r\}$$

is not guaranteed to be locally optimal solution to constrained optimization problem $\max_{c \in \mathcal{C}} U_k$*.*

Remark 51 (Computational tractability) *Again consider* X_k *to be a discrete-valued random variable that takes a value from an alphabet of size* $|\mathcal{X}|$*. An algorithm employing the reward function's gradient approximation* (5.13) *requires* $O(n_y|\mathcal{X}|)$ *time and* $O(n_y)$ *memory, where the memory requirement is due to precomputing the summation in the logarithm function for all joint observation samples. Hence, the members of the novel controller class scale linearly with respect to the sizes of the joint observation sample set. Moreover, their computational complexity remain constant with respect to the number of robots.*

Remark 52 (Details of the sensor model) *For the approximate gradient of the reward function to be well defined, we must have* $\mathbb{P}(\check{Y}_k^{[i]}|X_k) > 0$ *for all robot positions* $c_t^{[i]} \in \mathcal{C}^{[i]}$*. For this property, it is sufficient to have non-zero probabilities associated with all robots' observations. In addition, note that* $\mathbb{P}(\check{Y}_k^{[i]}|X_k)$ *in the denominator of* (5.13) *is evaluated at time* $t = (k-1)T_s$ *since the terms that make up* $\check{p}_k^{[i]}$ *are formed at the beginning of the time step (see Figure 5-5).*

5.3 Algorithmic implementations

In this section we give algorithmic implementations for the members of the novel controller class, then show their compatibility with any decentralized Bayesian filter

designed to infer the environment state.

5.3.1 Decentralized standard Bayesian estimation

We consider the two cases when the environment state X_k is a discrete-valued random variable and a continuous-valued random variable. Both algorithmic implementations assume a decentralized standard Bayesian estimation approach, i.e., when the posterior distribution calculation (5.7) is computed in full with respect to all (possibly discretized) environment state values. We leave the details of specialized Bayesian estimation approaches (e.g., Kalman filters) to the reader.

First consider the environment state X_k to be a discrete-valued random variable. The controller given in Algorithm 6 is a member of the novel controller class and is designed to distributively run on each robot to generate local velocity control commands.

Algorithm 6 DiscreteEnvironmentStateController($\check{p}_k^{[i]}$)

Require: Beliefs $\mathbb{P}^{[i]}(X_k = x)$ for all states $x \in \mathcal{X}$.

// initialize normalization tuple found in (5.13)
$\eta \leftarrow \emptyset$;
for $\ell = 1, \ldots, n_y$ **do**
  // augment normalization tuple
  $\eta \leftarrow (\eta, 0)$;
  for all $x \in \mathcal{X}$ **do**
    $\eta^{[\ell]} \leftarrow \eta^{[\ell]} + \mathbb{P}^{[i]}(X_k = x)\, \check{p}_k^{[i,\ell]}(x)$;
  end for
end for
// initialize reward function's gradient
$\partial \check{U} \leftarrow \mathbf{0}$;
for $\ell = 1, \ldots, n_y$ **do**
  for all $x \in \mathcal{X}$ **do**
    // sum elements of (5.13)
    $$\partial \check{U} \leftarrow \partial \check{U} + \frac{\partial \mathbb{P}(\check{Y}_k^{[i]} = \check{y}_k^{[i,\ell]} | X_k = x)}{\partial c_t^{[i]}} \frac{\mathbb{P}^{[i]}(X_k = x)\, \check{p}_k^{[i,\ell]}(x) \left(\log(\check{p}_k^{[i,\ell]}(x)) - \log(\eta^{[\ell]}) \right)}{\mathbb{P}(\check{Y}_k^{[i]} = \check{y}_k^{[i,\ell]} | X_k = x)\big|_{t=(k-1)T_s}};$$
  end for
end for
return $\gamma^{[i]}\, \partial \check{U}$;

Proposition 53 (Computational complexity for discrete case) *The time and space complexities of Algorithm 6 are $O(n_y|\mathcal{X}|)$ and $O(n_y)$, respectively. In words, the time (space) complexity of the controller is at worst linear (linear, respectively) in the size of the observation sample set and linear (independent, respectively) in the size of the environment state alphabet.*

Now consider the environment state X_k to be a continuous-valued random variable. To numerically calculate the integrals in (5.13), let $\widehat{\mathcal{X}}$ be a discretized set of environment state values representing the alphabet $\mathcal{X}$. In addition, let λ_x be the the numerical integration's resolution for this discretized set. The controller given in Algorithm 7 is again a member of the novel controller class and is designed to distributively run on each robot to generate local velocity control commands.

Algorithm 7 ContinuousEnvironmentStateController($\check{p}_k^{[i]}$)

Require: Beliefs $\mathbb{P}^{[i]}(X_k = x)$ for all states $x \in \mathcal{X}$.

// initialize normalization tuple found in (5.13)
$\eta \leftarrow \emptyset$;
for $\ell = 1, \ldots, n_y$ **do**
  // augment normalization tuple
  $\eta \leftarrow (\eta, 0)$;
  for all $x \in \widehat{\mathcal{X}}$ **do**
    $\eta^{[\ell]} \leftarrow \eta^{[\ell]} + \mathbb{P}^{[i]}(X_k = x)\, \check{p}_k^{[i,\ell]}(x)$;
  end for
  // include integration discretization measure
  $\eta^{[\ell]} \leftarrow \eta^{[\ell]}\, \lambda_x^{-1}$;
end for
// initialize reward function's gradient
$\widehat{\partial U} \leftarrow \mathbf{0}$;
for $\ell = 1, \ldots, n_y$ **do**
  for all $x \in \widehat{\mathcal{X}}$ **do**
    // sum elements of (5.13)
    $\widehat{\partial U} \leftarrow \widehat{\partial U} + \frac{\partial \mathbb{P}(\check{Y}_k^{[i]} = \check{y}_k^{[i,\ell]} | X_k = x)}{\partial c_t^{[i]}} \frac{\mathbb{P}^{[i]}(X_k = x)\, \check{p}_k^{[i,\ell]}(x) \left(\log(\check{p}_k^{[i,\ell]}(x)) - \log(\eta^{[\ell]})\right)}{\mathbb{P}(\check{Y}_k^{[i]} = \check{y}_k^{[i,\ell]} | X_k = x)}\Big|_{t=(k-1)T_s}$;
  end for
end for
return $\gamma^{[i]}\, \widehat{\partial U}\, \lambda_x^{-1}$;

Proposition 54 (Computational complexity for continuous case) *The time and space complexities of Algorithm 7 are $O(n_y \lambda_x)$ and $O(n_y)$, respectively. In words, the time (space) complexity of the controller is at worst linear (linear, respectively) in the size of the observation sample set and linear (independent, respectively) in the size of the numerical integration's resolution of the environment state alphabet.*

5.3.2 Non-parametric decentralized Bayesian estimation

Not all applications allow the posterior distribution calculation (5.7) to be computed in full. For example, consider the discrete-valued case when the alphabet size $|\mathcal{X}|$ makes the computational complexity of Algorithm 6 intractable, such as an occupancy grid mapping problem with *dependence* between cells. For these applications, principled approximations are needed to represent the robots' beliefs of the environment state. In the spirit of how likely observations are represented in Section 5.2.3, we now give an approach that uses a non-parametric sample-based method, and by doing so show that this and any other decentralized Bayesian filter is compatible with the members of the novel controller class.

Let each robot maintain a weighted environment state sample set

$$\check{\mathcal{X}}_k^{[i]} = \left\{ (\check{x}_k^{[i,j]}, \check{w}_k^{[i,j]}) : j \in \{1, \ldots, n_x\} \right\},$$

of size n_x, where each sample, $\check{x}_k^{[i,j]} \in \mathcal{X}$, has a corresponding weight[2], $\check{w}_k^{[i,j]} \in (0, 1)$. Each sample is a candidate instantiation of the environment state, and the pairing of the samples and their corresponding weights represents a non-parametric representation of the robot's belief of the environment state.

Once a joint observation $y_k \in \mathcal{Y}$ is received, an approximation for the joint measurement likelihood $\mathbb{P}(Y_k = y_k | X_k)$ needs to be distributively calculated. We will discuss consensus-based algorithms to do so in Chapter 6, but for now let the $p_k^{[i]}$

[2]We have for all robots that $\sum_{j \in \{1,\ldots,n_x\}} \check{w}_k^{[i,j]} = 1$.

Algorithm 8 SequentialImportanceResampling($\check{\mathcal{X}}_k^{[i]}$)

$\check{\mathcal{X}}_{k+1}^{[i]} \leftarrow \emptyset$.
for $j = 1$ to n_x **do**
  Sample $\check{x}_{k+1}^{[i,j]} \sim \mathbb{P}^{[i]}(X_{k+1}|X_k = \check{x}_k^{[i,j]})$.
  $\check{w}_{k+1}^{[i,j]} \leftarrow \frac{\check{w}_k^{[i,j]} p_k^{[i]}(\check{x}_{k+1}^{[i,j]})}{\sum_{j'=1}^{n_x} \check{w}_k^{[i,j']} p_k^{[i]}(\check{x}_{k+1}^{[i,j']})}$.
  $\check{\mathcal{X}}_{k+1}^{[i]} \leftarrow \check{\mathcal{X}}_{k+1}^{[i]} \cup \{(\check{x}_{k+1}^{[i,j]}, \check{w}_{k+1}^{[i,j]})\}$.
end for
Apply appropriate resampling technique.
return $\check{\mathcal{X}}_{k+1}^{[i]}$.

Algorithm 9 NonparametricEnvironmentStateController($\check{p}_k^{[i]}$)

Require: Non-parametric belief $\check{\mathcal{X}}_k^{[i]}$.
// initialize normalization tuple found in (5.13)
$\eta \leftarrow \emptyset$;
for $\ell = 1, \ldots, n_y$ **do**
  // augment normalization tuple
  $\eta \leftarrow (\eta, 0)$;
  for $j = 1, \ldots, n_x$ **do**
    $\eta^{[\ell]} \leftarrow \eta^{[\ell]} + \check{w}^{[i,j]} \check{p}_k^{[i,\ell]}(x)$;
  end for
end for
// initialize reward function's gradient
$\partial \check{U} \leftarrow \mathbf{0}$;
for $\ell = 1, \ldots, n_y$ **do**
  for $j = 1, \ldots, n_x$ **do**
    // sum elements of (5.13)
    $\partial \check{U} \leftarrow \partial \check{U} + \frac{\partial \mathbb{P}(\check{Y}_k^{[i]} = \check{y}_k^{[i,\ell]} | X_k = x)}{\partial c_t^{[i]}} \frac{\check{w}^{[i,j]} \check{p}_k^{[i,\ell]}(x) \left(\log(\check{p}_k^{[i,\ell]}(x)) - \log(\eta^{[\ell]})\right)}{\mathbb{P}(\check{Y}_k^{[i]} = \check{y}_k^{[i,\ell]} | X_k = x)\big|_{t=(k-1)T_s}}$;
  end for
end for
return $\gamma^{[i]} \partial \check{U}$;

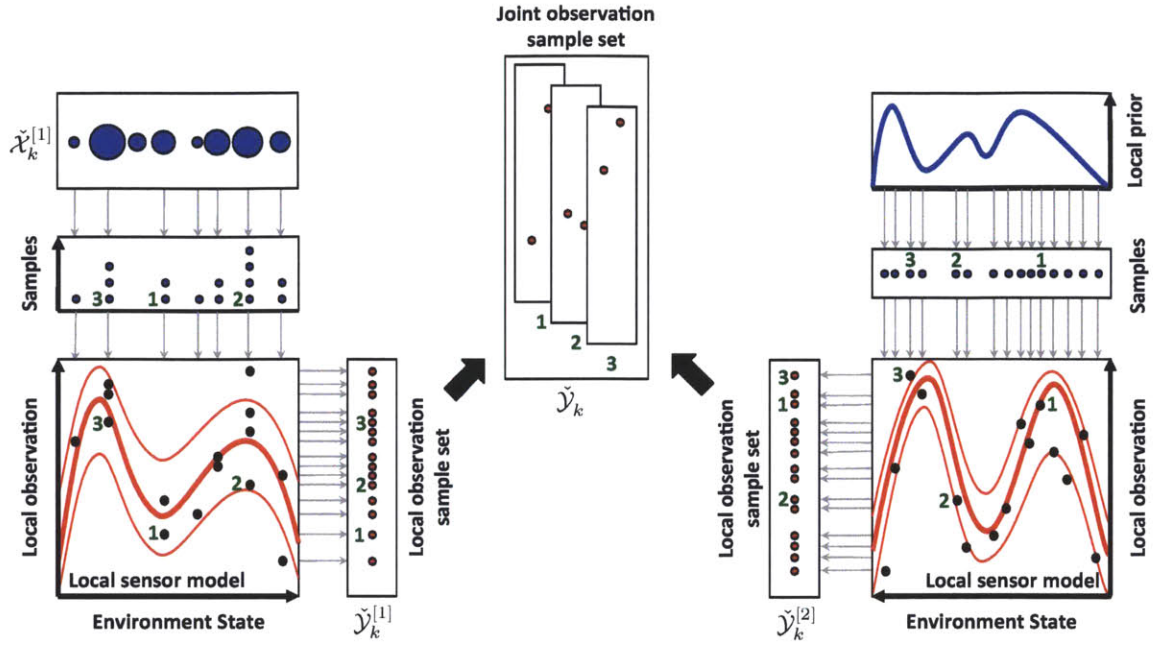


Figure 5-11: The joint observation sampling methodology using various decentralized Bayesian filters. The first robot employs a non-parametric Bayesian filter for inferring the environment state, while the second employs a standard Bayesian filter. Besides local variations in the controller implementation (Algorithm 9 versus Algorithm 7), the overall employment of the controllers remains identical among the robots.

denote this approximation. The posterior calculation from (5.7) becomes

$$\mathbb{P}(X_k = \check{x}_k^{[i,j]} | Y_k = y_k) \approx \frac{\check{w}_k^{[i,j]} p_k^{[i]}(\check{x}^{[i,j]})}{\sum_{j'=1}^{n_x} \check{w}_k^{[i,j']} p_k^{[i]}(\check{x}^{[i,j']})} \tag{5.15}$$

for all $j \in \{1, \ldots, n_x\}$, where $p_k^{[i]}(x)$ is the value of $p_k^{[i]}$ evaluated at x. Thus, each robot forms its weighted environment state sample set for the upcoming time step $k+1$ by drawing from its state transition distribution $\mathbb{P}^{[i]}(X_{k+1}|X_k)$, calculating the corresponding weights from (5.15), and applying an appropriate resampling technique. The process given in Algorithm 8 is the well-known sequential Monte Carlo method called sequential importance resampling (Thrun et al., 2005).

Using this non-parametric representation of the robot's belief, the controller given in Algorithm 9 is, once again, a member of the novel controller class and is designed to distributively run on each robot to generate local velocity control commands.

Proposition 55 (Computational complexity for non-parametric case) *The time and space complexities of Algorithm 9 are $O(n_x n_y)$ and $O(n_y)$, respectively. In words, the time (space) complexity of the controller is at worst linear (linear, respectively) in the size of the observation sample set and linear (independent, respectively) in the size of the weighted environment state sample set.*

Remark 56 (Compatibility of the novel controller class) *The members of the novel controller class are compatible with any type of decentralized Bayesian filter. In other words, as long as the local controller has access to a probabilistic representation of the environment state, it can perform mutual information-based gradient-ascent control. More over, different robots can employ different types of decentralized Bayesian filters within the same system, as illustrated in Figure 5-11.*

5.4 Experiments in multi-robot systems

Figure 5-12: A snapshot of the autonomous deployment of five quadrotor flying robots for an outdoor hardware experiment. Even though the robots were fully autonomous, a safety pilot was assigned to each robot to allow for manual overrides in the event of an emergency.

We first describe the technical details of our hardware system employing five quadrotor flying robots, then verified three members of the novel controller class by presenting results from two indoor experiments and one outdoor experiment.

5.4.1 Quadrotor flying robot system

We are working towards a multi-robot system that can rapidly assess the state of disaster-affected environments. In these cases the state can represent a wide spectrum of relevant information, ranging from the presence of fires and harmful substances to the structural integrity of buildings. Motivated by this goal, the task for the hardware experiments was to infer the state of a bounded, planar environment by deploying five Ascending Technologies Hummingbird quadrotor flying robots (Gurdan et al., 2007). Five heterogeneous sensors were simulated with measurement noise that was proportional to the field of view, meaning that sensors of larger footprints produced noisier observations. Lastly, an ideal disk network model was enforced on the system to limit the peer-to-peer communication range.

For each environment, we defined $\mathcal{W}$ to be an n_w cell partition[3], where for each cell $\mathcal{W}^{[m]}$, the state was modeled as a binary random variable $X_k^{[m]}$ that took values from the alphabet $\mathcal{X}^{[m]} = \{0, 1\}$. Thus for the environment state, we had an n_w-tuple random variable $X_k = (X_k^{[1]}, \ldots, X_k^{[n_w]})$ that took values from the alphabet $\mathcal{X} = \prod_{m=1}^{n_w} \mathcal{X}^{[m]}$. A first order Markov model was used for the state transition distributions, where a uniform probability represented the likelihood that the state of an environment discretization cell transitioned to any other state. We modeled the robot's observation as an n_w-tuple of binary random variables $Y_k^{[i]} = (Y_k^{[i,1]}, \ldots, Y_k^{[i,n_w]})$ that took values from the alphabet $\mathcal{Y}^{[i]} = \prod_{m=1}^{n_w} \mathcal{Y}^{[i,m]}$ with $\mathcal{Y}^{[i,m]} = \{0, 1\}$.

For simplicity, we assumed conditional independence between environment discretization cells for the measurement probabilities. This assumption resulted in the cell-wise sensor model $\mathbb{P}(Y_k^{[i,m]} | X_k)$ being dependent on the state of the correspond-

[3]The partition $\mathcal{W}$ is defined as a collection of closed connected subsets of $\mathcal{Q}$ satisfying $\prod_{m=1}^{n_w} \mathcal{W}^{[m]} = \mathcal{W}$, $\bigcup_{m=1}^{n_w} \mathcal{W}^{[m]} = \mathcal{Q}$ and $\bigcap_{m=1}^{n_w} \text{int}(\mathcal{W}^{[m]}) = \emptyset$, where $\text{int}(\cdot)$ denotes the subset of interior points.

ing environment discretization cell $X_k^{[m]}$ and conditionally independent from all other $Y_k^{[i,m']}$ with $m' \neq m$. More formally, we have that

$$\mathbb{P}(Y_k^{[i]}|X_k) = \prod_{m=1}^{n_w} \mathbb{P}(Y_k^{[i,m]}|X_k^{[m]}). \tag{5.16}$$

In words, the robot's observation was composed of n_w conditionally independent observation elements, where each element concerned a specific environment discretization cell. For all experiments, the robots had maximum cell-wise measurement probabilities $\mathbb{P}(Y_k^{[i,m]} = 0|X_k^{[m]} = 0)$ and $\mathbb{P}(Y_k^{[i,m]} = 1|X_k^{[m]} = 1)$ of $\{0.95, 0.9, 0.85, 0.8, 0.75\}$, which decreased quadratically (e.g., power decay of light) to 0.5 at the edge of the robot's field of view.

Parameter	Symbol	Value
Environment state alphabet	$\mathcal{X}^{[m]}$	$\{0, 1\}$
Local observation alphabet	$\mathcal{Y}^{[i,m]}$	$\{0, 1\}$
Min cell-wise measurement prob.	$\min \mathbb{P}(Y_k^{[i,m]} = 0\|X_k^{[m]} = 0)$ $\min \mathbb{P}(Y_k^{[i,m]} = 1\|X_k^{[m]} = 1)$	0.5
Max cell-wise measurement prob.	$\max \mathbb{P}(Y_k^{[i,m]} = 0\|X_k^{[m]} = 0)$ $\max \mathbb{P}(Y_k^{[i,m]} = 1\|X_k^{[m]} = 1)$	$\{0.95, 0.9, 0.85, 0.8, 0.75\}$
Observation sample set size	n_y	500

Table 5.1: Common parameters for the hardware experiments in Section 5.4.

Remark 57 (Complexity of environment state) *We do not assume independence between environment discretization cells, resulting in a state alphabet size that scales exponentially with respect to the number of cells, i.e.,* $\max_m O(|\mathcal{X}^{[m]}|^{n_w})$. *One can assume independence to have this size scale linearly with respect to the number of cells, i.e.,* $\max_m O(n_w|\mathcal{X}^{[m]}|)$, *which is a common assumption in the robot mapping literature, e.g., occupancy grid mapping in Chapter 4.*

Remark 58 (Using simulated sensors) *We note that the experiments are a validation of the novel controller class and not of the sensing capabilities of our hardware*

platforms. The selection of the simulated sensor properties was primarily motivated by their generality, in particular for systems employing passive sensors that measure the intensity of electromagnetic signals radiating from point sources. Nonetheless, we expect qualitatively good controller performance for systems that obtain noisier observations towards the boundaries of their sensors' limited range, even though convergence guarantees may not be provable for such systems.

5.4.2 Indoor experiment using a decentralized discrete Bayesian filter

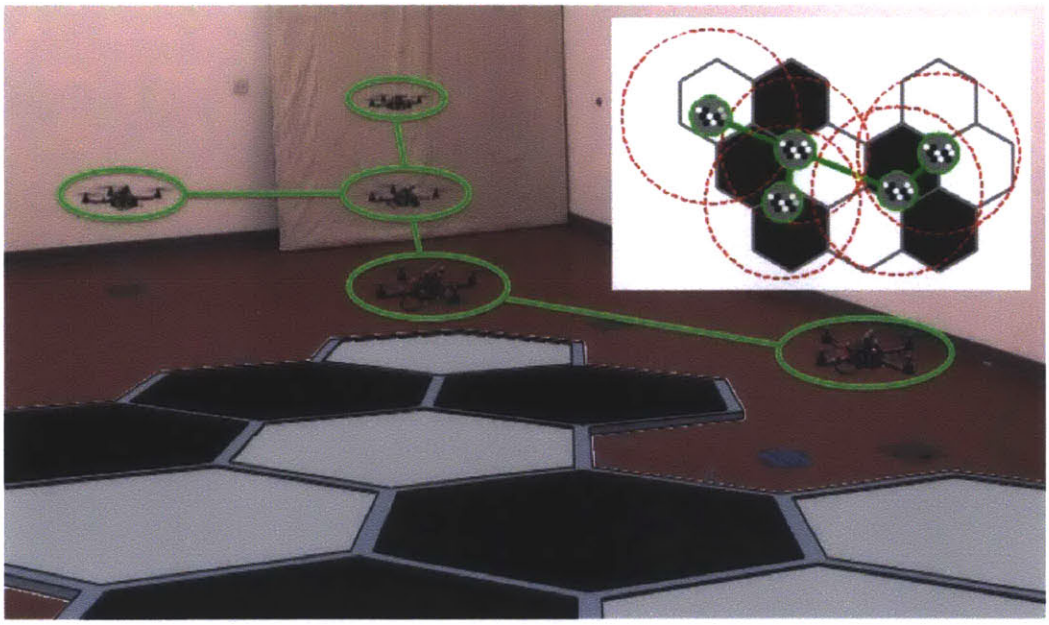


Figure 5-13: Five quadrotor flying robots inferring the state of an indoor environment using a decentralized discrete Bayesian filter. The hexagon cells overlaying the snapshot illustrate the state and location of the discretized environment. The green lines between robots represent network connectivity. *Top right:* This schematic illustrates the state of each robot's inference enabled by simulated sensors with their footprints drawn in red dashed circles.

For the first indoor experiment, the 10 m long environment (see Figure 5-13) was discretized into $n_w = 10$ hexagon cells, each being of inner radius 2 m and having a binary static state of either 0 (e.g., no fire) or 1 (e.g., fire). In addition, the random variables representing the robots' local observation elements $Y_k^{[i,m]}$ also took values of

0 (e.g., no heat observed) or 1 (e.g., heat observed). The experiment was conducted in an MIT CSAIL laboratory equipped with a Vicon motion capture system. The realtime software for each robot ran in distributed fashion on a single computer. This software included a low level linear-quadratic regulator position controller that accepted waypoint inputs from Algorithm 6 and sent low level control commands to the robots via 2.4 Ghz Xbee-Pro wireless modules. The five heterogeneous sensors were simulated with maximum sensing radii of $\{2.0, 2.1, 2.2, 2.3, 2.4\}$ m. We represented these sensing properties by setting the hovering heights proportional to the sensing radii. In other words, robots hovering closer to the environment had more accurate observations, but also had smaller fields of view.

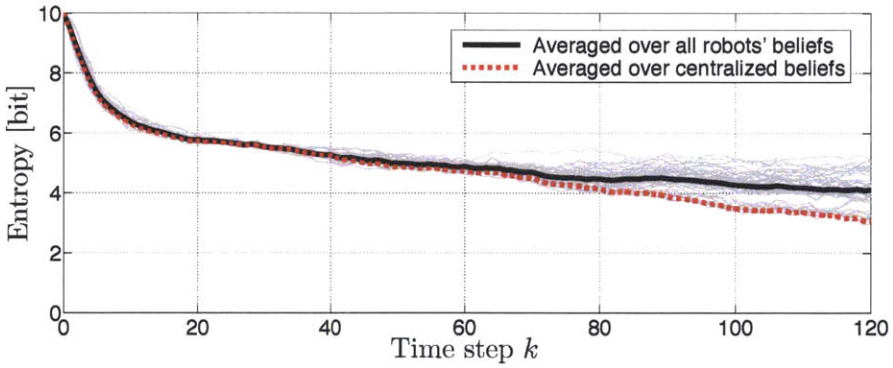


Figure 5-14: This plot shows the decrease in entropy of a five robot experiment using a decentralized discrete Bayesian filter. The entropy of the robots' beliefs and centralized belief are averaged over 10 consecutive runs. In addition, the light gray lines show the entropy of each robot's belief for every run.

For all robots, we used a control policy set of $\mathcal{U}_i = [-0.1, 0.1]^2$ m/s, a control gain of $\gamma_i = 10$, and a ideal disk network radius of 3 m radius. In addition, a safety radius of 1 m was enforced between neighboring robots, meaning the gradient projection of $u_t^{[i]}$ would be taken to prevent two directly communicating robots from moving closer than 1 m from each other. An observation sample set size of $n_y = 500$ was used, which allowed for a sampling interval of $T_s \approx 2$ s.

We recorded 10 consecutive runs deploying the five robots from the bottom of the environment, including one robot that started on the environment boundary and another outside. The plot in Figure 5-14 shows a decrease in average entropy (i.e.

Parameter	Symbol	Value
Control gain	$\gamma^{[i]}$	10
Control policy	$\mathcal{U}^{[i]}$	$[-0.1, 0.1]^2$ m/s
Ideal disk network radius	–	3.0 m
Number of environment discretization cells	n_w	10
Safety radius	–	1 m
Sample period	T_s	≈ 2 s
Sensor radii	–	$\{2.0, 2.1, 2.2, 2.3, 2.4\}$ m
State transition distribution	$\mathbb{P}^{[i]}(X_{k+1}^{[m]} = x \mid X_k^{[m]} = x)$	0.99, uniform otherwise

Table 5.2: Parameters for the indoor experiment in Sections 5.4.2 and 5.4.3.

uncertainty) of the robots' beliefs compared to a centralized one. The centralized inference considered observations from all robots, and can be interpreted as a baseline. On average the entropy of the robots' beliefs were within 1 bit of the centralized one over the 240 s. To date, we have over 100 successful runs with various starting positions and algorithm parameters, compared to one unsuccessful run caused by the motion capture system losing track of one robot. Even during this run, the decentralized controller continued to run properly for the other robots, showing the approach's robustness to individual robot failures.

5.4.3 Indoor experiment using a decentralized non-parametric Bayesian filter

For the second indoor experiment, most experimental details were identical to the first experiment except that a decentralized non-parametric Bayesian filter (Algorithm 8) and controller (Algorithm 9) ran onboard the robots using a weighted environment state sample set of size $n_x = 500$. Each onboard 2 GHz single board computer hosted its own independent Robot Operating System (ROS) environment (Quigley et al., 2009) and wirelessly communicated with other robots via UDP multicast over 802.11. Thus, the non-parametric version of our decentralized controller (Algorithm 9) ran

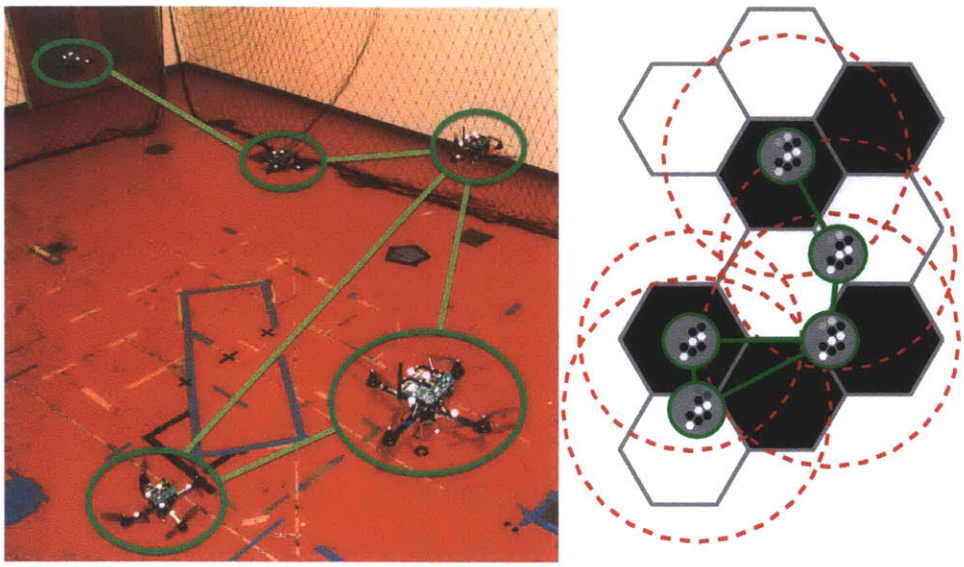


Figure 5-15: *Left:* Five quadrotor flying robots inferring the state of an indoor environment using a decentralized non-parametric Bayesian filter. The green lines between robots represent network connectivity. *Right:* This schematic illustrates the state of each robot's inference enabled by simulated sensors with their footprints drawn in red dashed circles.

in purely distributed fashion, for which we used the same parameters as in the first indoor experiment found in Section 5.4.2.

We recorded 10 consecutive runs deploying the five robots from the same starting location as in the first indoor experiment. Figure 5-16 shows the beginning, middle, and end configuration of a typical run, along with a plot showing the decrease in average entropy (i.e., uncertainty) of the robots' beliefs compared to a centrally computed one. Again, the centralized inferences considered observations from all robots, and can be interpreted as a baseline. Overall, the behavior of the non-parametric implementation compared similarly to the standard one, suggesting that enough computational resources were allocated to effectively approximate the probability distributions inherent to the information acquisition task.

Parameter	Symbol	Value
Control gain	$\gamma^{[i]}$	5
Control policy set	$\mathcal{U}^{[i]}$	$[-3,3]^2$ m/s
Environment state sample set size	n_x	500
Ideal disk network radius	–	50 m
Number of environment discretization cells	n_w	58
Safety radius	–	10 m
Sample period	T_s	≈ 1 s
Sensor radii	–	$\{30, 32.5, 35, 37.5, 40\}$ m
State transition distribution	$\mathbb{P}^{[i]}(X_{k+1}^{[m]} = x \mid X_k^{[m]} = x)$	0.95, uniform otherwise

Table 5.3: Parameters for the outdoor experiment in Section 5.4.4.

5.4.4 Outdoor experiment using a decentralized non-parametric Bayesian filter

For the outdoor experiment, a 150 m wide environment (see Figure 5-17) was discretized using a Voronoi partitioner into $n_w = 58$ heterogeneous cells. The decentralized non-parametric Bayesian filter (Algorithm 8) and controller (Algorithm 9) ran in distributed fashion for all robots at 1 Hz on a single ground workstation. The resulting GPS-based control commands were wirelessly transmitted via the Xbee-Pro modules to each robot's onboard autopilot. The five heterogeneous sensors were simulated with measurement noise proportional to the sensor radii of $\{30, 32.5, 35, 37.5, 40\}$ m, again meaning that sensors of larger footprints produced noisier observations. Each robot used a control policy set of $\mathcal{U}^{[i]} = [-3,3]^2$ m/s, a control gain of $\gamma^{[i]} = 5$, and an ideal disk network radius of 50 m.

In preparation for the outdoor experiment, reproducible results were recorded from multiple preliminary deployments, producing over 25 minutes of total flight time, e.g., see Figure 5-18. This initial effort was to verify the non-parametric implementation without any higher level control except for the manual override capabilities enabled by the Disaster Management Tool (DMT) developed at the German Aerospace Center

DLR (Frassl et al., 2010). Once we obtained qualitative validation for our approach, the algorithms were adjusted to handle binary event detection (e.g., fire or no fire) as described for the indoor experiment. In addition, a decentralized communication scheme continuously assigned robots to act as dynamic network relays, overriding the control actions produced by the distributed controller. For the experiment, the robots were deployed from outside the environment, and at any given point could have at most 58 bits of uncertainty concerning the environment state. The plot in Figure 5-17 shows the decrease in entropy over the extent of the experiment, even though the higher level control scheme at times was overriding the decentralized mutual information-based gradient-ascent controller.

5.5 Summary

Concerning the decentralized controllers presented in this chapter, scalability with respect to the number of robots comes from the algorithmic decoupling of the controller from the joint measurement likelihood calculations. By implementing distributed algorithms to approximate these likelihoods, we can induce scalability for the entire control approach. We next present in Chapter 6 novel consensus-based algorithms for approximating the robots' joint measurement likelihoods in both a discrete and continuous setting.

5.6 Proofs

This section contains proofs for all theorems, corollaries, and propositions found in this chapter.

Proof 59 (Theorem 38) *This proof was derived in collaboration with Schwager et al. (2011a), and a similar result in the context of channel coding was proved by Palomar and Verdú (2007). Concerning the partial derivative of* (5.2) *with respect to a robot's position* $c_t^{[i]}$*, we can move the differentiation inside the integrals since they do not depend on* $c_t^{[i]}$*. Applying the chain rule to the integrand and separating the two*

resulting terms, we have

$$\begin{aligned}\frac{\partial U_k}{\partial c_t^{[i]}} &= \int_{y\in\mathcal{Y}}\int_{x\in\mathcal{X}} \frac{\partial \mathbb{P}(X_k = x|Y_k = y)}{\partial c_t^{[i]}}\mathbb{P}(Y_k = y)\,dx\,dy \\ &+ \int_{y\in\mathcal{Y}}\int_{x\in\mathcal{X}} \frac{\partial \mathbb{P}(Y_k = y|X_k = x)}{\partial c_t^{[i]}}\mathbb{P}(X_k = x) \\ &\times \log\left(\frac{\mathbb{P}(X_k = x|Y_k = y)}{\mathbb{P}(X_k = x)}\right) dx\,dy, \end{aligned} \tag{5.17}$$

where

$$\mathbb{P}(Y_k) = \int_{x\in\mathcal{X}} \mathbb{P}(Y_k|X_k = x)\mathbb{P}(X_k = x)dx$$

from the law of total probability. We will now show that the first integration term on the right hand side of (5.17) *is equal to zero. First using the chain rule to take the partial derivative of* (5.1) *with respect to* $c_t^{[i]}$*, we have*

$$\frac{\partial \mathbb{P}(X_k|Y_k)}{\partial c_t^{[i]}} = \frac{\partial \mathbb{P}(Y_k|X_k)}{\partial c_t^{[i]}}\frac{\mathbb{P}(X_k)}{\mathbb{P}(Y_k)} - \frac{\partial \mathbb{P}(Y_k = y)}{\partial c_t^{[i]}}\frac{\mathbb{P}(Y_k|X_k)\mathbb{P}(X_k)}{\mathbb{P}(Y_k)^2}. \tag{5.18}$$

Substituting (5.18) *back into the first integration term on the right hand side of* (5.17) *and considering the two resulting integrals separately, we have*

$$\begin{aligned}&\int_{y\in\mathcal{Y}}\int_{x\in\mathcal{X}} \frac{\partial \mathbb{P}(Y_k = y|X_k = x)}{\partial c_t^{[i]}}\mathbb{P}(X_k = x)\,dx\,dy \\ &= \frac{\partial}{\partial c_t^{[i]}}\int_{y\in\mathcal{Y}}\int_{x\in\mathcal{X}} \mathbb{P}(Y_k = y|X_k = x)\mathbb{P}(X_k = x)\,dx\,dy \\ &= \frac{\partial}{\partial c_t^{[i]}}1 \\ &= 0\end{aligned}$$

and

$$\begin{aligned}
&\int_{y\in\mathcal{Y}}\int_{x\in\mathcal{X}} \frac{\partial \mathbb{P}(Y_k=y)}{\partial c_t^{[i]}} \frac{\mathbb{P}(Y_k=y|X_k=x)\mathbb{P}(X_k=x)}{\mathbb{P}(Y_k=y)}\,dx\,dy \\
&\quad= \int_{y\in\mathcal{Y}} \frac{\partial \mathbb{P}(Y_k=y)}{\partial c_t^{[i]}} \int_{x\in\mathcal{X}} \frac{\mathbb{P}(Y_k=y|X_k=x)\mathbb{P}(X_k=x)}{\mathbb{P}(Y_k=y)}\,dx\,dy \\
&\quad= \int_{y\in\mathcal{Y}} \frac{\partial \mathbb{P}(Y_k=y)}{\partial c_t^{[i]}}\,dy \\
&\quad= \frac{\partial}{\partial c_t^{[i]}} \int_{y\in\mathcal{Y}} \mathbb{P}(Y_k=y)\,dy \\
&\quad= \frac{\partial}{\partial c_t^{[i]}} 1 \\
&\quad= 0.
\end{aligned}$$

Proof 60 (Theorem 44) *Let*

$$V_k = -U_k = -\int_{y\in\mathcal{Y}}\int_{x\in\mathcal{X}} \mathbb{P}(Y_k=y|X_k=x)\mathbb{P}(X_k=x)\log\left(\frac{\mathbb{P}(X_k=x|Y_k=y)}{\mathbb{P}(X_k=x)}\right)dx\,dy$$

be a Lypanov-type function candidate whose partial derivative with respect to $c_t^{[i]}$ is the negative of $\partial U_k/\partial c_t^{[i]}$ from (5.3)*. Firstly, the closed loop dynamics*

$$\frac{dc_t^{[i]}}{dt} = -\gamma^{[i]}\frac{\partial V_k}{\partial c_t^{[i]}}$$

are autonomous. Assumption 43 implies that these dynamics are continuous on the robot's configuration space $\mathcal{C}^{[i]}$, as well as that V_k is differentiable on these spaces. Therefore, we have that V_k is continuously differentiable on the joint configuration space $\mathcal{C}$.

Taking the Lie derivative of V_k along the trajectories of the system, we have

$$\mathcal{L}_t V_k = \sum_{i=1}^{n_r} \frac{\partial V_k}{\partial c_t^{[i]}}\frac{dc_t^{[i]}}{dt} = -\sum_{i=1}^{n_r}\gamma^{[i]}\left(\frac{\partial V_k}{\partial c_t^{[i]}}\right)^2 \leq 0.$$

Now consider robots "far" enough away from the environment $\mathcal{Q}$ such that for all robots and $q \in \mathcal{Q}$, we have $\text{dist}(c_t^{[i]}, q) \geq d^{[i]}$. *Assumption 42 implies for all robots* $i \in \{1, \ldots, n_r\}$ *that*

$$\frac{\partial \mathbb{P}(Y_k^{[i]}|X_k)}{\partial c_t^{[i]}} = 0,$$

and thus we have that $dc_t^{[i]}/dt = 0$ *for all time. Hence, all evolutions of the system are bounded.*

Finally, consider the set of all $c_* = (c_*^{[1]}, \ldots, c_*^{[n_r]}) \in \mathcal{C}$ *such that for all robots* $i \in \{1, \ldots, n_r\}$ *we have that*

$$\left.\frac{\partial V_k}{\partial c_t^{[i]}}\right|_{c_t^{[i]}=c_*^{[i]}} = 0.$$

This set is invariant since it implies $\partial U_k/\partial c_t^{[i]} = 0$ *for all* $i \in \{1, \ldots, n_r\}$. *Thus, all conditions of LaSalle's Invariance Principle are satisfied and the trajectories will converge to this invariant set (LaSalle, 1960; Bullo et al., 2009).*

In addition, we have that c_* *is either a local minimum, maximum, or saddle point for the constrained optimization problem* $\max_{c \in \mathcal{C}} U_k$. *However, for a gradient system of this type, we know that this configuration is a Lyapunov stable equilibrium if and only if it is a local maximum, and thus locally optimal (Hirsche and Smale, 1974).*

Proof 61 (Corollary 47) *The proof directly follows the convergence proof for Theorem 44, using the Lypanov-type function candidate*

$$V_k = -\sum_{i=1}^{n_r} \check{U}_k^{[i]},$$

where

$$\check{U}_k^{[i]} := \sum_{\ell=1}^{n_y} \int_{x\in\mathcal{X}} \mathbb{P}(\check{Y}_k^{[i]} = \check{y}_k^{[i,\ell]} | X_k = x) \times \frac{\mathbb{P}^{[i]}(X_k = x)\, \check{p}_k^{[i,\ell]}(x)}{\mathbb{P}(\check{Y}_k^{[i]} = \check{y}_k^{[i,\ell]} | X_k = x)\Big|_{t=(k-1)T_s}} \times \log\left(\frac{\check{p}_k^{[i,\ell]}(x)}{\int_{x'\in\mathcal{X}} \mathbb{P}^{[i]}(X_k = x')\, \check{p}_k^{[i,\ell]}(x')dx'}\right) dx$$

can be thought of as an the ith robot's approximation of the reward function U_k given its local measurement likelihoods at time $t = kT_s$.

Proof 62 (Proposition 53) *By definition, the observation sample set $\check{\mathcal{Y}}_k^{[i]}$ has n_y elements for any robot i, while the environment state alphabet has $|\mathcal{X}|$ elements. Thus, the normalization tuple η has n_y scalar elements and takes $O(n_y|\mathcal{X}|)$ time to construct. More specifically for the latter, each element of η is calculated using a expression of $O(1)$ time complexity within a* for *loop of size $|\mathcal{X}|$.*

To finish the proof, we note that the expression on line 11 *has a time complexity of $O(1)$. Since this expression is within two* for *loops of sizes n_y and $|\mathcal{X}|$, the time complexity for Algorithm 6 is $O(n_y|\mathcal{X}|)$.*

Proof 63 (Proposition 54) *The proof directly follows the proof for Proposition 53, except that the normalization tuple η takes $O(n_y\lambda_x)$ time to construct and the* for *loop on line* 11 *in Algorithm 7 is of size $O(\lambda_x)$.*

Proof 64 (Proposition 55) *The proof directly follows the proof for Proposition 53, except that the normalization tuple η takes $O(n_x n_y)$ time to construct and the* for *loop on line* 11 *in Algorithm 9 is of size n_x.*

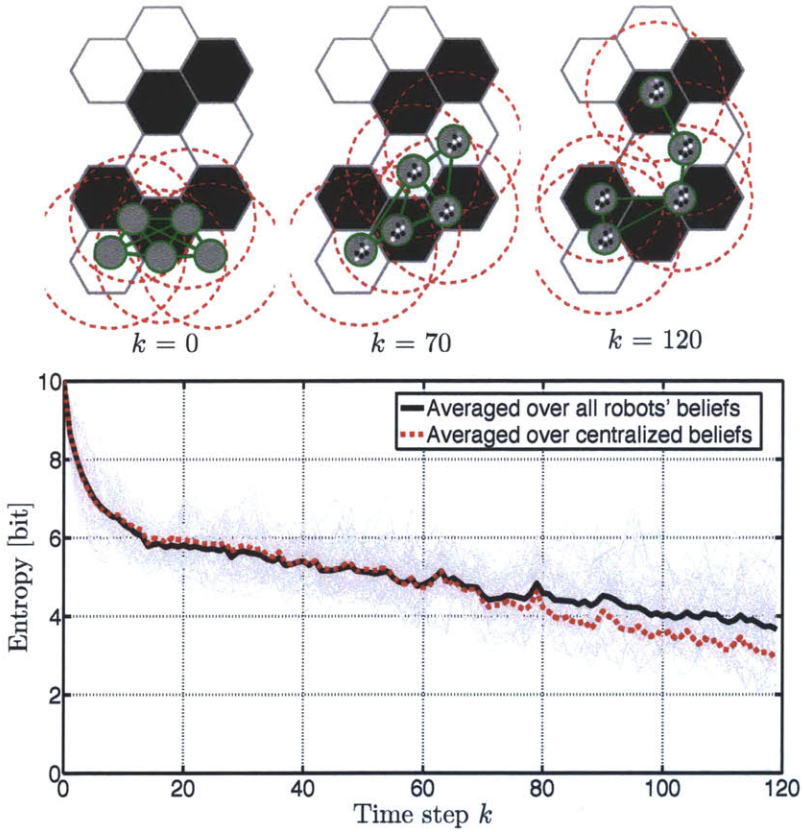


Figure 5-16: *Top:* The beginning, middle, and end configuration of a five robot experiment over a 10 cell environment, where the state of each cell is either 1 (black) or 0 (white). The robots are represented by the gray circles, within which their prior distributions can be visualized. The green lines represent network connectivity, and the dashed red circles represent the simulated sensors' footprints. *Bottom:* This plot shows the decrease in entropy of the inferences averaged over 10 consecutive runs. In addition, the light gray lines show the entropy of each robot's belief for every run. We note much more variability with the individual runs compared with the runs of Figure 5-14, however, the averaged entropies are similar between two different experiments.

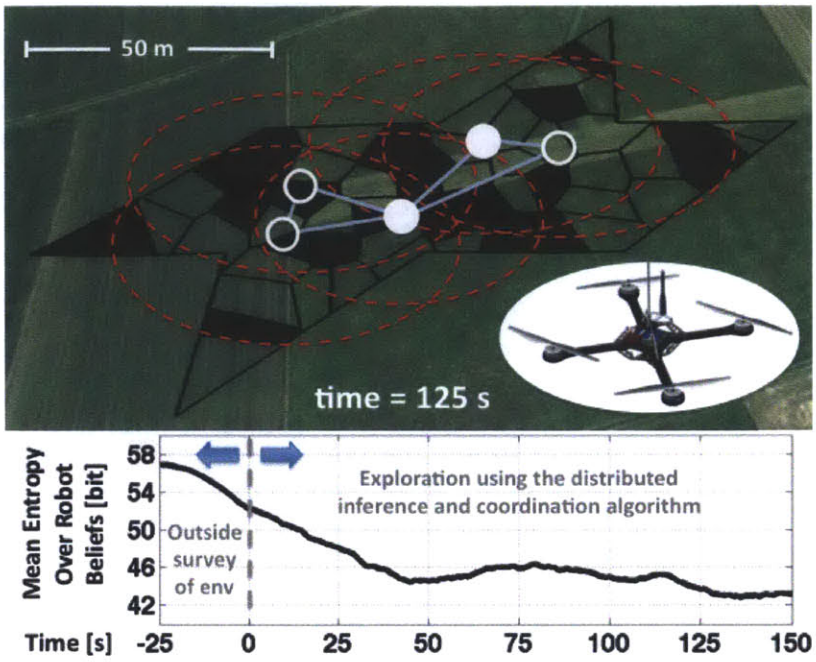


Figure 5-17: The deployment of five quadrotor flying robots (white ○) tasked to explore a 150 m wide outdoor environment containing 58 discretized cells of binary state. Exploration by the robotic sensor network (blue lines) is accomplished by a decentralized mutual-information-based gradient-ascent controller that continuously moves the robots to minimize the uncertainty associated with the inference. In parallel, the robots can be assigned by a higher level control scheme to act as dynamic network relays (white filled ○). The end result is a decrease in average entropy over time, as shown in the lower plot.

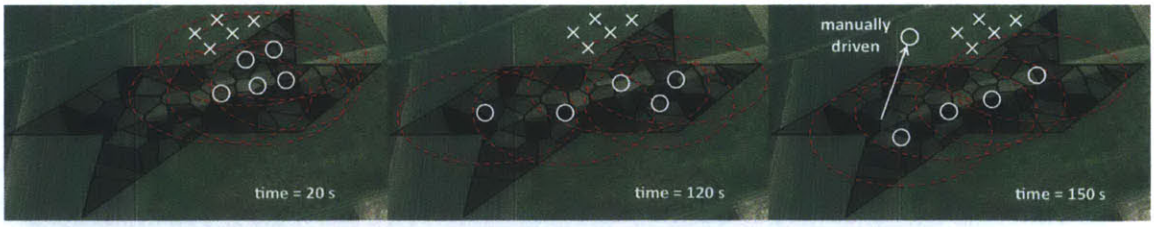


Figure 5-18: Preliminary deployment in preparation for the outdoor experiment (left). One robot is manually tasked to exit the environment (middle), which is compensated for by the other robots running a decentralized mutual information-based gradient-ascent controller (right).

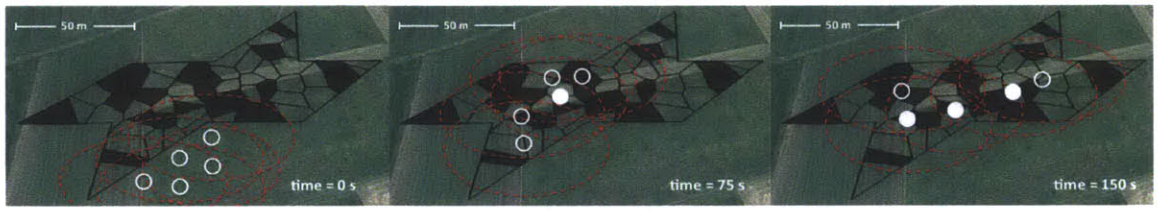


Figure 5-19: The time evolution of a constellation of five quadrotor flying robots (white circles) with simulated sensor (red dashed circles). These robots are tasked to explore a 150 m wide outdoor environment containing 58 discretized cells of binary state. *Left:* The experiment starts with all robots hovering at their starting positions. *Middle:* The robots at time = 75 s have begun to explore the environment. In addition, one robot is assigned by a higher level communication scheme to act as a dynamic network relay (white filled circle), and thus the control actions produced by the distributed controller are overridden for that robot. *Right:* At time = 150 s, even though three robots are assigned as dynamic relays, the distributed controller has driven the system into a configuration that covers a large portion of the cells.

Chapter 6

Distributed algorithms enabling decentralized inference and coordination

This thesis is dedicated to the development of scalable control solutions to environmental state estimation tasks such as tracking, surveillance, and exploration using large teams of autonomous robots equipped with sensors. Consider the task of using many aerial robots to monitor the flow of objects into and out of a major seaport (e.g., ships, containers, ground vehicles). To collectively estimate the objects' positions, one approach is to wirelessly communicate all sensor measurements to a data fusion center, perform the Bayesian estimation calculations in a centralized manner, and then globally broadcast the results to enable the robots to better position their sensors. For large systems, the central processor quickly becomes a computational and communication bottleneck, and thus is not considered to be scalable (Durrant-Whyte et al., 1990). Even if the robot controllers are implemented in a distributed fashion, the system can still experience scalability issues with sequential Bayesian estimation, and vice versa.

By design, the members of the novel controller class (Definition 48) require the same type of probabilistic information to be "fused" over the communication network when compared to decentralized Bayesian filters. More specifically, the filter and con-

troller require measurement likelihoods describing the realized joint observation (the former) and a sampled set of likely observations (the latter). The implication is that the same type of distributed algorithm (Lynch, 1997) can be used for approximating these distributions during each time step, although this property is not necessary (see Figure 5-5). In this chapter, we introduce two consensus-based algorithms (Olfati-Saber et al., 2005; Xiao et al., 2007), one to approximate discrete joint measurement probabilities and the other to approximate continuous joint measurement likelihood distributions. These distributed algorithms allow for resource adaptive Bayesian estimation and decentralized mutual information-based gradient-ascent control, for which the computational complexities do not depend on the number of robots.

We prove for all robots on a static and connected network graph that i) the approximations to the joint measurement probabilities converge to the joint of all the robots' local measurement probabilities; and ii) the approximations to the scaled Gaussian joint measurement likelihood distributions converge weakly[1] to the joint of all the robots' local measurement likelihood distributions. The given restrictions on the graph are used to derive bounds and convergence rates, specifically for the latter case. Yet, the implementation works on arbitrary networks without risk of catastrophic failures (e.g., robustness against robot failures), and without restriction on the number of communication rounds that the robots need to use for the consensus-based algorithm. An attractive aspect of this work is that expected performance provably improves as more system resources are allocated, i.e., as the computational and communication capabilities increase. We believe these theoretical contributions can drive the development of application specific sensor fusion approaches that are unbiased, convergent, and scalable. In addition, we used the provided consensus-based algorithms to support the decentralized hardware experiments previously described in Section 5.4.

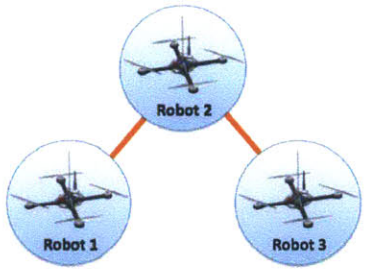
[1]Also known as convergence in distribution.

6.1 Distributed robot network

We introduce the multi-robot communication model used throughout this chapter, then present a distributed algorithm to discover the maximum in/out degree of the system.

6.1.1 Communication model

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$



$$\mathbf{W} = \begin{bmatrix} 2/3 & 1/3 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 2/3 \end{bmatrix}$$

Figure 6-1: An unweighted undirected network graph describing the communication capabilities of a three robot system. Also shown are the corresponding adjacency matrix (left) and the Metropolis-Hastings matrix (right).

Let the robots simultaneously transmit and receive messages on the network at a much faster rate than the rate at which joint observations are realized. Suppose that this message exchange happens multiple times between consecutive joint observations. We refer to a time period when a single message exchange occurs as a communication round, denoted κ, and assume that the exchanges within the robot system occur according to an undirected, unweighted network graph $\mathcal{G}$. Let the graph consist of a vertex set $\mathcal{V} = \{1, \ldots, n_r\}$ and an unordered edge set $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$; that is, $\{i, v\} \in \mathcal{E}$ if the ith and vth robots can directly communicate with one another, i.e., are neighbors. Let $N^{[i]}$ denote the set of neighbors of the ith robot, which has an in/out degree of $\Delta^{[i]} := |N^{[i]}|$. Let Δ denote the maximum in/out degree among all robots. The corresponding symmetric unweighted adjacency matrix $\mathbf{A} \in \{0, 1\}^{n_r \times n_r}$ for the network graph is defined as

$$[\mathbf{A}]_{iv} = \begin{cases} 1, & \text{if } \{i, v\} \in \mathcal{E}, \\ 0, & \text{otherwise,} \end{cases}$$

while the corresponding Metropolis-Hastings weight matrix $\mathbf{W} \in [0,1]^{n_r \times n_r}$ is

$$[\mathbf{W}]_{iv} = \begin{cases} 1 - \sum_{v' \in N^{[i]}} \frac{1}{\max\{\Delta^{[i]}, \Delta^{[v']}\}+1}, & i = v, \\ \frac{1}{\max\{\Delta^{[i]}, \Delta^{[v]}\}+1}, & \{i, v\} \in \mathcal{E}, \\ 0, & \text{otherwise.} \end{cases}$$

Note that $[\,\cdot\,]_{iv}$ and $[\,\cdot\,]_i$ denote the (i,v) matrix entry and ith vector entry. Figure 6-1 gives both the adjacency matrix and the Metropolis-Hastings matrix for an example three robot system.

Given the volatile nature of mobile networks, we expect the network graph to be incomplete, time-varying, and stochastic. The algorithms presented in this chapter work in practice even when properties of the network graph cannot be formalized. However, to allow for meaningful analysis from a theoretical perspective, we assume that this graph remains connected and is time-invariant between consecutive joint observations. Connectivity allows the system to be analyzed as a single unit instead of separate independent subsystems. The property of time-invariance between joint observations is more strict, however, this assumption is used to formalize the convergence of our consensus-based algorithm. We also assume an upper bound exists on each robot's in/out degree. This assumption comes from the property that physical communication devices support a finite bandwidth, and thus the number of neighbors a robot can have cannot increase without bound. We now formally state these assumptions below.

Assumption 65 (Connected network graph) *For any two vertices in $\mathcal{V}$, there exists a path[2] in the network graph $\mathcal{G}$.*

Assumption 66 (Time-invariant network graph) *Between two consecutive joint observations, we have that the network graph $\mathcal{G}$ is static.*

Assumption 67 (Bounded in/out degree) *For any number of robots n_r, there*

[2]A *path* in a graph is an ordered sequence of vertices such that any pair of consecutive vertices in the sequence is an edge of the graph.

exists a finite integer M such that the in/out degree $\Delta^{[i]}$ is less than M for any robot $i \in \{1 \dots n_r\}$.

Remark 68 (Simplification of notation) *Since the network graph is time-varying over multiple time steps k, we would normally augment its symbol $\mathcal{G}$ as well as other symbols with the subscript $\sqcup_k$. In this chapter, we only consider a single yet arbitrary time step, and thus remove this subscript for clarity.*

6.1.2 Discovery of the maximum in/out degree

The FloodMax algorithm is a well studied distributed algorithm used in leader elect problems (Lynch, 1997). Traditionally implemented, each robot would transmit the maximum unique identifier (UID) it received up to the given communication round κ. After $\text{diam}(\mathcal{G})$ communication rounds, where $\text{diam}(\cdot)$ represents the diameter of a graph, all robots would know the maximum UID in the network. To solve the leader elect problem, the robot whose own UID matches the maximum UID of the network would declare itself the leader.

For one of our consensus-based algorithms, the robots do not need to select a leader, but instead need to discover the maximum in/out degree, denoted Δ. Moreover, we assume that each robot only know characteristics that describe its local neighborhood, e.g., its own in/out degree $\Delta^{[i]}$. In other words, the robots do not know characteristics describing the overall network topology, such as the number of robots and the network diameter. This restriction implies that the robots may never identify that the maximum in/out degree has been found. Regardless, the robots can still reach an agreement during consensus by using in parallel the FloodMax algorithm described in the following lemma.

Lemma 69 (FloodMax for degree discovery) *Suppose Assumptions 65–66 hold, and consider the system of robots running a FloodMax algorithm of the form*

$$\vartheta_{\kappa+1}^{[i]} = \max\left\{\{\vartheta_{\kappa}^{[i]}\} \cup \{\vartheta_{\kappa}^{[v]} : v \in N^{[i]}\}\right\}, \tag{6.1}$$

where for each robot $\vartheta_\kappa^{[i]}$ *is initialized to the robot's in/out degree plus one, i.e.,* $\Delta^{[i]}+1$. *Then for all robots after* $\mathrm{diam}(\mathcal{G})$ *communication rounds,* $\vartheta_\kappa^{[i]}$ *is equal to the maximum in/out degree plus one, i.e.,* $\Delta+1$.

6.2 Consensus of joint measurement probabilities

Suppose Assumptions 65–66 hold, and consider the system of robots running a consensus algorithm (Olfati-Saber et al., 2005) of the form

$$\psi_{\kappa+1}^{[i]} = \psi_\kappa^{[i]} + \epsilon \sum_{v\in N^{[i]}} (\psi_\kappa^{[v]} - \psi_\kappa^{[i]}), \tag{6.2}$$

where $\epsilon \in (0, \Delta^{-1})$ guarantees for all robots that the state $\psi_\kappa^{[i]}$ exponentially converges to the average initial state of all robots, i.e., $\sum_{i=1}^{n_r} \psi_0^{[i]}/n_r$. For the robots to select a valid ϵ, they need to know either the maximum in/out degree of the network graph or the number of robots. Since we are assuming that neither parameter is known, the consensus algorithm is modified to use in parallel the FloodMax algorithm of the form (6.1). As a result, convergence to the average initial state is preserved as described in the following lemmas.

Lemma 70 (Convergence of modified consensus algorithm) *Suppose Assumptions 65–66 hold, and consider the system of robots running a FloodMax algorithm of the form* (6.1) *in parallel with two consensus algorithms of the form*

$$\psi_{\kappa+1}^{[i]} = \frac{\vartheta_{\kappa+1}^{[i]} - \vartheta_\kappa^{[i]}}{\vartheta_{\kappa+1}^{[i]}}\psi_0^{[i]} + \frac{\vartheta_\kappa^{[i]} - \Delta^{[i]}}{\vartheta_{\kappa+1}^{[i]}}\psi_\kappa^{[i]} + \frac{1}{\vartheta_{\kappa+1}^{[i]}} \sum_{v\in N^{[i]}} \psi_\kappa^{[v]}, \tag{6.3}$$

and

$$\pi_{\kappa+1}^{[i]} = (\pi_0^{[i]})^{\frac{\vartheta_{\kappa+1}^{[i]} - \vartheta_\kappa^{[i]}}{\vartheta_{\kappa+1}^{[i]}}} (\pi_\kappa^{[i]})^{\frac{\vartheta_\kappa^{[i]} - \Delta^{[i]}}{\vartheta_{\kappa+1}^{[i]}}} \prod_{v\in N^{[i]}} (\pi_\kappa^{[v]})^{\frac{1}{\vartheta_{\kappa+1}^{[i]}}}. \tag{6.4}$$

Then for all robots $i \in \{1, \dots n_r\}$, *the states* $\psi_\kappa^{[i]}$ *and* $\pi_\kappa^{[i]}$ *converge to* $\sum_{v=1}^{n_r} \psi_0^{[v]}/n_r$

and $\prod_{v=1}^{n_r} \left(\pi_0^{[v]}\right)^{1/n_r}$, *respectively, in the limit as* κ *tends to infinity.*

Lemma 71 (Convergence on complete network graphs) *For a complete network graph,*[3] $\psi_\kappa^{[i]}$ *and* $\pi_\kappa^{[i]}$ *converge for all robots after one communication round.*

Remark 72 (Order of calculations) *We require that the state* $\vartheta_\kappa^{[i]}$ *is available when calculating both* $\psi_\kappa^{[i]}$ *and* $\pi_\kappa^{[i]}$. *In other words,* (6.1) *is computed prior to* (6.3) *and* (6.4) *during a given communication round. See Algorithm 10 for an example algorithmic implementation of this concept.*

6.2.1 Distributive approximations for the sampled joint measurement probabilities

Suppose we wish to distribute a member of the novel controller class among robots sensing an environment of discrete-valued state, e.g., Algorithm 6. From Section 5.2.4, each robot must locally approximate the measurement probabilities $\mathbb{P}(\check{Y} = y | X = x)$ describing the sampled set of likely joint observations $y \in \check{\mathcal{Y}}$ for all $x \in \mathcal{X}$. As previously discussed, each robot represents these approximations with the n_y-tuple $\check{p}^{[i]}$, which for the current case has $|\mathcal{X}|$-tuple elements $\check{p}^{[i,\ell]}$ for all $\ell \in \{1, \ldots, n_y\}$. Let their be a predetermined ordering of elements in $|\mathcal{X}|$ that is known to the robots, i.e.,

$$\mathcal{X} = \left\{ x^{[j]} : j \in \{1, \ldots, |\mathcal{X}|\} \right\}.$$

For the robots to reach a consensus on the sampled joint measurement probabilities, let $\pi_\kappa^{[i]}$ be a belief matrix[4] representing the unnormalized approximated n_rth root of these probabilities known by the ith robot after κ communication rounds, i.e.,

$$\frac{\left[\pi_\kappa^{[i]}\right]_{j,\ell}}{\eta} \approx \mathbb{P}(\check{Y} = \check{y}^{[\ell]} | X = x^{[j]}),$$

[3] A complete network graph is a graph such that any pair of vertices is an edge of the graph (Bullo et al., 2009).

[4] We are using terminology introduced by Pearl (1988).

for all $j \in \{1, \ldots, |\mathcal{X}|\}$ and $\ell \in \{1, \ldots, n_y\}$, where η is a normalization factor. In addition, let the belief matrix be initialized as

$$[\pi_0^{[i]}]_{j,\ell} = \mathbb{P}(\check{Y}^{[i]} = \check{y}^{[i,\ell]} | X = x^{[j]}).$$

In words, the belief matrix is initialized to the ith robot's conditionally independent contribution to the unnormalized sampled joint measurement probabilities

$$\frac{\mathbb{P}(\check{Y}|X)}{\eta} = \prod_{i=1}^{n_r} \mathbb{P}(\check{Y}^{[i]}|X). \tag{6.5}$$

By using (6.4) to evolve the belief matrix over a finite number of communication rounds, termed the communication round size and denoted by n_π, we can formally define the approximation for the sampled joint measurement probabilities as

$$\check{p}^{[i,\ell]}(x^{[j]}) := \frac{[\pi^{[i]}]_{j\ell}^{\beta^{[i]}}}{\sum_{\ell'=1}^{n_y} [\pi^{[i]}]_{j\ell'}^{\beta^{[i]}}} \approx \mathbb{P}(\check{Y} = \check{y}^{[\ell]} | X = x^{[j]}) \tag{6.6}$$

for all $j \in \{1, \ldots, |\mathcal{X}|\}$ and $\ell \in \{1, \ldots, n_y\}$, where $\pi^{[i]}$ and $\beta^{[i]}$ are shorthand denoting $\pi_{n_\pi}^{[i]}$ and $\beta_{n_\pi}^{[i]}$, respectively. Here, $\beta_\kappa^{[i]}$ is an exponential factor accounting for the fact that the consensus-based algorithm may terminate before converging. More specifically, $\pi_\kappa^{[i]}$ can be thought of as the ith robot's weighted logarithmic summation of $\mathbb{P}(\check{Y}^{[v]}|X)$ over all $v \in \{1, \ldots, n_r\}$ at communication round κ, while $\beta_\kappa^{[i]}$ is the inverse of the largest weight to ensure that no single measurement probability in the right hand side product of (6.5) has an exponent of value larger than one. In other words, no observation "gets counted" more than once.

Remark 73 (Size of the belief matrix) *Due to the construction of the joint observation sample set $\check{\mathcal{Y}}$, the column size of the belief matrix remains constant with respect to the number of robots. However, the row dimension is linear with respect to the environment state alphabet size $|\mathcal{X}|$, which can be misleading since this alphabet size is usually exponential with respect to other quantities. For example, the alphabet*

size for an environment that is partitioned into cells can scale exponentially with respect to the number of cells, e.g., see the experiments in Section 5.4. Thus, a system designer will need to consider both lossless and lossy compression techniques specific for the application at hand.

Remark 74 (Anytime approximations) *Note that there is no restriction on how many communication rounds are needed for the distributed approximation* (6.6)*. In fact, if no communication rounds are performed, the each robot will believe it exists in a "lonely world" and* (6.6) *will correctly return that robot's local measurement properties. Therefore, consensus-based algorithms of this form can be considered as anytime (Russell et al., 1995).*

To calculate the exponential factor $\beta_\kappa^{[i]}$ in parallel with the belief matrix $\pi_\kappa^{[i]}$, let the state $\psi_\kappa^{[i]}$ evolve by using (6.3) and be initialized to $\mathbf{e}_i$. From the discussion above, we have that

$$\beta_\kappa^{[i]} = \|\psi_\kappa^{[i]}\|_\infty^{-1}. \tag{6.7}$$

Note that $\psi_0^{[i]}$ does not need to be initialized with a size of n_r, which would induce the requirement that the number of robots must be known. Instead, let each robot maintain a tuple of indices, which can be arbitrarily augmented when unknown indices are received during the communication round. We leave the details of this implementation to the reader.

From Lemma 70, we have for all robots that $\beta_\kappa^{[i]}$ converges to n_r in the limit as κ tends to infinity, or after one communication round if the network is complete (Lemma 71). This property is required for the convergence of the approximations to the true sampled joint measurement probabilities, which is stated in the following theorem and corollary.

Theorem 75 (Convergence of distributed approximations) *Suppose Assumptions 65–66 hold. Then for all robots* $i \in \{1, \ldots, n_r\}$, $j \in \{1, \ldots, |\mathcal{X}|\}$, *and* $\ell \in \{1, \ldots, n_y\}$, *we have that* $\check{p}^{[i,\ell]}(x^{[j]})$ *converges to* $\mathbb{P}(\check{Y} = \check{y}^{[\ell]} | X = x^{[j]})$ *in the limit as*

the total number of communication rounds n_π tends to infinity.

Corollary 76 (Convergence on a complete network graph) *Suppose Assumption 66 holds, and suppose the network graph $\mathcal{G}$ is complete. Then for all robots $i \in \{1, \ldots, n_r\}$, $j \in \{1, \ldots, |\mathcal{X}|\}$, and $\ell \in \{1, \ldots, n_y\}$, we have that $\check{p}^{[i,\ell]}(x^{[j]})$ is equal to $\mathbb{P}(\check{Y} = \check{y}^{[\ell]} | X = x^{[j]})$ after one communication round.*

With Theorem 75, we have that the robots can employ a consensus-based algorithm of form (6.6) to enable a decentralized mutual information-based gradient-ascent controller for sensing discrete-valued environment states. Moreover, this distributed approximation provably converges to the centralized solution as more communication rounds are performed, i.e., as network bandwidth increases. With Corollary 76, we have that the decentralized control problem reduces to the centralized problem for a complete network graph.

6.2.2 Algorithmic implementation

The algorithmic implementation to approximate the sampled joint measurement probabilities is summarized in Algorithm 10, and a discussion of its computational complexity follows in Proposition 77. Note that for the experiments discussed in Section 5.4, we have that Algorithm 10 was used for the decentralized controller and Bayesian filter. Please refer to Figure 5-5 for a system level overview of these processes.

Proposition 77 (Computational complexity) *Suppose Assumption 67 holds for a fix number of communication rounds n_π. Then the time and space complexities of Algorithm 10 are both $O(n_y|\mathcal{X}|)$. In words, both the time and space complexities to approximate the sampled joint measurement probabilities are at worst linear in the sizes of the joint observation sample set and the environment state alphabet. Moreover, these computational complexities are independent of the number or robots.*

Algorithm 10 MeasurementProbabilitiesConsensus($\check{\mathcal{Y}}^{[i]}$)

// Initialize FloodMax state
$\vartheta_0^{[i]} \leftarrow (\Delta^{[i]} + 1)$;
// Initialize consensus state (shorthand)
$\psi_0^{[i]} \leftarrow \mathbf{e}_i$;
for $j = 1$ to $|\mathcal{X}|$ **do**
  for $\ell = 1$ to n_y **do**
    // Initialize belief matrix
    $[\pi_0^{[i]}]_{j\ell} \leftarrow \mathbb{P}(\check{Y}^{[i]} = \check{y}^{[i,\ell]} | X = x^{[j]})$;
  end for
end for
for $\kappa = 1$ to n_π **do**
  $\vartheta_\kappa^{[i]} \leftarrow \max\left\{\{\vartheta_{\kappa-1}^{[i]}\} \cup \{\vartheta_{\kappa-1}^{[v]} : v \in N^{[i]}\}\right\}$;
  // Update consensus state $\psi_\kappa^{[i]}$ from (6.3), note calculations are element-wise
  $\psi_\kappa^{[i]} \leftarrow \frac{\vartheta_\kappa^{[i]} - \vartheta_{\kappa-1}^{[i]}}{\vartheta_\kappa^{[i]}} \psi_0^{[i]} + \frac{\vartheta_{\kappa-1}^{[i]} - \Delta^{[i]}}{\vartheta_\kappa^{[i]}} \psi_{\kappa-1}^{[i]} + \frac{1}{\vartheta_\kappa^{[i]}} \sum_{v \in N^{[i]}} \psi_{\kappa-1}^{[v]}$;
  // Update belief matrix $\pi_\kappa^{[i]}$ from (6.4), note calculations are element-wise
  $\pi_\kappa^{[i]} \leftarrow (\pi_0^{[i]})^{\frac{\vartheta_\kappa^{[i]} - \vartheta_{\kappa-1}^{[i]}}{\vartheta_\kappa^{[i]}}} (\pi_{\kappa-1}^{[i]})^{\frac{\vartheta_{\kappa-1}^{[i]} - \Delta^{[i]}}{\vartheta_\kappa^{[i]}}} \prod_{v \in N^{[i]}} (\pi_{\kappa-1}^{[v]})^{\frac{1}{\vartheta_\kappa^{[i]}}}$;
end for
// Calculate exponential factor $\beta^{[i]}$ from (6.7)
$\beta^{[i]} \leftarrow \|\psi^{[i]}\|_\infty^{-1}$;
for $j = 1$ to $|\mathcal{X}|$ **do**
  // Initialize normalization term
  $\eta \leftarrow 0$;
  for $\ell = 1$ to n_y **do**
    // Build normalization term found in (6.6)
    $\eta \leftarrow \eta + [\pi^{[i]}]_{j\ell}^{\beta^{[i]}}$;
  end for
  for $\ell = 1$ to n_y **do**
    // Insert entries into tuple $\check{p}^{[i]}$
    $\check{p}^{[i,\ell]}(x^{[j]}) \leftarrow \frac{[\pi^{[i]}]_{j\ell}^{\beta^{[i]}}}{\eta}$;
  end for
end for
return $\check{p}^{[i]}$;

6.3 Consensus of joint measurement likelihood distributions

We will be employing a similar consensus-based approach as in Section 6.2.1 to approximate joint measurement likelihood distributions. To emphasize the overall flexibility of the general approach, we consider the system of robots running consensus algorithms of the form

$$\psi_{\kappa+1}^{[i]} = [\mathbf{W}]_{ii}\psi_{\kappa}^{[i]} + \sum_{v \in N^{[i]}} [\mathbf{W}]_{iv}\psi_{\kappa}^{[v]} \tag{6.8}$$

and

$$\pi_{\kappa+1}^{[i]} = (\pi_{\kappa}^{[i]})^{[\mathbf{W}]_{ii}} \prod_{v \in N^{[i]}} (\pi_{\kappa}^{[v]})^{[\mathbf{W}]_{iv}}. \tag{6.9}$$

Similarly to (6.3) and (6.4), we have for all robots that the states $\psi_{\kappa}^{[i]}$ and $\pi_{\kappa}^{[i]}$ converge to $\sum_{i=1}^{n_r} \psi_0^{[i]}/n_r$, and $\prod_{i=1}^{n_r} \left(\psi_0^{[i]}\right)^{1/n_r}$, respectively (Xiao et al., 2007). Again, this convergence happens for robots without knowledge of the network's maximum in/out degree or the number of robots. Note that the distributive algorithms research community is quite large, and that many other consensus-based algorithms yielding asymptotic averages can be modified for our approach.

6.3.1 Distributive approximations for scaled Gaussian joint measurement likelihood distributions

Suppose we wish to distribute a member of the novel controller class among robots sensing an environment of continuous-valued state, e.g., Algorithm 7. Similar to Section 6.2.1, each robot must locally approximate the likelihood distribution $\mathbb{P}(\check{Y} = y|X = x)$ describing the sampled set of likely joint observations $y \in \check{\mathcal{Y}}$ for all $x \in \mathcal{X}$. Consider the ℓth element $y^{[\ell]}$ from the joint observation sample set $\check{\mathcal{Y}}$. Now suppose for all robots $i \in \{1, \ldots, n_r\}$, the corresponding local likelihood distribution $\mathbb{P}(\check{Y}^{[i]} =$

$y^{[i,\ell]}|X)$ can be accurately represented by a non-degenerate[5] n_g-dimensional scaled Gaussian

$$\mathbb{P}(\check{Y}^{[i]} = y^{[i,\ell]}|X) = \rho^{[i,\ell]}\mathcal{N}_c(\xi_0^{[i,\ell]}, \mathbf{\Omega}_0^{[i,\ell]}),$$

with

$$\mathcal{N}_c(\xi, \mathbf{\Omega}) := \eta_c(\xi, \mathbf{\Omega})e^{x^T\xi - \frac{1}{2}x^T\mathbf{\Omega}x}$$

and

$$\eta_c(\xi, \mathbf{\Omega}) := \frac{\det(\mathbf{\Omega})^{\frac{1}{2}}}{\left((2\pi)^{n_g}e^{\xi^T\mathbf{\Omega}^{-1}\xi}\right)^{\frac{1}{2}}},$$

where for each robot, we have that $\xi_0^{[i,\ell]} \in \mathbb{R}^{n_g}$ is its information vector, $\mathbf{\Omega}_0^{[i,\ell]} \in \mathbb{R}^{n_g \times n_g}$ is its information matrix, and $\rho^{[i,\ell]} := \mathbb{P}(Y^{[i]} = y^{[i,\ell]})$ is its scaling factor. For such distributions, we have for the sampled joint measurement likelihood distribution that

$$\mathbb{P}(\check{Y} = y^{[\ell]}|X) = p^{[\ell]}\mathcal{N}_c(\xi^{[\ell]}, \mathbf{\Omega}^{[\ell]}),$$

where $\xi^{[\ell]} = \sum_{i=1}^{n_r} \xi_0^{[i,\ell]}$ is the joint information vector, $\mathbf{\Omega}^{[\ell]} = \sum_{i=1}^{n_r} \mathbf{\Omega}_0^{[i,\ell]}$ is the joint information matrix, and

$$\rho^{[\ell]} = \frac{1}{\eta_c(\xi, \mathbf{\Omega})}\prod_{i=1}^{n_r} \rho^{[i,\ell]}\eta_c(\xi_0^{[i,\ell]}, \mathbf{\Omega}_0^{[i,\ell]})$$

is the joint scaling factor.

For a given environment state $x \in \mathcal{X}$, let $\pi_\kappa^{[i,\ell]} \in \mathbb{R}_{\geq 0}$ be initialized to the corresponding robot's local measurement likelihood distribution evaluated at x, i.e., $\pi_0^{[i,\ell]} = \mathbb{P}(\check{Y}^{[i]} = y^{[i,\ell]}|X = x)$. In addition, let $\psi_\kappa^{[i]} \in [0,1]^{n_r}$ be initialized to the standard basis $\mathbf{e}_i$ pointing in the ith direction of $\mathbb{R}^{n_r}$. Similar to the approach in Section 6.2.1, we can use (6.9) and (6.8) at each communication round to have $(\pi_\kappa^{[i,\ell]})^{\beta_\kappa^{[i]}}$

[5]By non-degenerate we mean that the information matrix of a Gaussian is a real positive-definite symmetric matrix.

converge to the joint measurement likelihood distribution evaluated at x in the limit as $\kappa \to \infty$, where again $\beta_\kappa^{[i]} := \|\psi_\kappa^{[i]}\|_\infty^{-1}$ is a scalar exponential factor that converges to n_r. The expansion of $(\pi_\kappa^{[i,\ell]})^{\beta_\kappa^{[i]}}$ leads to the following theorem.

Theorem 78 (Consensus of scaled Gaussian likelihood distributions) *Suppose Assumptions 65–66 hold. For all robots $i \in \{1 \dots n_r\}$ and samples $\ell \in \{1, \dots, n_y\}$, let $\psi_\kappa^{[i]} \in [0,1]^{n_r}$, $\xi_\kappa^{[i,\ell]} \in \mathbb{R}^{n_g}$, and $\mathbf{\Omega}_\kappa^{[i,\ell]} \in \mathbb{R}^{n_g \times n_g}$ be initialized to $\boldsymbol{e}_i$, $\xi_0^{[i,\ell]}$, and $\mathbf{\Omega}_0^{[i,\ell]}$, respectively, and have all evolve according to* (6.8)*. In addition, let $\alpha_\kappa^{[i,\ell]} \in \mathbb{R}_{>0}$ be initialized to $\rho^{[i,\ell]}\eta_c(\xi_0^{[i,\ell]}, \mathbf{\Omega}_0^{[i,\ell]})$ and evolve according to* (6.9)*. We then have that*

$$\frac{(\alpha_\kappa^{[i,\ell]})^{\beta_\kappa^{[i]}}}{\eta_c(\beta_\kappa^{[i]}\xi_\kappa^{[i,\ell]}, \beta_\kappa^{[i]}\mathbf{\Omega}_\kappa^{[i,\ell]})}\mathcal{N}_c(\beta_\kappa^{[i]}\xi_\kappa^{[i,\ell]}, \beta_\kappa^{[i]}\mathbf{\Omega}_\kappa^{[i,\ell]}) \to \rho^{[\ell]}\mathcal{N}_c(\xi^{[\ell]}, \mathbf{\Omega}^{[\ell]}), \quad \forall x \in \mathcal{X}, \tag{6.10}$$

as κ tends to infinity. In other words, the expression on the left hand side of (6.10) *converges weakly to the sampled joint measurement likelihood distribution.*

Given Theorem 78, we can formally define the approximations to the sampled joint measurement likelihood distributions as

$$\check{p}^{[i,\ell]} := \frac{(\alpha^{[i,\ell]})^{\beta^{[i]}}}{\eta_c(\beta^{[i]}\xi^{[i,\ell]}, \beta^{[i]}\mathbf{\Omega}^{[i,\ell]})}\mathcal{N}_c(\beta^{[i]}\xi^{[i,\ell]}, \beta^{[i]}\mathbf{\Omega}^{[i,\ell]}) \tag{6.11}$$

for all $j \in \{1, \dots, |\mathcal{X}|\}$ and $\ell \in \{1, \dots, n_y\}$, where $\alpha^{[i,\ell]}$, $\xi^{[i,\ell]}$, and $\mathbf{\Omega}^{[i,\ell]}$ are shorthand denoting $\alpha_{n_\pi}^{[i,\ell]}$, $\xi_{n_\pi}^{[i,\ell]}$, and $\mathbf{\Omega}_{n_\pi}^{[i,\ell]}$, respectively. Again, the exponential factor $\beta_\kappa^{[i]}$ assures that no observation "gets counted" more than once – a concept that will be discussed in Section 6.3.3.

6.3.2 Algorithmic implementation

The algorithmic implementation to approximate the sampled joint measurement likelihood distributions is summarized in Algorithm 11, and a discussion of its computational complexity follows in Proposition 79.

Proposition 79 (Computational complexity) *Suppose Assumption 67 holds for a fix the number of communication rounds n_π. Then the time and space complexities*

Algorithm 11 LikelihoodDistributionConsensus($\check{\mathcal{Y}}^{[i]}$)

// Initialize consensus state
$\psi_0^{[i]} \leftarrow \mathbf{e}_i$;
for $\ell = 1$ to n_y **do**
  // Initialize information vector, information matrix, and normalization factor
  $\xi_0^{[i,\ell]} \leftarrow \text{InformationVector}\left(\mathbb{P}(\check{Y}^{[i]} = \check{y}^{[i,\ell]}|X)\right)$;
  $\mathbf{\Omega}_0^{[i,\ell]} \leftarrow \text{InformationMatrix}\left(\mathbb{P}(\check{Y}^{[i]} = \check{y}^{[i,\ell]}|X)\right)$;
  $\alpha_0^{[i,\ell]} \leftarrow \mathbb{P}(\check{Y}^{[i]} = \check{y}^{[i,\ell]})\eta_c(\xi_0^{[i,\ell]}, \mathbf{\Omega}_0^{[i,\ell]})$;
end for
for $\kappa = 1$ to n_π **do**
  // Update consensus states $\psi_\kappa^{[i,\ell]}$ from (6.8), note calculations are element-wise
  $\psi_\kappa^{[i]} \leftarrow [\mathbf{W}]_{ii}\psi_\kappa^{[i]} + \sum_{v\in N^{[i]}}[\mathbf{W}]_{iv}\psi_\kappa^{[v]}$;
  for $\ell = 1$ to n_y **do**
    // Update consensus states $\xi_\kappa^{[i,\ell]}$ and $\mathbf{\Omega}_\kappa^{[i,\ell]}$ using (6.8)
    // note calculations are element-wise
    $\xi_\kappa^{[i,\ell]} \leftarrow [\mathbf{W}]_{ii}\xi_\kappa^{[i,\ell]} + \sum_{v\in N^{[i,\ell]}}[\mathbf{W}]_{iv}\xi_\kappa^{[v,\ell]}$;
    $\mathbf{\Omega}_\kappa^{[i,\ell]} \leftarrow [\mathbf{W}]_{ii}\mathbf{\Omega}_\kappa^{[i,\ell]} + \sum_{v\in N^{[i,\ell]}}[\mathbf{W}]_{iv}\mathbf{\Omega}_\kappa^{[v,\ell]}$;
    // Update consensus state $\alpha_\kappa^{[i,\ell]}$ using (6.9)
    // note calculations are element-wise
    $\alpha_\kappa^{[i,\ell]} \leftarrow (\alpha_\kappa^{[i,\ell]})^{[\mathbf{W}]_{ii}} \prod_{v\in N^{[i,\ell]}}(\alpha_\kappa^{[v,\ell]})^{[\mathbf{W}]_{iv}}$;
  end for
end for
// Calculate exponential factor $\beta^{[i]}$ from (6.7)
$\beta^{[i]} \leftarrow \|\psi^{[i]}\|_\infty^{-1}$;
for $\ell = 1$ to n_y **do**
  // Insert entries into tuple $\check{p}^{[i]}$ via (6.11)
  $\check{p}^{[i,\ell]} \leftarrow \frac{(\alpha^{[i,\ell]})^{\beta^{[i]}}}{\eta_c(\beta^{[i]}\xi^{[i,\ell]}, \beta^{[i]}\mathbf{\Omega}^{[i,\ell]})}\mathcal{N}_c(\beta^{[i]}\xi^{[i,\ell]}, \beta^{[i]}\mathbf{\Omega}^{[i,\ell]})$;
end for
return $\check{p}^{[i]}$;

of Algorithm 11 are both $O(n_y n_g^2)$. In words, both the time and space complexities to approximate the sampled joint measurement probabilities are at worst linear in the size of the joint observation sample set and quadratic in the dimension of the representative Gaussians. Moreover, these computational complexities are independent of the number or robots.

6.3.3 Performance guarantees

In the following subsection, we simplify notation by dropping sample index ℓ. For example, we have that $\xi_\kappa^{[i]}$ and $\mathbf{\Omega}_\kappa^{[i]}$ denote $\xi_\kappa^{[i,\ell]}$ and $\mathbf{\Omega}_\kappa^{[i,\ell]}$, respectively.

We begin to characterize the approximations to the sampled joint measurement likelihood distributions by proving that the corresponding Gaussian-like distributions 'make sense.' Since we are forming these distributions from the canonical parameters $\beta_\kappa^{[i]}\xi_\kappa^{[i]}$ and $\beta_\kappa^{[i]}\mathbf{\Omega}_\kappa^{[i]}$, this making sense objective is equivalent to proving for all robots $i \in \{1,\ldots,n_r\}$ and communication rounds $\kappa \in \mathbb{Z}_{\geq 0}$ that $\beta_\kappa^{[i]}\xi_\kappa^{[i]}$ is a real vector and $\beta_\kappa^{[i]}\check{\mathbf{\Omega}}_\kappa^{[i]}$ is a real positive-definite symmetric matrix. Since the collection of real vectors and the collection of positive-definite symmetric matrices are both closed under addition and positive scalar multiplication ($\beta_\kappa^{[i]} \in [1, n_r]$ from the upcoming Lemma 81), it holds that the likelihood approximation is composed of non-degenerate scaled Gaussians.

The guarantee of non-degeneracy is fundamental to many of the claims to come. More interestingly, the mathematical structure of (6.8) that allows this guarantee also allows for intuitive interpretations of how the approximations evolve over time, especially concerning the rate of convergence of the canonical parameters. We will discuss these discussed shortly, but first we review the concept of exponentially decreasing disagreement (Olfati-Saber et al., 2005; Xiao et al., 2007).

Lemma 80 (Exponentially decreasing disagreement) *For all robots and communication rounds, we have that*

$$\left\| \psi_\kappa^{[i]} - 1\frac{1}{n_r} \right\|_2 \leq U_\kappa^\psi := \left\| \boldsymbol{W} - \mathbf{1}\mathbf{1}^T\frac{1}{n_r} \right\|_2^\kappa \left(1 - \frac{1}{n_r}\right)^{\frac{1}{2}},$$

where lefthand side of the inequality is termed disagreement and $\|\cdot\|_2$ *for a matrix denotes the spectral norm.*

Lemma 81 (Properties of consensus state) *For all robots and communication rounds, we have that* $\psi_\kappa^{[i]} \in [0,1]^{n_r}$, $\|\psi_\kappa^{[i]}\|_1 = 1$, *and* $\|\psi_\kappa^{[i]}\|_\infty \geq 1/n_r$.

Except when run on a complete network graph, we expect Algorithm 11 to prematurely terminate before the canonical parameters converge, and thus the exponential factor $\beta_\kappa^{[i]}$ indicates how 'close' the approximated information vector $\xi_\kappa^{[i]}$ and information matrix $\mathbf{\Omega}_\kappa^{[i]}$ are to the true joint canonical parameters ξ and $\mathbf{\Omega}$, respectively. In the following, we provide a strictly increasing lower bound for the exponential factor that equals one at $\kappa = 0$ and converges to n_r in the limit as κ tends to infinity.

Proposition 82 (Lower bound for the exponential factor) *For all robots and communication rounds, we have that*

$$\beta_\kappa^{[i]} \geq L_\kappa^\beta := \left(U_\kappa^\psi \sqrt{1 - \frac{1}{n_r}} + \frac{1}{n_r} \right)^{-1}.$$

We now focus our attention to the geometric interpretation of the information matrix $\beta_\kappa^{[i]} \mathbf{\Omega}_\kappa^{[i]}$, which describes ellipsoidal contours of equal density for the corresponding scaled Gaussian. The squared lengths of the contours' principal axes are given by the inverse of the information matrix eigenvalues, with larger values representing distribution axes of higher certainty. As more communication rounds are performed and the information matrix converges element-wise, we expect this certainty to increase and also converge. This is in fact the case, and by using the lower bound for the exponential factor, we provide a strictly increasing lower bound for the information matrix eigenvalues.

Proposition 83 (Lower bound for the information matrix eigenvalues) *Let* $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{n_g}$. *Then for all robots, communication rounds, and* $m \in \{1, \ldots, n_g\}$, *we*

have that

$$\lambda_m(\beta_\kappa^{[i]}\boldsymbol{\Omega}_\kappa^{[i]}) \geq L_{k,m}^{\boldsymbol{\Omega}} := \max\{L_{k,m}^{\boldsymbol{\Omega}-}, L_{k,m}^{\boldsymbol{\Omega}+}\},$$

where

$$L_{k,m}^{\boldsymbol{\Omega}-} := \sum_{\ell=1}^{\lfloor L_\kappa^\beta \rfloor} \lambda_{n_g}(\boldsymbol{\Omega}_0^{[\ell]}) + \left(L_\kappa^\beta - \lfloor L_\kappa^\beta \rfloor\right)\lambda_{n_g}(\boldsymbol{\Omega}_0^{[\lceil L_\kappa^\beta \rceil]})$$

with the robot indices ordered such that $\lambda_1(\boldsymbol{\Omega}_0^{[1]}) \leq \lambda_1(\boldsymbol{\Omega}_0^{[2]}) \leq \cdots \leq \lambda_1(\boldsymbol{\Omega}_0^{[n_r]})$, *and where*

$$L_{k,m}^{\boldsymbol{\Omega}+} := \lambda_m(\boldsymbol{\Omega}) - \sum_{\ell=\lceil L_\kappa^\beta \rceil+1}^{n_r} \lambda_{n_g}(\boldsymbol{\Omega}_0^{[\ell]}) - \left(\lceil L_\kappa^\beta \rceil - L_\kappa^\beta\right)\lambda_{n_g}(\boldsymbol{\Omega}_0^{[\lceil L_\kappa^\beta \rceil]})$$

with $\lambda_{n_g}(\boldsymbol{\Omega}_0^{[1]}) \leq \lambda_{n_g}(\boldsymbol{\Omega}_0^{[2]}) \leq \cdots \leq \lambda_{n_g}(\boldsymbol{\Omega}_0^{[n_r]})$.

Remark 84 (Maximum of two bounds) *The use of both* $L_{k,m}^{\boldsymbol{\Omega}-}$ *and* $L_{k,m}^{\boldsymbol{\Omega}+}$ *yields an intuitive bound on* $\lambda_m(\beta_\kappa^{[i]}\boldsymbol{\Omega}_\kappa^{[i]})$ *in the instances where* $\kappa = 0$ *and* $\kappa \to \infty$, *respectively. The former implies* $\lambda_m(\boldsymbol{\Omega}_0^{[i]}) \geq \min_v \lambda_m(\boldsymbol{\Omega}_0^{[v]})$ *and the latter with Lemma 81 implies* $\lim_{k\to\infty} \lambda_m(\beta_\kappa^{[i]}\boldsymbol{\Omega}_\kappa^{[i]}) = \lambda_m(\boldsymbol{\Omega})$, *both of which are obvious. In addition, the two bounds are equivalent for univariate Gaussians (i.e.,* $n_g = 1$*).*

Lastly, we derive the strictly shrinking range for the information vector elements, which when combined with the bounds on the information matrix eigenvalues well characterizes the convergence behavior of the resulting scaled Gaussians. We believe such characterizations can lead to bounds on such information theoretic metrics such as Kullback-Leibler divergence of the mixture of Gaussians, however, such efforts are reserved for future work.

Proposition 85 (Bounds on the information vector elements) *For all robots, samples, communication rounds, and* $m \in \{1, \ldots, n_g\}$, *we have that*

$$L_{k,m}^{\boldsymbol{\xi}} \leq [\beta_\kappa^{[i]}\boldsymbol{\xi}_\kappa^{[i]}]_m \leq U_{k,m}^{\boldsymbol{\xi}},$$

where

$$L^{\xi}_{k,m} := \sum_{v=1}^{\lfloor L^{\beta}_{\kappa} \rfloor} [\xi_0^{[v]}]_m + (L^{\beta}_{\kappa} - \lfloor L^{\beta}_{\kappa} \rfloor)[\xi_0^{[\lceil L^{\beta}_{\kappa} \rceil]}]_m$$

with the robot indices arranged such that $[\xi_0^{[1]}]_m \leq [\xi_0^{[2]}]_m \leq \cdots \leq [\xi_0^{[n_r]}]_m$, *and where* $U^{\xi}_{k,m}$ *is defined the same as* $L^{\xi}_{k,m}$ *but with* $[\xi_0^{[1]}]_m \geq [\xi_0^{[2]}]_m \geq \cdots \geq [\xi_0^{[n_r]}]_m$.

6.4 Parallelized numerical simulations

Using the LLGrid computer cluster system at MIT Lincoln Laboratory, numerical simulations employing the various algorithms discussed in this thesis were performed in distributed fashion. More specifically, both the decentralized algorithms and the simulated dynamics for each robot ran on an independent computer cluster node, which exchanged messages with other nodes using MatlabMPI. In the following section, we discuss both consensus-only simulations and consensus-enabled control simulations.

6.4.1 Consensus-only simulations approximating measurement likelihood distributions

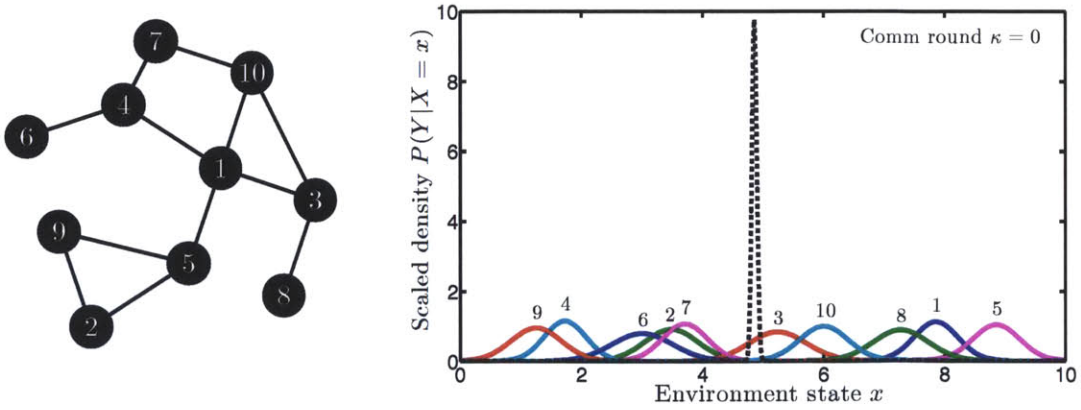


Figure 6-2: *Left:* A connected network graph on which ten robots distributively estimate their joint measurement likelihood distribution. *Right:* One dimensional normalized local measurement likelihood distributions for ten robots with respect to their normalized joint (dashed black curve).

We first consider ten robots distributively approximating their joint measurement likelihood distributions from local distributions that can be accurately represented by univariate scaled Gaussians. Such a simplified task best illustrates how each robot's joint approximation converges weakly to the true one. Figure 6-2 shows the normalized local measurement likelihood distributions, which for robots $1, \ldots, 10$ have distributions of $\mathcal{N}_c(22, 2.8)$, $\mathcal{N}_c(8, 2.3)$, $\mathcal{N}_c(11, 2.1)$, $\mathcal{N}_c(5, 2.9)$, $\mathcal{N}_c(23, 2.6)$, $\mathcal{N}_c(6, 2.0)$, $\mathcal{N}_c(10, 2.7)$, $\mathcal{N}_c(16, 2.2)$, $\mathcal{N}_c(16, 2.2)$, $\mathcal{N}_c(3, 2.4)$, $\mathcal{N}_c(15, 2.5)$. Note that we selected the canonical parameters to separate the distributions for illustrative purposes, as one should not expect such initial disagreement within a fielded robot system. Figure 6-2 also shows the normalized joint of nonzero mean, since the assumption of zero mean can lead to misleadingly tight bounds, e.g., bounds that are not invariant under translation.

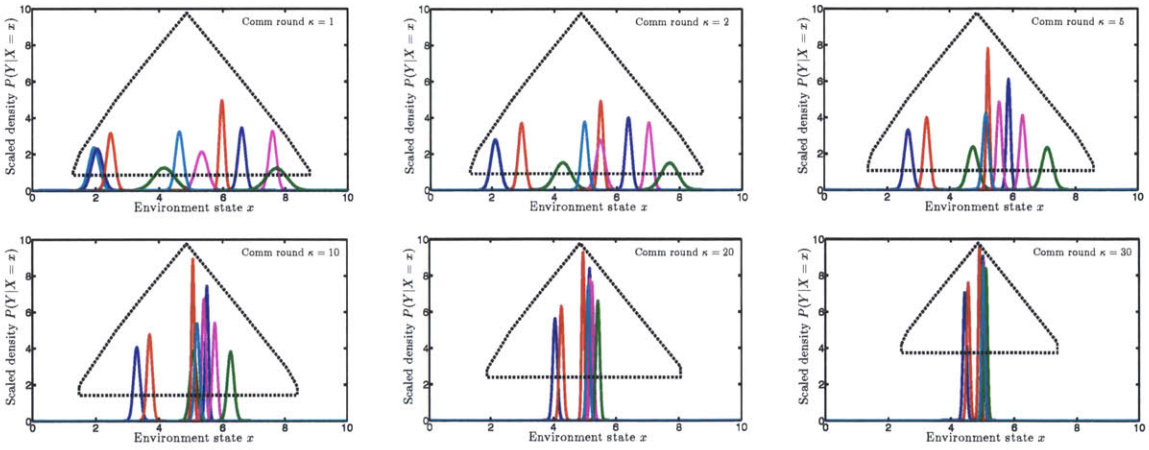


Figure 6-3: *Left to right, top to bottom:* The evolution of each robot's normalized joint approximation on the connected network graph in Figure 6-2 at communication rounds of $\kappa \in \{1, 2, 5, 10, 20, 30\}$. The dashed envelope represents the feasible region within which the peak of every robot's normalized approximation must lie.

We evaluated the performance of Algorithm 11 on the connected network graph shown. Figure 6-3 shows the evolution of each robot's normalized joint approximation with respect to a strictly shrinking envelope derived from bounds given in Propositions 83–85. These envelopes can be interpreted as feasible regions within which the peaks of all the robots' approximations must lie, intuitively highlighting the performance guarantees discussed in Section 6.3.3. We note that these bounds

for this particular network graph are conservative; we found that graphs with higher algebraic connectivity tend to produce tighter bounds, as shown in Figure 6-3.

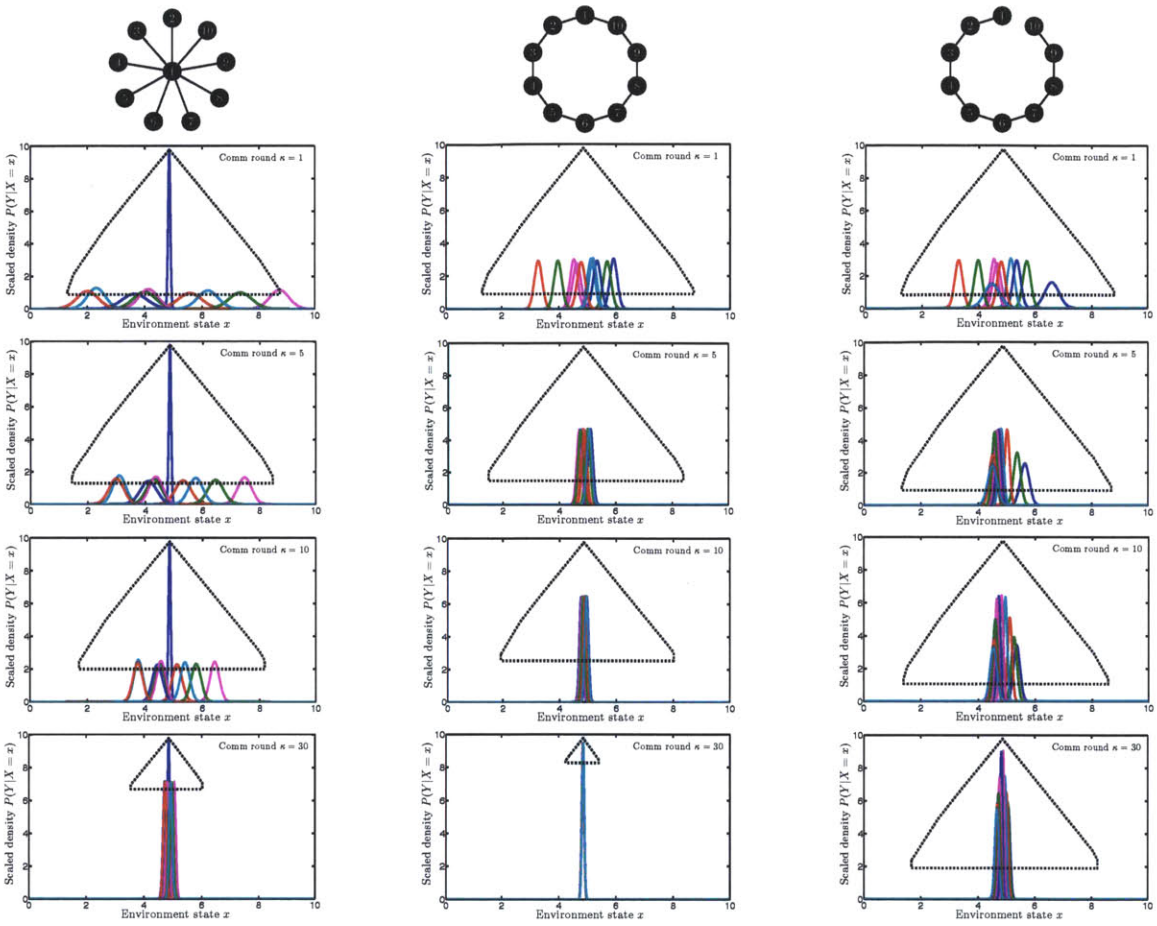


Figure 6-4: *Top to bottom:* The evolution of each robot's normalized joint approximation on a star network graph (left column), chain network graph (middle column), and a ring network graph (right column) at communication rounds of $\kappa \in \{1, 5, 10, 30\}$. The dashed envelope represents the feasible region within which the peak of every robot's approximation must lie.

6.4.2 Control simulations performing consensus for the measurement probabilities

To demonstrate the scalability of our complete control approach with respect to the number of robots, we simulated a $n_r = 100$ robot system using different communication round sizes of n_π for Algorithm 11. For each run, the heterogeneous sensors

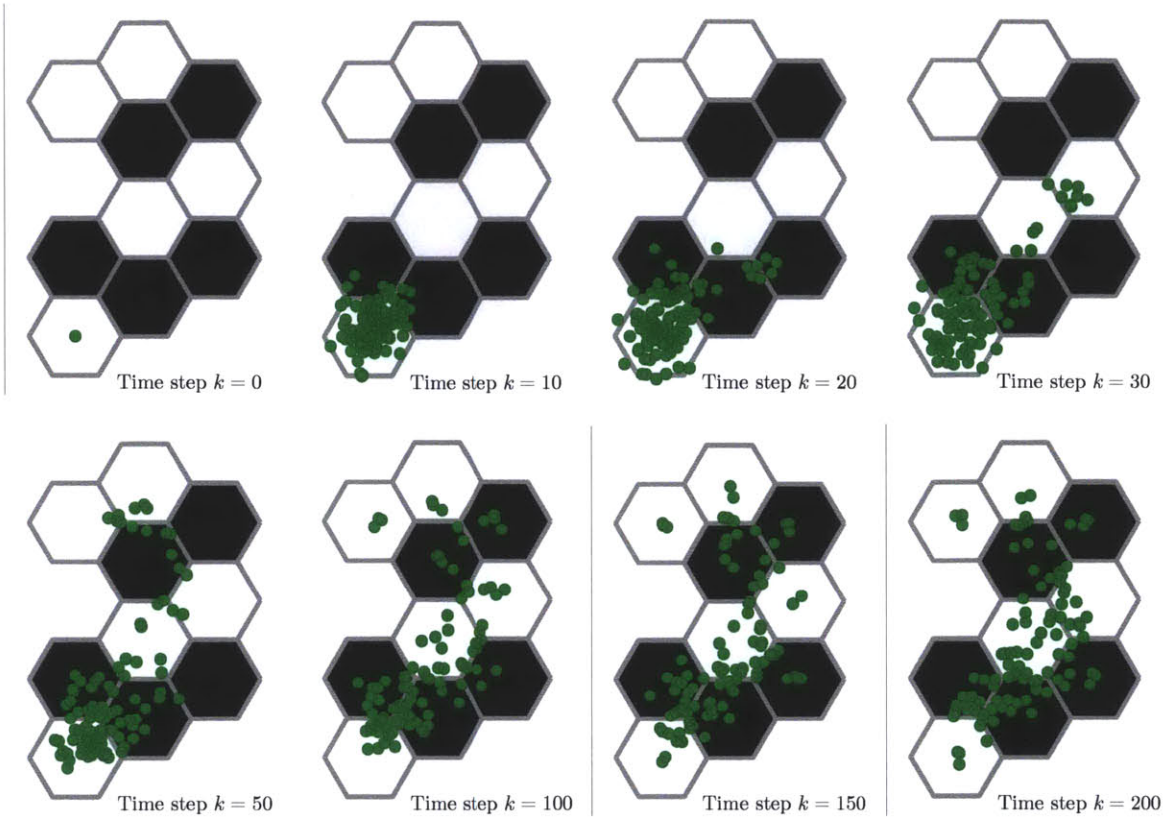


Figure 6-5: The position evolution of a simulated system of 100 heterogeneous robots employing a decentralized standard Bayesian filter and a decentralized mutual information-based gradient-ascent controller, both using a consensus round size of $n_\pi = 5$. *Left to right, top to bottom:* The positions of all robots (green $\circ$) at time steps $k \in \{0, 10, 20, 30, 50, 100, 150, 200\}$.

were uniformly selected from the sensor set used in the indoor hardware experiments from Section 5.4, and the robots were deployed from a single location at the bottom of the environment, e.g., see Figure 6-5. To emulate a physically larger environment for the simulation, no safety radius was used and the network radius was fixed to 1.5 m – half the value used in the experiments. All other parameters remained the same from the corresponding hardware experiments.

Implementing both Algorithm 6 (for a standard Bayesian filter) and Algorithm 9 (for a non-parametric Bayesian filter), we verified that the runtime for the simulation remained constant as more robots were simulated. We then performed Monte-Carlo simulations for a 100 robot system with i) all robots employing the decentralized standard Bayesian filter (see Figure 6-5); ii) all robots employing the decentralized

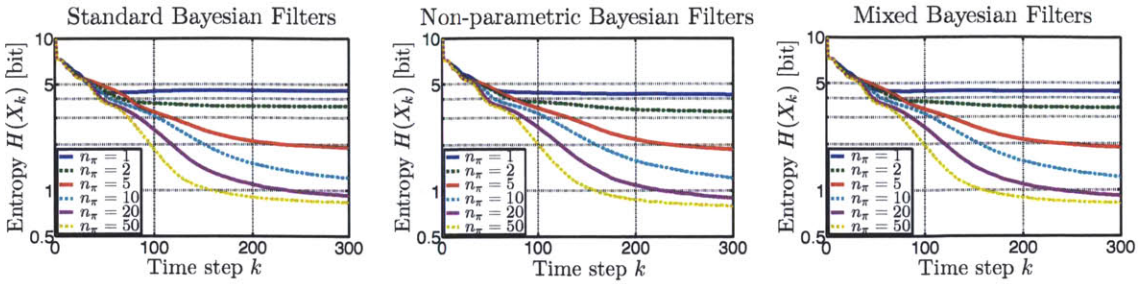


Figure 6-6: Plots from 1000 Monte Carlo simulations showing the average entropy over the 100 robots' beliefs for various communication round sizes. *Left:* All robots employed the decentralized standard Bayesian filter. *Middle:* All robots employed the decentralized non-parametric Bayesian filter. *Right:* The first 50 robots employed the decentralized standard Bayesian filter, while the second 50 employed a non-parametric one.

non-parametric Bayesian filter; and iii) the first 50 (second 50) robots employing the decentralized standard (non-parametric, respectively) Bayesian filter. Figure 6-6 shows the decrease in the average entropy of the robots' beliefs over 1000 Monte-Carlo simulations for each communication round size.

Considering the use of two different types of Bayesian filters (standard versus non-parametric) and the drastic difference in the number of robots (100 versus 5 in the hardware experiments), it is important emphasize that the decentralized mutual information-based gradient-ascent controllers required no modifications to parameters such as controller gain $\gamma^{[i]}$, control space $\mathcal{U}^{[i]}$, and observation sample set size n_y. This validation of generality (former) and scalability (latter) best highlights the practicality of such a control approach. We have also scaled up the simulations to 1000 robots using a single workstation computer equipped with a multicore GPU-enabled graphics card. As shown in Figure 6-7, we get qualitatively similar trajectories when using the same control and inference parameters as in previous simulations.

With respect to the inference, we are not surprised that larger communication round sizes resulted in lower overall uncertainty within the system. In other words, the allocation of more network resources increases filter performance, however, there is clearly evidence of diminishing returns. In addition, the simulations highlight the importance of the network topology; even though many more robots are deployed in

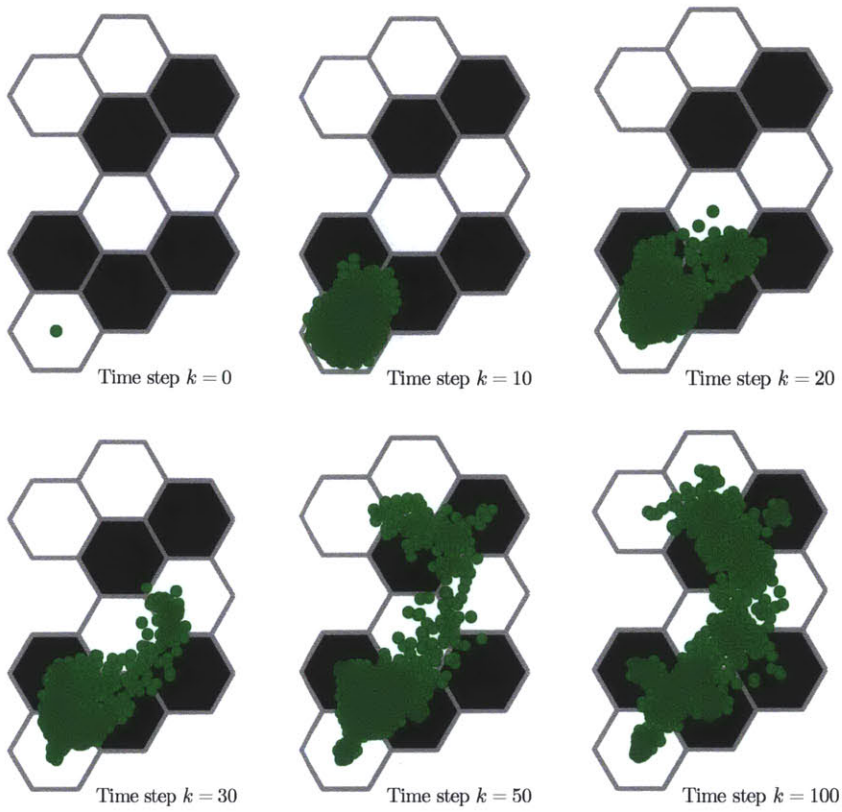


Figure 6-7: The position evolution of a simulated system of 1000 heterogeneous robots employing a decentralized non-parametric Bayesian filter and a decentralized mutual information-based gradient-ascent controller, both using the same parameters as in Figure 6-5. *Left to right, top to bottom:* The positions of all robots (green ∘) at time steps $k \in \{0, 10, 20, 30, 50, 100\}$.

comparison to the hardware experiments, the propagation of information throughout the system is hindered by the sparsity of the network when using small communication round sizes. This result raises interesting questions about fundamental limitations that cannot be overcome by simply deploying more robots.

6.5 Summary

Although theoretic in nature, the distributed algorithms presented in this chapter should inspire application specific approaches to joint measurement likelihood approximation that are unbiased, convergent, and scalable. More importantly, the derivation, analysis, and implementation of the novel class of decentralized mutual

information-based gradient ascent controllers reveals a long list of open research questions. We discuss such future work in Chapter 7, as well as give concluding remarks and lessons learned.

6.6 Proofs

Proof 86 (Lemma 69) *The proof is a simple extension of the proof for Theorem 4.1 in (Lynch, 1997).*

Proof 87 (Lemma 70) *From Lemma 69, we have for all robots and* $\kappa \geq \text{diam}(\mathcal{G})$ *that* $\vartheta_{\kappa}^{[i]}$ *is equal to* $(1+\Delta)$. *Substituting* $(1+\Delta)$ *into* (6.3) *for all* $\vartheta_{\kappa}^{[i]}$ *and* $\vartheta_{\kappa+1}^{[i]}$ *results in a consensus algorithm that is equivalent to* (6.2) *with* $\epsilon = 1/(1+\Delta)$. *In addition, we know for* $\kappa < \text{diam}(\mathcal{G})$ *that the time varying nonlinear system will not have worse than exponential divergence since all coefficients in the right hand side of* (6.3) *are bounded below and above by* 0 *and* 1, *respectively. Thus, since* $\psi_{\kappa}^{[i]}$ *in* (6.2) *was proven to converge to* $\sum_{v=1}^{n_r} \psi_0^{[v]}/n_r$, *we have that* $\psi_{\kappa}^{[i]}$ *in* (6.3) *will do the same if and only if* $\sum_{v=1}^{n_r} \psi_{\kappa}^{[v]}$ *equals* $\sum_{v=1}^{n_r} \psi_0^{[v]}$ *for* $\kappa = \text{diam}(\mathcal{G})$.

For this aim, consider the summation

$$\sum_{i=1}^{n_r} \vartheta_{\kappa+1}^{[i]} \psi_{\kappa+1}^{[i]} = \sum_{i=1}^{n_r} (\vartheta_{\kappa+1}^{[i]} - \vartheta_{\kappa}^{[i]}) \psi_0^{[i]} + \sum_{i=1}^{n_r} \vartheta_{\kappa}^{[i]} \psi_{\kappa}^{[i]} + \sum_{i=1}^{n_r} \sum_{v \in N^{[i]}} (\psi_{\kappa}^{[v]} - \psi_{\kappa}^{[i]})$$

In an undirected graph, the last term on the right hand side of the last equation is equal to zero, and thus we have for all communication rounds that

$$\begin{aligned} \sum_{i=1}^{n_r} \vartheta_{\kappa+1}^{[i]} \psi_{\kappa+1}^{[i]} &= \sum_{i=1}^{n_r} (\vartheta_{\kappa+1}^{[i]} - \vartheta_{\kappa}^{[i]}) \psi_0^{[i]} + \sum_{i=1}^{n_r} (\vartheta_{\kappa}^{[i]} - \vartheta_{\kappa-1}^{[i]}) \psi_0^{[i]} \\ &\quad + \cdots + \sum_{i=1}^{n_r} (\vartheta_1^{[i]} - \vartheta_0^{[i]}) \psi_0^{[i]} + \sum_{i=1}^{n_r} \vartheta_0^{[i]} \psi_0^{[i]} = \sum_{i=1}^{n_r} \vartheta_{\kappa+1}^{[i]} \psi_0^{[i]}, \end{aligned}$$

implying that $\sum_{v=1}^{n_r} \psi_{\kappa}^{[v]}$ *equals* $\sum_{v=1}^{n_r} \psi_0^{[v]}$ *when* $\vartheta_{\kappa+1}^{[i]} = \Delta$. *Thus, we have that* $\psi_{\kappa}^{[i]}$ *in* (6.3) *converges to* $\sum_{v=1}^{n_r} \psi_0^{[v]}/n_r$, *which also implies that* $\pi_{\kappa}^{[i]}$ *in* (6.4) *converges to* $\prod_{v=1}^{n_r} (\pi_0^{[v]})^{1/n_r}$.

Proof 88 (Lemma 71) *We have for all robots that* $\vartheta_0^{[i]} = \vartheta_1^{[i]} = (1+\Delta)$ *for a complete network graph. Thus,* $\psi_1^{[i]}$ *and* $\pi_1^{[i]}$ *are equal to* $\sum_{v=1}^{n_r} \psi_0^{[v]}/n_r$ *and* $\prod_{v=1}^{n_r} \left(\pi_0^{[v]}\right)^{1/n_r}$, *respectively. The proof follows by induction on* κ.

Proof 89 (Theorem 75) *For all robots,* $j \in \{1, \ldots, |\mathcal{X}|\}$, *and* $\ell \in \{1, \ldots, n_y\}$, *let* $[\pi_\kappa^{[i]}]_{j\ell}$ *be initialized to*

$$[\pi_0^{[i]}]_{j\ell} = \mathbb{P}(\check{Y}^{[i]} = \check{y}^{[i,\ell]} | X = x^{[j]})$$

and evolve using (6.4). *From Lemma 70, we have that*

$$[\pi^{[i]}]_{j\ell} \to \prod_{v=1}^{n_r} \mathbb{P}(\check{Y}^{[v]} = \check{y}^{[v,\ell]} | X = x^{[j]})^{1/n_r}$$

and

$$\beta^{[i]} \to n_r$$

in the limit as n_π *tends to infinity. Hence from* (6.5) *we have that*

$$[\pi^{[i]}]_{j\ell} \to \eta\, \mathbb{P}(\check{Y} = \check{y}^{[\ell]} | X = x^{[j]}).$$

Lastly, using the definition of $\check{p}^{[i]}$ *from* (6.6), *we have that*

$$\check{p}^{[i,\ell]}(x^{[j]}) \to \mathbb{P}(\check{Y} = \check{y}^{[\ell]} | X = x^{[j]}).$$

The proof for $p^{[i]}$ *follows in the same manner with* $[\pi_\kappa^{[i]}]_j$ *being initialized to* $\mathbb{P}(Y^{[i]} = y^{[i]} | X = x^{[j]})$ *and converging to* $\mathbb{P}(Y = y | X = x^{[j]})$.

Proof 90 (Corollary 76) *Following the proof for Theorem 75, convergence after one communication round for a complete network graph is a direct consequence of Corollary 71.*

Proof 91 (Proposition 77) *By definition, the observation sample set* $\check{\mathcal{Y}}_k^{[i]}$ *has* n_y *el-*

ements for any robot i, while the environment state alphabet has $|\mathcal{X}|$ elements. Thus, the belief matrix $\pi_\kappa^{[i]}$ has $n_y|\mathcal{X}|$ entries and takes $O(n_y|\mathcal{X}|)$ time to update per communication round. Moreover, since the number of communication rounds n_π is fixed and Assumption 67 bounds the maximum number of neighbors $\Delta^{[i]}$ by a constant integer M, the time and space complexities are independent of the number of robots n_r.

For thoroughness, again note that the consensus state $\phi_0^{[i]}$ is not initialized to a tuple of size n_r but instead an indexed tuple of initial size 1.

Proof 92 (Theorem 78) *We first note for all robots and communication rounds that $\pi_\kappa^{[i]}$ is a product of values taken from unnormalized Gaussians. Hence $(\pi_{\kappa+1}^{[i]})^{\beta_{\kappa+1}^{[i]}}$ is itself a value that is taken from an unnormalized Gaussian proportional to*

$$\prod_{v \in \left\{\{i\} \cup N^{[i]}\right\}} e^{\beta_{\kappa+1}^{[i]}[W]_{iv}(x^T \xi_\kappa^{[v]} - \frac{1}{2} x^T \Omega_\kappa^{[v]} x)}$$

which gives us the desired consensus update expressions for $\xi_{\kappa+1}^{[i]}$ and $\Omega_{\kappa+1}^{[i]}$. Lastly, from (Xiao et al., 2007) we have for every $x \in \mathcal{X}$ that $\pi_\kappa^{[i]}$, $\beta_\kappa^{[i]}$, and $\alpha_\kappa^{[i]}$ converge to $\mathcal{N}_c(\xi, \Omega)^{\frac{1}{n_r}}$, n_r, and $\left(p\eta(\xi, \Omega)\right)^{\frac{1}{n_r}}$, respectively.

Proof 93 (Proposition 79) *The proof follows the proof for Proposition 77, except that i) there exists no for loop of size $|\mathcal{X}|$ and ii) the information matrix $\Omega_\kappa^{[i,\ell]}$ is of size n_g^2 for all samples $\ell \in \{1, \ldots, n_y\}$.*

Proof 94 (Lemma 80) *The proof follows from Xiao et al. (2007) with $\psi_0^{[i]} = e_i$, since*

$$\left\| \psi_0^{[i]} - \mathbf{1}\frac{1}{n_r} \right\|_2^2 = \left(1 - \frac{1}{n_r}\right)^2 + \frac{n_r - 1}{n_r^2} = 1 - \frac{1}{n_r}.$$

Proof 95 (Lemma 81) *Since for all robots* $\|\psi_0^{[i]}\|_1 = \|\boldsymbol{e}_i\|_1 = 1$, *we have that*

$$\begin{aligned}\|\psi_1^{[i]}\|_1 &= [\boldsymbol{W}]_{ii}\|\psi_0^{[i]}\|_1 + \sum_{v\in N^{[i]}} [\boldsymbol{W}]_{iv}\|\psi_0^{[v]}\|_1 \\ &= [\boldsymbol{W}]_{ii} + \sum_{v\in N^{[i]}} [\boldsymbol{W}]_{iv} \\ &= 1 - \sum_{v\in N^{[i]}} [\boldsymbol{W}]_{iv} + \sum_{v\in N^{[i]}} [\boldsymbol{W}]_{iv} \\ &= 1.\end{aligned}$$

In addition, $\psi_1^{[i]}$ *is nonnegative since it is an element-wise summation of nonnegative terms, which from the previous equation implies* $\psi_1^{[i]} \in [0,1]^{n_r}$. *Lastly, we have from the relationship between the* 1*-norm and the* ∞*-norm that*

$$\|\psi_1^{[i]}\|_\infty \geq \frac{\|\psi_1^{[i]}\|_1}{n_r} = \frac{1}{n_r}.$$

The proof follows by induction on κ.

Proof 96 (Proposition 82) *Consider the optimization problem of maximizing* $\|\psi_\kappa^{[i]}\|_\infty$ *with* $\psi_\kappa^{[i]}$ *being a free variable subject to the constraints in Lemma 81 and subject to*

$$\left\|\psi_\kappa^{[i]} - \boldsymbol{1}\frac{1}{n_r}\right\|_2 \leq U_\kappa^\psi \in \left[0, \sqrt{1-\frac{1}{n_r}}\right] = [0, U_0^\psi].$$

Note that an optimal solution ψ_κ^* *always exists. Put* $c \geq 0$, *and without loss of generality suppose* $\|\psi_\kappa^*\|_\infty = [\psi_\kappa^*]_1$ *and* $\|\psi_\kappa^* - \boldsymbol{1}\frac{1}{n_r}\|_2^2 = c^2$. *We define*

$$f(\psi_\kappa^{[i]}, \mu_1, \mu_2) := [\psi_\kappa^{[i]}]_1 + \mu_1\left(\left\|\psi_\kappa^{[i]} - \boldsymbol{1}\frac{1}{n_r}\right\|_2^2 - c^2\right) + \mu_2\big(\|\psi_\kappa^{[i]}\|_1 - 1\big),$$

and by using Lagrange multipliers obtain

$$\frac{1}{n_r - 1}\left([\psi_\kappa^{[i]}]_1 - \frac{1}{n_r}\right)^2 + \left([\psi_\kappa^{[i]}]_1 - \frac{1}{n_r}\right)^2 - c^2 = 0.$$

Thus, we have that

$$[\psi_\kappa^{[i]}]_1 = c\sqrt{1 - \frac{1}{n_r}} + \frac{1}{n_r} \tag{6.12}$$

and

$$c \leq \sqrt{1 - \frac{1}{n_r}} \tag{6.13}$$

since $[\psi_\kappa^{[i]}]_1 \in [0, 1]$. *From* (6.12), *we have that* $[\psi_\kappa^{[i]}]_1$ *is proportional to* c, *and thus by* (6.13) *we conclude that* $c = U_\kappa^\psi$. *Thus, the corresponding value of* $[\psi_\kappa^{[i]}]_1 = \|\psi_\kappa^*\|_\infty$ *is*

$$U_\kappa^\psi \sqrt{1 - \frac{1}{n_r}} + \frac{1}{n_r}.$$

Lastly, consider $\psi_\kappa^{[i]}$ *as consensus term rather than a free variable. From above, we can interpret* $\|\psi_\kappa^*\|_\infty^{-1}$ *as a lower bound for* $\beta_\kappa^{[i]}$ *given* U_κ^ψ, *which gives* L_κ^β.

Proof 97 (Proposition 83) *We first prove that* $\lambda_m(\beta_\kappa^{[i]}\mathbf{\Omega}_\kappa^{[i]})$ *is bounded below by* $L_{\kappa,m}^{\mathbf{\Omega}-}$. *Note that*

$$\beta_\kappa^{[i]}\mathbf{\Omega}_\kappa^{[i]} = \beta_\kappa^{[i]} \sum_{v=1}^{n_r} [\psi_\kappa^{[i]}]_v \mathbf{\Omega}_0^{[v]}.$$

Recursively applying Weyl's Theorem (Horn and Johnson, 1985), we have that

$$\lambda_m(\beta_\kappa^{[i]}\mathbf{\Omega}_\kappa^{[i]}) \geq \beta_\kappa^{[i]} \sum_{v=1}^{n_r} [\psi_\kappa^{[i]}]_v \lambda_1(\mathbf{\Omega}_0^{[v]}). \tag{6.14}$$

Ordering the robot indices for λ_1 *and using the lower bound from Proposition 82, we have that*

$$\beta_\kappa^{[i]} \sum_{v=1}^{n_r} [\psi_\kappa^{[i]}]_v \lambda_1(\mathbf{\Omega}_0^{[v]}) \geq \sum_{\ell=1}^{\lfloor L_\kappa^\beta \rfloor} \lambda_1(\mathbf{\Omega}_0^{[\ell]}) + (L_\kappa^\beta - \lfloor L_\kappa^\beta \rfloor)\lambda_1(\mathbf{\Omega}_0^{[\lceil L_\kappa^\beta \rceil]}).$$

Substituting this inequality into (6.14) *gives* $L_{\kappa,m}^{\mathbf{\Omega}-}$.

Lastly we prove in similar fashion that $\lambda_m(\beta_\kappa^{[i]}\mathbf{\Omega}_\kappa^{[i]})$ *is bounded below by* $L_{\kappa,m}^{\mathbf{\Omega}+}$. *Note that*

$$\beta_\kappa^{[i]}\mathbf{\Omega}_\kappa^{[i]} = \mathbf{\Omega} - \sum_{v=1}^{n_r}(1 - \beta_\kappa^{[i]}[\psi_\kappa^{[i]}]_v)\mathbf{\Omega}_0^{[v]}.$$

Recursively applying Weyl's Theorem, we have that

$$\lambda_m(\mathbf{\Omega}) \leq \lambda_m(\mathbf{\Omega}_\kappa^{[i]}) + \sum_{v=1}^{n_r}(1 - \beta_\kappa^{[i]}[\psi_\kappa^{[i]}]_v)\lambda_{n_g}(\mathbf{\Omega}_0^{[v]}). \tag{6.15}$$

Ordering the robot indices for λ_{n_g} *and using the lower bound from Proposition 82, we have that*

$$\sum_{v=1}^{n_r}(1 - \beta_\kappa^{[i]}[\psi_\kappa^{[i]}]_v)\lambda_{n_g}(\mathbf{\Omega}_0^{[v]}) \leq \left(\lceil L_\kappa^\beta \rceil - L_\kappa^\beta\right)\lambda_{n_g}(\mathbf{\Omega}_0^{[\lceil L_\kappa^\beta \rceil]}) + \sum_{\ell=\lceil L_\kappa^\beta \rceil+1}^{n_r} \lambda_{n_g}(\mathbf{\Omega}_0^{[\ell]}).$$

Subtracting the summation term from both sides of (6.15), *substituting the result into the previous inequality gives* $L_{\kappa,m}^{\mathbf{\Omega}+}$.

Proof 98 (Corollary 85) *The proof follows from applying Proposition 82 to Theorem 78.*

Chapter 7

Conclusions and remarks

In this thesis we present the derivation, analysis, and implementation of a novel class of decentralized controllers that continuously move robots equipped with sensors to better observe their environment. These controllers are designed to maximize a mutual information reward function, which is shown to be relevant to many information acquisition tasks. One such task is occupancy grid mapping, for which we proved that such controllers are eventually attracted to unexplored space – a geometrically intuitive behavior for exploration. More importantly, the insight we obtained in deriving this proof is exemplified by the results of our experiments. This result solidifies the intuition that inherently guides our logic when designing motion planning algorithms, however, we believe that much more can be proven. In particular, we leave as a future work proving (under additional assumptions) that obstacles eventually repel these controllers due to properties such as sensor obscuration, e.g., Figure 4-1. Such a property would result in trajectories that naturally avoid obstacles and, more importantly, give insight into the overall performance of these controllers in unstructured environments.

This thesis also gives insight into the information-theoretic foundation supporting the novel controller class. We showed that these controllers are intractable in their general form, and that principled approximations forfeit optimality to induce computational tractability. We also proved that each robot can distributively approximate the joint measurement likelihoods such that they converge to the true values

as more computational resources are allocated for inference and coordination. With any real system, however, the amount of computational resources available to each robot is always finite and stochastic. To this end, the derivation and implementation of these decentralized controllers and distributed approximation algorithms should work *regardless*[1] of the available resources. This anytime and any-input mentality to decentralized control and inference is essential to making fielded multi-robot systems with autonomous surveillance capabilities a reality.

7.1 Lessons learned

Seemingly with all robotics research, analysis is only useful up to a point. The knowledge I gained through implementing (or at times, failing to implement) our algorithms on real hardware is invaluable for two reasons. The obvious one is to validate the analysis, where the implementation acts as the final step of the proof. Not only is the real world too complex to be perfectly captured by any tractable mathematical model, but the omnipresent probability space should not be insulted by one relying too heavily on a random number generator. The second reason is, as an engineer, I should always strive to "make stuff work." I have lost count of how many times the implementation process has helped me better understand a mathematical derivation and, as a result, the final analysis.

Although I have always tried to not "reinvent the wheel," I have learned to expand my literature search beyond the robotics community. By construction, robotics is a multi-disciplinary field that builds upon many other research fields. I am often tempted to use existing robotics publications as "cheat sheets" representing the vast knowledge-base offered by these other fields – this temptation is dangerous for countably infinite reasons. For one, these "cheat sheets" exist to act as the technical basis for solutions to *solved* robotics problems, and obviously cannot cover all potential solutions to *unsolved* problems. I have found that exploring this unfamiliar territory

[1]For the extreme case when no computational resources are available, we can only hope that these algorithms do not catastrophically fail nor cause catastrophic failures within the robot team.

has not only provided me with great technical growth and a powerful research arsenal, but has also lead to some of my most valued collaborations and friendships.

Lastly, I have learned that the most powerful controllers are both information-based and geometric-based. We implicitly constructed such a mixture for our quadrotor experiments in Section 5.4 by establishing the notion of a safety radius. Without this geometric-based policy, the potential for midair collisions is not captured by the environment state and effectiveness of the decentralized controller is reduced. We can envision incorporating such hazards into the Bayesian inference, e.g., (Dames et al., 2012), but this may not always be the ideal approach. For example, a downward leading stairwell often results in attractive tendencies for mutual information-based controllers (Chapter 4), which in turn leads to epic crashes for most ground robots. Due to the simplicity of the occupancy grid map random variable, there is no inherent way of representing such stairwells. Thus, we can either i) add complexity to the map and lose much of the computational benefits the occupancy grid mapping algorithm offers, or ii) penalize the mutual information reward function by a geometric-based cost. From an implementation standpoint, I certainly prefer the latter.

7.2 Future work

The submission of this thesis is my first step towards addressing a long list of open research questions revealed to me over the extent of my graduate career.

Although the environment state is continuous-time in general, the digital inference calculations occur at discrete-time instances, which is the motivation behind representing the environment state as a discrete-time random variable with samples taken at constant time intervals. In order to properly represent the temporal evolution of the environment state and avoid aliasing, the Nyquist-Shannon sampling theorem states that the sampling frequency be at least twice the highest temporal frequency component of interest. Analogous to discretizing the environment state in the temporal domain, the spatial domain is sampled in a similar manner to capture frequency components in all dimensions.

I am interested in formalizing the procedure of spatially and temporally partitioning the environment state for multi-robot information acquisition tasks. More importantly, we wish to best represent an environment in an adaptive way based on the information sensed given limited computational resources. We believe information-theoretic compression techniques will be an invaluable tool in not only computationally processing and wirelessly communicating state information, but also to capture the frequency components necessary for good inference and control performance, especially in dynamic environments.

Given the environment representation (or representation methodology if adaptive), we believe information-theoretic bounds naturally arise as a function of system resources. I hope future works provide upper bounds on the quality of information acquired before a system is deployed, especially concerning the number of robots employed. For example, if a systems designer wishes to track up to five separate cars of known maximum velocity in a city area, I want to provide the necessary number of robots to successful perform the task. I am also interested in how accurate our sample-based non-parametric approximations are from an information-theoretic perspective, something Charrow et al. (2013) have begun investigating in target tracking applications using Gaussian mixture models.

We have seen much evidence in our numerical simulations and hardware experiments that given enough system resources (e.g., sensor, computational, network), the robots tend to 'converge' to a configuration that mimics sensor coverage in geometric-based control. This phenomenon makes sense – given that the robots can reduce uncertainty in a way that dominates the rate at which the environment produces uncertainty, the optimization surface on which the mutual information gradient is calculated does not significantly change after 'convergence between observations.'

Thus, I am inspired to go beyond the idea of 'convergence between observations' due to recent works on stochastic stability analysis over a finite-time horizons. Steinhardt and Tedrake (2011) provided bounds on the probability of motion planning failure using a variant of the classical supermartingale result, which falls under the framework of stochastic verification. Although I am interested in verification results

(e.g., probability that robots explore a given percent of an environment within a given time), my true interest lies in probabilistic analogs to geometric steady state behaviors, such as flocking and coverage (Bullo et al., 2009).

I am also interested in periodic behavior with finite periods, which have been shown in linear Gaussian processes to be independent of initial states (Zhang et al., 2010). I believe periodic behavior is possible in expectation if the robots and environment are somehow matched in their uncertainty reduction / production, leading to cyclic dynamics of the same optimization manifold. Given a relatively large amount of uncertainty production by the environment, we hypothesize that robot motion does not converge in expectation. This concept relates back to the bounds on system performance discussed above.

With respect to the influence of time scales on latent networks, researchers have made considerable progress on formalizing the interplay between decentralized controllers and mesh protocols. Schwager et al. (2011c) showed that systems made up of robots with stable first order dynamics are stable for all network update times, positive feedback gains, and connected communication graphs. Although our robot model employs first order integrator dynamics, we again are interested in the expected behavior of the robots over stochastic events. From this viewpoint, the overall control strategy introduces nonlinear characteristics over numerous time steps and thus may lead to 'divergence in expectation.' I believe the work of Smith and Hadaegh (2007) on formation control with consensus-based state estimation algorithms will provide a starting point working within our probabilistic framework. In addition, results from research considering bounded (Olfati-Saber and Murray, 2004; Blondel et al., 2005) and random network delays (Xiao et al., 2000) may be instrumental in the investigation of stability.

Bibliography

M. Alighanbari and J. P. How. An unbiased Kalman consensus algorithm. In *Proc. Amer. Control Conf.*, Minneapolis, MN, USA, Jun 2006.

F. Amigoni and V. Caglioti. An information-based exploration strategy for environment mapping with mobile robots. *Robotics and Autonomous Syst.*, 58(5):684–699, May 2010.

R. Aragués, J. Cortés, and C. Sagués. Dynamic consensus for merging visual maps under limited communications. In *Proc. IEEE Int. Conference on Robotics and Automation*, Anchorage, AK, USA, May 2010.

S. Bhattacharya, N. Michael, and V. Kumar. Distributed coverage and exploration in unknown non-convex environments. In *Proc. Int. Symp. on Distributed Autonomous Systems*, Lausanne, Switzerland, Nov 2010.

V. D. Blondel, J. M. Hendrickx, A. Olshevsky, and J. N. Tsitsiklis. Convergence in multiagent coordination, consensus and flocking. In *Proc. IEEE Conference on Decision and Control*, Seville, Spain, Dec 2005.

J. Borenstein, L. Feng, and H. R. Everett. *Navigating Mobile Robots: Systems and Techniques.* Data Management Systems Series. Morgan Kaufmann Publishers, Inc., 1997.

F. Bourgault, A. Makarenko, S. B. Williams, B. Grocholsky, and H. F. Durrant-Whyte. Information based adaptive robotic exploration. In *Proc. IEEE/RSJ Int. Conference on Intelligent Robots and Syst.*, Lausanne, Switzerland, Sep 2002.

F. Bullo, J. Cortés, and S. Martínez. *Distributed Control of Robotic Networks*. Applied Mathematical Series. Princeton University Press, 2009.

W. Burgard, D. Fox, and S. Thrun. Active mobile robot localization by entropy minimization. In *EUROMICRO Workshop on Advanced Mobile Robots*, Brescia, Italy, Oct 1997.

W. Burgard, M. Moors, D. Fox, R. Simmons, and S. Thrun. Collaborative multi-robot exploration. In *Proc. IEEE Int. Conference on Robotics and Automation*, San Francisco, CA, USA, Apr 2000.

A. Cameron and H. Durrant-Whyte. A bayesian approach to optimal sensor placement. *The Int. J. of Robotics Research*, 9(5):70–88, Oct 1990.

A. R. Cassandra, L. P. Kaelbling, and J. A. Kurien. Acting under uncertainty: discrete Bayesian models for mobile-robot navigation. In *Proc. IEEE/RSJ Int. Conference on Intelligent Robots and Syst.*, Osaka, Japan, Nov 1996.

D. A. Castanon. Optimal search strategies in dynamic hypothesis testing. In *Proc. IEEE Conference on Decision and Control*, San Antonio, TX, USA, Dec 1993.

B. Charrow, V. Kumar, and N. Michael. Approximate representations for multi-robot control policies that maximize mutual information. In *Proc. Robotics: Sci. and Syst. Conference*, Berlin, Germany, 2013 2013.

R. Chatila and J. Laumond. Position referencing and consistent world modeling for mobile robots. In *Proc. IEEE Int. Conference on Robotics and Automation*, St. Louis, MO, USA, Mar 1985.

H. L. Choi and J. P. How. Continuous trajectory planning of mobile sensors for informative forecasting. *Automatica*, 46(8):1266–1275, Aug 2010.

H. I. Christensen. A roadmap for US robotics: from internet to robotics. Technical report, Computing Community Consortium, Mar 2013.

T. H. Chung, V. Gupta, J. W. Burdick, and R. M. Murray. On a decentralized active sensing strategy using mobile sensor platforms in a network. In *Proc. IEEE Conference on Decision and Control*, Atlantis, Paradise Island, Bahamas, Dec 2004.

J. Cortés. Distributed algorithms for reaching consensus on general functions. *Automatica*, 44(3):727–737, Mar 2008.

J. Cortés. Distributed Kriged Kalman filter for spatial estimation. *IEEE Trans. on Automatic Control*, 54(12):2816–2827, Dec 2009.

J. Cortés, S. Martínez, T. Karatas, and F. Bullo. Coverage control for mobile sensing networks. *IEEE Trans. on Robotics*, 20(2):243–255, Apr 2004.

T. M. Cover and J. A. Thomas. *Elements of Information Theory.* John Wiley, New York, 1991.

P. Dames and V. Kumar. Cooperative multi-target localization with noisy sensors. In *Proc. IEEE Int. Conference on Robotics and Automation*, Karlsruhe, Germany, May 2013.

P. Dames, M. Schwager, V. Kumar, and D. Rus. A decentralized control policy for adaptive information gathering in hazardous environments. In *Proc. IEEE Conference on Decision and Control*, Grand Wailea, HI, USA, Dec 2012.

W. J. deGroot, T. J. Lynham, M. A. Brady, I. A. Csiszar, D. Davies, C. O. Justice, E. M. Prins, and J. G. Goldammer. Global Observation of Forest and Land Cover Dynamics (GOFC-GOLD): monitoring and early warning systems for wildland fire disaster reduction. In *The Full Picture*, pages 164–167. Group on Earth Observations (GEO), 2007.

H. F. Durrant-Whyte, B. Y. S. Rao, and H. Hu. Toward a fully decentralized architecture for multi-sensor data-fusion. In *Proc. IEEE Int. Conference on Robotics and Automation*, Cincinnati, OH, USA, May 1990.

A. Elfes. Sonar-based real-world mapping and navigation. *IEEE J. of Robotics and Automation*, 3(3):249–265, Jun 1987.

A. Elfes. Robot navigation: integrating perception, environmental constraints and task execution within a probabilistic framework. In *Reasoning with Uncertainty in Robotics*, Amsterdam, The Netherlands, Dec 1995.

S. P. Engelson and D. V. McDermott. Error correction in mobile robot map learning. In *Proc. IEEE Int. Conference on Robotics and Automation*, Nice, France, May 1992.

D. Fox, W. Burgard, H. Kruppa, and S. Thrun. A probabilistic approach to collaborative multi-robot localization. *Autonomous Robots*, 8(3):325–344, Jun 2000.

C. S. R. Fraser, L. F. Bertuccelli, H. L. Choi, and J. P. How. A hyperparameter consensus method for agreement under uncertainty. *Automatica*, 48(2):374–380, Feb 2012.

M. Frassl, M. Lichtenstern, M. Khider, and M. Angermann. Developing a system for information management in disaster relief - methodology and requirements. In *Proc. Int. Community on Inform. Syst. for Crisis Response and Manage. Conference*, Seattle, WA, USA, May 2010.

G. Grisetti, C. Stachniss, and W. Burgard. Improved techniques for grip mapping with Rao-Blackwellized particle filters. *IEEE Trans. on Robotics*, 23(1):34–46, Feb 2007.

B. Grocholsky. *Information-theoretic control of multiple sensor platforms.* PhD thesis, University of Sydney, 2002.

B. Grocholsky, A. Makarenko, and H. F. Durrant-Whyte. Information-theoretic control of multiple sensor platforms. In *Proc. IEEE Int. Conference on Robotics and Automation*, Taipei, Taiwan, Sep 2003.

J. A. Gubner. Distributed estimation and quantization. *IEEE Trans. on Inform. Theory*, 39(4):1456–1459, Jul 1993.

D. Gurdan, J. Stumpf, M. Achtelik, K. M. Doth, G. Hirzinger, and D. Rus. Energy-efficient autonomous four-rotor flying robot controlled at 1 khz. In *Proc. IEEE Int. Conference on Robotics and Automation*, Roma, Italy, Apr 2007.

P. R. Halmos. *Measure Theory.* Springer-Verlag, 1974.

M. W. Hirsche and S. Smale. *Differential Equations, Dynamical Systems, and Linear Algebra.* Applied Mathematical Series. Academic Press, Inc., 1974.

G. M. Hoffmann and C. J. Tomlin. Mobile sensor network control using mutual information methods and particle filters. *IEEE Trans. on Automatic Control*, 55 (1):32–47, Jan 2010.

R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge University Press, 1985.

A. Howard. Multi-robot simultaneous localization and mapping using particle filters. *The Int. J. of Robotics Research*, 25(12):1243–1256, Dec 2006.

A. T. Ihler, J. W. Fischer III, R. L. Moses, and A. S. Willsky. Nonparametric belief propagation for self-localization of sensor networks. *IEEE J. on Selected Areas in Commun.*, 23(4):809–819, Apr 2005.

A. E. Johnson and S. B. Kang. Registration and integration of textured 3D data. *Image and Vision Computing*, 17(2):135–147, Feb 1999.

B. J. Julian, M. Schwager, M. Angermann, and D. Rus. A location-based algorithm for multi-hopping state estimates within a distributed robot team. In *Proc. Field and Service Robotics*, Cambridge, MA, USA, Jul 2009.

B. J. Julian, M. Angermann, M. Frassl, M. Lichtenstern, and D. Rus. Towards a unifying information theoretic framework for multi-robot exploration and surveillance. In *RSS Workshop on 3D Exploration, Mapping, and Surveillance with Aerial Robots*, Los Angeles, CA, USA, Jun 2011a.

B. J. Julian, M. Angermann, M. Schwager, and D. Rus. A scalable information theoretic approach to distributed robot coordination. In *Proc. IEEE/RSJ Int. Conference on Intelligent Robots and Syst.*, San Francisco, CA, USA, Sep 2011b.

B. J. Julian, M. Angermann, and D. Rus. Nonparametric inference and coordination for distributed robotics. In *Proc. IEEE Conference on Decision and Control*, Grand Wailea, HI, USA, Dec 2012a.

B. J. Julian, M. Angermann, M. Schwager, and D. Rus. Distributed robotic sensor networks: an information-theoretic approach. *The Int. J. of Robotics Research*, 31 (10):1134–1154, Sep 2012b.

B. J. Julian, S. L. Smith, and D. Rus. Distributed approximation of joint measurement distributions using mixtures of Gaussians. In *Proc. Robotics: Sci. and Syst. Conference*, Sydney, Australia, Jul 2012c.

B. J. Julian, S. Karaman, and D. Rus. On mutual information-based control of range sensing robots for mapping applications. *The Int. J. of Robotics Research*, 2013a. Submitted.

B. J. Julian, S. Karaman, and D. Rus. On mutual information-based control of range sensing robots for mapping applications. In *Proc. IEEE/RSJ Int. Conference on Intelligent Robots and Syst.*, Tokyo, Japan, Nov 2013b. Accepted.

C. O. Justice and S. Korontzi. A review of the status of satellite fire monitoring and the requirements for global environmental change research. In *Global and Regional Vegitation Fire Monitoring From Space: Planning a Coordinate International Effort.* SPB Academic Publishing, 2001.

V. Kelha, Y. Rauste, T. Hame, T. Sephton, O. Frauenberger, K. Soini, A. Venalainen, J. San-Miguel-Ayanz, and T. Vainio. Promotion of space technologies for supporting the management of natural disasters: Earth observation technologies for decision support, 'fire risk management in Finland'. Technical report, VTT Automation, 2001.

A. Krause, A. Singh, and C. Guestrin. Near-optimal sensor placements in Gaussian processes: theory, efficient algorithms and empirical studies. *J. of Mach. Learning Research*, 9:235–284, Jun 2008.

H. Kretzschmar and C. Stachniss. Information-theoretic compression of pose graphs for laser-based SLAM. *The Int. J. of Robotics Research*, 31(11):1219–1230, Aug 2012.

B. J. A. Kröse, N. Vlassis, R. Bunschoten, and Y. Motomura. A probabilistic model for appearance-based robot localization. *Image and Vision Computing*, 19(6):381 – 391, Apr 2001.

B. Kuipers and Y. T. Byun. A robot exploration and mapping strategy based on a semantic hierarchy of spatial representations. *J. of Robotics and Autonomous Systems*, 8:47–63, 1991.

O. Kung, C. Strecha, A. Beyeler, J. C. Zufferey, D. Floreano, P. Fua, and F. Gervaiz. The accuracy of automatic photogrammetic techniques on ultra-light uav imagery. In *Proc. Int. Conference on Unmanned Aerial Vehicles in Geomatics*, Zurich, Switzerland, Sep 2011.

J. LaSalle. Some extensions of lyapunov's second method. *IRE Trans. on Circuit Theory*, 7(4):520–527, Dec 1960.

S. M. LaValle. *Planning Algorithms*. Cambridge University Press, Cambridge, U.K., 2006.

J. Leonard and H. F. Durrant-Whyte. Simultaneous map building and localization for an autonomous mobile robot. In *Proc. IEEE/RSJ Int. Conference on Intelligent Robots and Syst.*, Osaka, Japan, Nov 1991.

K. Leung, T. Barfoot, and H. Liu. Decentralized cooperative SLAM for sparsely-communicating robot networks: a centralized-equivalent approach. *j. of Artificial Intell. Research*, 66(3):321–342, May 2012.

K. M. Lynch, I. B. Schwartz, P. Yang, and R. A. Freeman. Decentralized environmental modeling by mobile sensor networks. *IEEE Trans. on Robotics*, 24(3):710–724, Jun 2008.

N. A. Lynch. *Distributed Algorithms.* Data Management Systems Series. Morgan Kaufmann Publishers, Inc., 1997.

A. Manyika and H. F. Durrant-Whyte. *Data Fusion and Sensor Management: A Decentralized Information-Theoretic Approach.* Prentice Hall, 1994.

N. Michael, M. Schwager, V. Kumar, and D. Rus. An experimental study of time scales and stability in networked multi-robot systems. In *Proc. Int. Symp. on Experimental Robotics*, Delhi, India, Dec 2010.

M. Montemerlo, S. Thrun, D. Koller, and B. Wegbreit. FastSLAM: a factored solution to the simultaneous localization and mapping problem. In *Proc. Nat. Conference on Artificial Intell.*, Edmonton, Alberta, Canada, Jul 2002.

S. J. Moorehead, R. Simmons, and W. L. Whittaker. Autonomous exploration using multiple sources of information. In *Proc. IEEE Int. Conference on Robotics and Automation*, Seoul, Korea, May 2001.

H. Moravec and A. Elfes. High resolution maps from wide angle sonar. In *Proc. IEEE Int. Conference on Robotics and Automation*, St. Louis, MO, USA, Mar 1985.

H. P. Moravec and M. C. Martin. Robot navigation by 3D spatial evidence grids. 1994.

P. Moutarlier and R. Chatila. Stochastic multisensory data fusion for mobile robot location and environment modeling. In *Proc. Int. Symp. on Robotics Research*, Tokyo, Japan, Aug 1989.

E. W. Nettleton, H. F. Durrant-Whyte, P. W. Gibbens, and A. H. Goektogan. Multiple platform localisation and map building. In *Proc. Soc. of Photo-Optical Instrumentation Engineers*, Oct 2000.

J. L. Ny and G. J. Pappas. On trajectory optimization for active sensing in Gaussian process models. In *Proc. IEEE Conference on Decision and Control*, Shanghai, China, Dec 2009.

R. Olfati-Saber and R. M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Trans. on Automatic Control*, 49(9): 1520–1533, Sep 2004.

R. Olfati-Saber, E. Franco, E. Frazzoli, and J. S. Shamma. Belief consensus and distributed hypothesis testing in sensor networks. In *Proc. Network Embedded Sensing and Control Workshop*, South Bend, IN, USA, Oct 2005.

D. P. Palomar and Sergio Verdú. Representation of mutual information via input estimates. *IEEE Trans. on Inform. Theory*, 53(2):453–470, Feb 2007.

J. Pearl. *Probabilistic Reasoning in Intelligent Systems : Networks of Plausible Inference.* Morgan Kaufmann Publishers, Inc., 1988.

M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng. ROS: an open-source robot operating system. In *ICRA Workshop on Open Source Software*, Kobe, Japan, May 2009.

Y. Rauste, A. J. Sephton, V. Kelha, T. Vainio, M. Heikinheimo, K. Soini, and O. Frauenberger. Requirements and analysis report, rar. Technical report, ESA Forest Fire Operational Study AO/1-3468/98/I-DC, 1999.

W. Ren, R. W. Beard, and D. B. Kingston. Multi-agent Kalman consensus with relative uncertainty. In *Proc. Amer. Control Conf.*, Atlanta, GA, USA, Jun 2005.

R. Rocha, J. Dias, and A. Carvalho. Cooperative multi-robot systems: a study of vision-based 3D mapping using information theory. *Robotics and Autonomous Syst.*, 53(3-4):282–311, Nov 2005.

N. Roy, W. Burgard, D. Fox, and S. Thrun. Costal navigation - mobile robot navigation with uncertainty in dynamic environments. In *Proc. IEEE/RSJ Int. Conference on Intelligent Robots and Syst.*, Kyongju, Korea, Oct 1999.

S. J. Russell, P. Norvig, J. F. Canny, J. M. Malik, and D. D. Edwards. *Artificial Intelligence: A Modern Approach.* Prentice Hall, 1995.

A. Ryan and J. K. Hedrick. Particle filter based information-theoretic active sensing. *Robotics and Autonomous Syst.*, 58(5):574–584, May 2010.

J. San-Miguel-Ayanz, N. Ravail, V. Kelha, and A. Ollero. Active fire detection for fire emergency management: potential and limitations for the operational use of remote sensing. *Natural Hazards*, 35(3):361–376, Jul 2005.

D. Schulz, W. Burgard, D. Fox, and A. B. Cremers. Tracking multiple moving targets with a mobile robot using particle filters and statistical data association. In *Proc. IEEE Int. Conference on Robotics and Automation*, Seoul, Korea, May 2001.

M. Schwager. *A gradient optimization approach to adaptive multi-robot control.* PhD thesis, Massachusetts Institute of Technology, 2009.

M. Schwager, B. J. Julian, and D. Rus. Optimal coverage for multiple hovering robots with downward facing cameras. In *Proc. IEEE Int. Conference on Robotics and Automation*, Kobe, Japan, May 2009.

M. Schwager, P. Dames, D. Rus, and V. Kumar. A multi-robot control policy for information gathering in the presence of unknown hazards. In *Proc. Int. Symp. on Robotics Research*, Flagstaff, AZ, USA, Aug 2011a.

M. Schwager, B. J. Julian, M. Angermann, and D. Rus. Eyes in the sky: decentralized control for the deployment of robotic camera networks. *Proc. of the IEEE*, 99(9): 1541–1561, July 2011b.

M. Schwager, N. Michael, V. Kumar, and D. Rus. Time scales and stability in networked multi-robot systems. In *Proc. IEEE Int. Conference on Robotics and Automation*, Shanghai, China, May 2011c.

R. Simmons, D. Apfelbaum, W. Burgard, D. Fox, M. Moors, S. Thrun, and H. Younes. Coordination for multi-robot exploration and mapping. In *Proc. Nat. Conference on Artificial Intell.*, Austin, TX, USA, Jul 2000.

A. Singh, A. Krause, C. Guestrin, W. Kaiser, and M. Batalin. Efficient planning of informative paths for multiple robots. In *Proc. Int. Joint Conference on Artificial Intell.*, Hyderabad, India, Jan 2007.

A. Singh, A. Krause, C. Guestrin, and W. J. Kaiser. Efficient informative sensing using multiple robots. *J. of Artificial Intell. Research*, 34(1):707–755, Jan 2009.

R. Smith and F. Y. Hadaegh. Closed-loop dynamics of cooperative vehicle formations with parallel estimators and communication. *IEEE Trans. on Automatic Control*, 52(8):1404–1414, Aug 2007.

R. Smith, M. Self, and P. Cheeseman. Estimating uncertain spatial relationships in robotics. In *Autonomous Robot Vehicles*, pages 167–193. Springer-Verlag, 1990.

R. C. Smith and P. Cheeseman. On the representation and estimation of spatial uncertainty. *The Int. J. of Robotics Research*, 5(4):56–68, Dec 1986.

C. Stachniss. *Exploration and mapping with mobile robots.* PhD thesis, University of Freiburg, 2006.

J. Steinhardt and R. Tedrake. Finite-time regional verification of stochastic nonlinear systems. In *Proc. Robotics: Sci. and Syst. Conference*, Los Angeles, CA, USA, Jun 2011.

E. Stump, V. Kumar, B. Grocholsky, and P. M. Shiroma. Control for localization of targets using range-only sensors. *The Int. J. of Robotics Research*, 28(6):764–757, May 2009.

S. Thrun and Y. Liu. Multi-robot SLAM with sparse extended information filters. In *Robotics Research*, volume 15 of *Springer Tracts in Advanced Robotics*, pages 254–266. Springer Berlin / Heidelberg, 2005.

S. Thrun, A. Bücken, W. Burgard, D. Fox, T. Fröhlinghaus, D. Hennig, T. Hofmann, M. Krell, and T. Schimdt. Map learning and high-speed navigation in rhino. In *AI-based Mobile Robots: Case studies of successful robot systems.* MIT Press, 1998a.

S. Thrun, W. Burgard, and D. Fox. A probabilistic approach to concurrent mapping and localization for mobile robots. *Autonomous Robots*, 5(3-4):253–271, Jul 1998b.

S. Thrun, J. S. Gutmann, D. Fox, W. Burgard, and B. Kuipers. Integrating topological and metric maps for mobile robot navigation: a statistical approach. In *Proc. Nat. Conference on Artificial Intell.*, Madison, WI, USA, Jul 1998c.

S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics.* Intelligent Robotics and Autonomous Agents Series. The MIT Press, 2005.

P. A. Viola. *Alignment by maximization of mutual information.* PhD thesis, Massachusetts Institute of Technology, 1995.

A. Visser and B. A. Slamet. Including communication success in the estimation of information gain for multi-robot exploration. In *Proc. Int. Symp. on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks*, Berlin, Germany, Mar 2008.

P. Whaite and F. P. Ferrie. Autonomous exploration: driven by uncertainty. *IEEE Trans. on Pattern Anal. and Mach. Intell.*, 19(3):193–205, Mar 1997.

L. Xiao, A. Hassibi, and J. P. How. Control with random communication delays via a discrete-time jump system approach. In *Proc. Amer. Control Conf.*, Chicago, IL, USA, Jun 2000.

L. Xiao, S. Boyd, and S. Lall. A scheme for robust distributed sensor fusion based on average consensus. In *Symp. on Inform. Process. of Sensor Networks*, Los Angeles, CA, USA, Apr 2005.

L. Xiao, S. Boyd, and S. J. Kim. Distributed average consensus with least-mean-square deviation. *J. of Parallel and Distributed Computing*, 67(1):33–46, Jan 2007.

B. Yamauchi. Frontier-based exploration using multiple robots. In *Proc. Int. Conference on Autonomous Agents*, St. Paul, MN, USA, May 1998.

P. Yang, R A. Freeman, and K. M. Lynch. Distributed cooperative active sensing using consensus filters. In *Proc. IEEE Int. Conference on Robotics and Automation*, Roma, Italy, Apr 2007.

M. Yguel, O. Aycard, and C. Laugier. Update policy of dense maps: efficient algorithms and sparse representation. In *Proc. Field and Service Robotics*, Chamonix, France, Jul 2007.

M. M. Zavlanos and G. J. Pappas. Distributed connectivity control of mobile networks. *IEEE Trans. on Robotics*, 24(6):1416–1428, Dec 2008.

W. Zhang, M. Vitus, J. Hu, A. Abate, and C. J. Tomlin. On the optimal solution of the infinite-horizon linear sensor scheduling problem. In *Proc. IEEE Conference on Decision and Control*, Atlanta, GA, USA, Dec 2010.