Decentralized Active Robotic Exploration and Mapping for Probabilistic Field Classification in Environmental Sensing

Kian Hsiang Low[†], Jie Chen[†], John M. Dolan[§], Steve Chien[‡], and David R. Thompson[‡]
Department of Computer Science, National University of Singapore, Republic of Singapore[†]
Robotics Institute, Carnegie Mellon University, Pittsburgh PA 15213 USA[§]
Jet Propulsion Laboratory, California Institute of Technology, Pasadena CA 91109 USA[‡]
{lowkh, chenjie}@comp.nus.edu.sg[†], jmd@cs.cmu.edu[§]
{steve.chien, david.r.thompson}@jpl.nasa.gov[‡]

ABSTRACT

A central problem in environmental sensing and monitoring is to classify/label the hotspots in a large-scale environmental field. This paper presents a novel *decentralized active robotic exploration* (DARE) strategy for probabilistic classification/labeling of hotspots in a *Gaussian process* (GP)-based field. In contrast to existing state-of-the-art exploration strategies for learning environmental field maps, the time needed to solve the DARE strategy is independent of the map resolution and the number of robots, thus making it practical for *in situ*, real-time active sampling. Its exploration behavior exhibits an interesting formal trade-off between that of boundary tracking until the hotspot region boundary can be accurately predicted and wide-area coverage to find new boundaries in sparsely sampled areas to be tracked. We provide a theoretical guarantee on the active exploration performance of the DARE strategy: under reasonable conditional independence assumption, we prove that it can optimally achieve two formal cost-minimizing exploration objectives based on the misclassification and entropy criteria. Importantly, this result implies that the uncertainty of labeling the hotspots in a GP-based field is greatest at or close to the hotspot region boundaries. Empirical evaluation on real-world plankton density and temperature field data shows that, subject to limited observations, DARE strategy can achieve more superior classification of hotspots and time efficiency than state-of-the-art active exploration strategies.

Categories and Subject Descriptors

G.3 [**Probability and Statistics**]: Stochastic processes; I.2.9 [**Robotics**]: Autonomous vehicles

General Terms

Algorithms, Performance, Experimentation, Theory

Keywords

Multi-robot exploration and mapping, adaptive sampling, active learning, Gaussian process

1. INTRODUCTION

A fundamental problem in environmental sensing and monitoring is to identify and delineate the hotspot regions in a large-scale environmental field [2, 11]. It involves partitioning the area spanned by the field into one class of regions called the hotspot regions in which the field measurements exceed a predefined threshold, and the other class of regions where they do not. Such a problem arises in many real-world applications such as precision agriculture, monitoring of ocean and freshwater phenomena (e.g., plankton bloom), forest ecosystems, rare species, pollution (e.g., oil spill), or contamination (e.g., radiation leak). In these applications, it is necessary to assess the spatial extent and shape of the hotspot regions accurately due to severe economic, environmental, and health implications, as discussed in [11]. In practice, this aim is non-trivial to achieve because the constraints on the sampling assets' resources (e.g., energy consumption, mission time, sensing range) limit the number and coverage of *in situ* observations over the large field that can be used to infer the hotspot regions. Subject to limited observations, the most informative ones should therefore be selected in order to minimize the uncertainty of estimating the hotspot regions (or, equivalently, classifying/labeling the hotspots) in the large field, which motivates our adaptive sampling work in this paper.

Mobile robot teams are particularly desirable for performing the above environmental sensing task because they can actively explore to map the hotspot regions at high resolution. On the other hand, static sensors lack mobility and are therefore not capable of doing this well unless a large quantity is deployed. While research in multi-robot exploration and mapping have largely focused on the conventional task of building occupancy grids [10], some recent efforts are put into the more complex, general task of sampling spatially distributed environmental fields [4, 5, 6]. In contrast to occupancy grids that assume *discrete, independent* cell occupancies, environmental fields are characterized by *continuous-valued, spatially correlated* measurements, properties of which cannot be exploited by occupancy grid mapping strategies to select the most informative observation paths. To exploit such properties, exploration strategies for learning environmental field maps have recently been developed and can be classified into two regimes: (a) *wide-area coverage* strategies [3, 4, 5] consider sparsely sampled (i.e., largely unexplored) areas to be of high uncertainty and consequently spread observations evenly across the field; (b) *hotspot sampling* strategies [7] assume areas of high uncertainty and interest to contain extreme, highly-varying measurements and hence produce clustered observations. Formal, principled approaches of exploration [4, 5] have also

Appears in: *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*, Conitzer, Winikoff, Padgham, and van der Hoek (eds.), 4-8 June 2012, Valencia, Spain.

been devised to simultaneously perform hotspot sampling when a hotspot region is found as well as wide-area coverage to search for new hotspot regions in sparsely sampled areas. These strategies optimize their observation paths to minimize the uncertainty (e.g., in terms of mean-squared error or entropy) of mapping the entire continuous-valued field. They are, however, suboptimal for classifying/labeling the hotspots in the field, which we will discuss and demonstrate theoretically and empirically in this paper. More details of their exploration behavior and properties will be provided in Section 6.1.

This paper proposes a novel *decentralized active robotic exploration* (DARE) strategy for probabilistic classification/ labeling of hotspots in a large-scale environmental field (Section 5). The environmental field is assumed to be realized from a rich class of probabilistic spatial models called *Gaussian process* (GP) (Section 2) that can formally characterize its spatial correlation structure. More importantly, it can provide formal measures of classification/labeling uncertainty (i.e., in the form of cost functions) such as the misclassification and entropy criteria (Section 3) for directing the robots to explore highly uncertain areas of the field. The chief impediment to using these formal criteria is that they result in cost-minimizing exploration strategies (Section 4), which cannot be solved in closed form. To resolve this, they are reformulated as reward-maximizing dual strategies, from which we can then derive the approximate DARE strategy to be solved in closed form efficiently. The specific contributions of our work include:

- analyzing the time complexity of solving the DARE strategy (Section 5): we prove that its incurred time is independent of the map resolution and the number of robots, thus making it practical for *in situ*, real-time active sampling. In contrast, existing state-of-the-art exploration strategies [3, 4, 5] for learning environmental field maps scale poorly with increasing map resolution and/or number of robots (Section 6.1);
- analyzing the exploration behavior of the DARE strategy through its formulation (Section 5): it exhibits an interesting formal trade-off between that of boundary tracking until the hotspot region boundary can be accurately predicted and wide-area coverage to find new boundaries in sparsely sampled areas to be tracked. In contrast, ad hoc, reactive boundary tracking strategies [9, 12] typically require a hotspot region boundary to be located manually or via random exploration and are not driven by the need to maximize the fidelity of estimating multiple hotspot regions given limited observations;
- providing theoretical guarantee on the active exploration performance of the DARE strategy (Section 5): we prove that, under reasonable conditional independence assumption, it produces the same optimal observation paths as that of the centralized cost-minimizing strategies, the latter of which otherwise cannot be solved in closed form. This result has a simple but important implication: the uncertainty of labeling the hotspots in a GP-based field is greatest at or close to the hotspot region boundaries;
- empirically evaluating the active exploration performance and time efficiency of the DARE strategy on real-world plankton density and temperature field data (Section 6): subject to limited observations, the DARE strategy can achieve better classification of the hotspots than state-of-the-art active exploration strategies [1, 5] while being significantly more time-efficient than those performing wide-area coverage and hotspot sampling.

2. GAUSSIAN PROCESS-BASED ENVIRONMENTAL FIELD

The *Gaussian process* (GP) can be used to model an environmental field as follows: the environmental field is defined to vary as a realization of a GP. Let $\mathcal{X}$ be a set of sampling locations representing the domain of the environmental field such that each location $x \in \mathcal{X}$ is associated with a realized (random) measurement y_x (Y_x) if x is sampled/observed (unobserved). Let $\{Y_x\}_{x\in\mathcal{X}}$ denote a GP, that is, every finite subset of $\{Y_x\}_{x\in\mathcal{X}}$ has a multivariate Gaussian distribution [8]. The GP is fully specified by its prior mean $\mu_x \triangleq \mathbb{E}[Y_x]$ and covariance $\sigma_{xs} \triangleq \mathrm{cov}[Y_x, Y_s]$ for all $x, s \in \mathcal{X}$. In the experiments (Section 6), we assume that the GP is second-order stationary, i.e., it has a constant *prior* mean and a stationary *prior* covariance structure (i.e., σ_{xs} is a function of $x - s$ for all $x, s \in \mathcal{X}$). The prior mean and covariance structure of the GP are assumed to be known. Let $\mathcal{S}$ denote a subset of locations of $\mathcal{X}$ sampled a priori (either by the robot team or other sampling assets) and $y_\mathcal{S}$ be a row vector of corresponding measurements. Given the set $\mathcal{S}$ of sampled locations and corresponding measurements $y_\mathcal{S}$, the distribution of Y_x at any unobserved location $x \in \mathcal{X} \setminus \mathcal{S}$ remains Gaussian with the following *posterior* mean and variance

$$\mu_{x|\mathcal{S}} = \mu_x + \Sigma_{x\mathcal{S}}\Sigma_{\mathcal{S}\mathcal{S}}^{-1}(y_\mathcal{S} - \mu_\mathcal{S})^\top \quad (1)$$

$$\sigma_{x|\mathcal{S}}^2 = \sigma_x^2 - \Sigma_{x\mathcal{S}}\Sigma_{\mathcal{S}\mathcal{S}}^{-1}\Sigma_{\mathcal{S}x} \quad (2)$$

where $\mu_\mathcal{S}$ is a row vector with mean components μ_s for every location $s \in \mathcal{S}$, $\Sigma_{x\mathcal{S}}$ is a row vector with covariance components σ_{xs} for every location $s \in \mathcal{S}$, $\Sigma_{\mathcal{S}x}$ is the transpose of $\Sigma_{x\mathcal{S}}$, and $\Sigma_{\mathcal{S}\mathcal{S}}$ is a covariance matrix with components $\sigma_{ss'}$ for every pair of locations $s, s' \in \mathcal{S}$. To map the entire field, the measurements at its unobserved areas can be predicted using the posterior mean (1) and the uncertainty of each of these point-based predictions is represented by the posterior variance (2). An important property of GP is that the posterior variance $\sigma_{x|\mathcal{S}}^2$ (2) is independent of the observed measurements $y_\mathcal{S}$.

If the environmental field evolves over time, then its domain is extended to include the temporal dimension: let $\mathcal{X}$ instead denote a set of *spatiotemporal* inputs such that each input $x \in \mathcal{X}$ comprises both the spatial location and time. The rest of the GP model formulation remains unchanged.

3. COST FUNCTIONS

Recall that the exploration objective is to select observation paths that minimize the uncertainty of estimating the hotspot regions in the field. To achieve this, formal measures of uncertainty (specifically, in the form of cost functions) have to be defined. Let us first consider the feasibility of using cost functions that quantify the uncertainty of mapping the entire continuous-valued field, such as (a) sum of posterior variances (2) over the unobserved locations in $\mathcal{X} \setminus \mathcal{S}$ [4]

$$\sum_{x\in\mathcal{X}\setminus\mathcal{S}} \sigma_{x|\mathcal{S}}^2 \quad (3)$$

and (b) posterior joint entropy of the measurements $Y_{\mathcal{X}\setminus\mathcal{S}}$ at the unobserved locations in $\mathcal{X} \setminus \mathcal{S}$ [5]

$$\mathbb{H}[Y_{\mathcal{X}\setminus\mathcal{S}}|y_\mathcal{S}] \triangleq -\int P(y_{\mathcal{X}\setminus\mathcal{S}}|y_\mathcal{S}) \log P(y_{\mathcal{X}\setminus\mathcal{S}}|y_\mathcal{S}) \, \mathrm{d}y_{\mathcal{X}\setminus\mathcal{S}} \ .$$

These cost functions have been utilized in [4, 5] to guide exploration: the resulting active exploration strategies for learning GP-based field maps are non-adaptive and perform wide-area coverage, that is, observation paths are distributed evenly across the field. Do these wide-area coverage strategies also optimize our exploration objective or should observation paths be directed to sample specific features of the field instead? In the rest of this paper, we will show that, by defining cost functions to measure the uncertainty of classifying the hotspots in the field, our objective can be better achieved by performing the latter.

Let us begin by framing the problem of estimating the hotspot regions in a field formally as one of classifying/labeling the hotspots in the field: A location x is defined as a hotspot if its corresponding field measurement Y_x is greater than or equal to a predefined threshold, denoted by γ. Let $\{Z_x\}_{x\in\mathcal{X}}$ denote a binary random process such that Z_x is an indicator variable of label 1 if $Y_x \geq \gamma$ (i.e., location x is a hotspot), and label 0 otherwise. Then, our problem of estimating the hotspot regions is equivalent to one of labeling the hotspots in the field, specifically, by predicting the label of Z_x for every location $x \in \mathcal{X}$. As a result, our exploration objective can be achieved through the use of cost functions that measure the uncertainty of labeling the hotspots in the field. Two such cost functions will be defined next.

Let $\widehat{Z}_x$ be the predicted label of Z_x for every location $x \in \mathcal{X}$ and the cost of predicting (or, more precisely, misclassifying) the label of Z_x with $\widehat{Z}_x$ be denoted by the following $0-1$ loss function

$$L(Z_x, \widehat{Z}_x) = \left|Z_x - \widehat{Z}_x\right| = \begin{cases} 1 & \text{if } Z_x \neq \widehat{Z}_x \ , \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

That is, (4) counts a false positive (i.e., the location x is labeled as a hotspot but it is not) or false negative (i.e., x is not labeled as a hotspot but it is) as a misclassification. If Z_x is unlabeled (i.e., location x is unobserved), then we calculate the *expected* cost (or risk) of predicting the label of Z_x with $\widehat{Z}_x$ instead, which is denoted by

$$\begin{aligned} R_{\widehat{Z}_x|\mathcal{S}} = & \ \textstyle\sum_{i=0}^{1} L(Z_x = i, \widehat{Z}_x)\ P(Z_x = i|y_\mathcal{S}) \\ = & \ \widehat{Z}_x\ (1 - P(Z_x = 1|y_\mathcal{S})) + (1 - \widehat{Z}_x)\ P(Z_x = 1|y_\mathcal{S}) \\ = & \ P(\widehat{Z}_x \neq Z_x|y_\mathcal{S}) \end{aligned} \tag{5}$$

where $P(Z_x = 1|y_\mathcal{S}) = P(Y_x \geq \gamma|y_\mathcal{S})$, the second equality results from $P(Z_x = 0|y_\mathcal{S}) = 1 - P(Z_x = 1|y_\mathcal{S})$, and the last equality states that the risk (5) is equal to the probability of misclassification.

The risk (5) is minimized by the Bayes decision/classification rule

$$\begin{aligned} \widehat{Z}^*_x = & \ \begin{cases} 1 & \text{if } P(Z_x = 1|y_\mathcal{S}) \geq 0.5 \ , \\ 0 & \text{otherwise.} \end{cases} \\ = & \ \arg\max_{i\in\{0,1\}} P(Z_x = i|y_\mathcal{S}) \ . \end{aligned}$$

Using $\widehat{Z}^*_x$ as the predicted label of Z_x, the risk (5) reduces to

$$R_{\widehat{Z}^*_x|\mathcal{S}} = \min\left(P(Z_x = 1|y_\mathcal{S}), 1 - P(Z_x = 1|y_\mathcal{S})\right) \ . \tag{6}$$

Consequently, the sum of risks (or expected number of misclassifications) over the unobserved locations in $\mathcal{X} \setminus \mathcal{S}$ is

$$\sum_{x\in\mathcal{X}\setminus\mathcal{S}} R_{\widehat{Z}^*_x|\mathcal{S}} \ , \tag{7}$$

which defines our first cost function. We call this (7) the *misclassification criterion*.

The second cost function, which we call the *entropy criterion*, is defined as the posterior joint entropy of the labels of $Z_{\mathcal{X}\setminus\mathcal{S}}$ at the unobserved locations in $\mathcal{X} \setminus \mathcal{S}$

$$\mathbb{H}[Z_{\mathcal{X}\setminus\mathcal{S}}|y_\mathcal{S}] \ . \tag{8}$$

4. CENTRALIZED ACTIVE EXPLORATION

In this section, we will formulate greedy cost-minimizing exploration strategies based on the misclassification (7) and entropy (8) criteria defined in Section 3. Unfortunately, these centralized strategies cannot be evaluated in closed form, as explained in this section. To resolve this, these cost-minimizing strategies must first be reformulated as reward-maximizing dual strategies, from which we can then derive the approximate DARE strategy (Section 5) to be solved in closed form efficiently.

Supposing the misclassification criterion (7) is used and a set $\mathcal{S}$ of locations are previously sampled, the exploration strategy for directing a team of k robots has to select the next set $\mathcal{O} \subseteq \mathcal{X} \setminus \mathcal{S}$ of k locations to be observed that minimize the sum of *expected* risks:

$$\min_{\mathcal{O}} \sum_{x\in\mathcal{X}\setminus\mathcal{S}} \mathbb{E}_{Y_\mathcal{O}|y_\mathcal{S}}\left\{R_{\widehat{Z}^*_x|\mathcal{S}\cup\mathcal{O}}\right\} \ . \tag{9}$$

This cost-minimizing strategy (9) can be reformulated as the following reward-maximizing dual strategy, which selects the next set $\mathcal{O}$ of locations to be observed that maximize the sum of expected risk reductions:

$$\max_{\mathcal{O}} \sum_{x\in\mathcal{X}\setminus\mathcal{S}} R_{\widehat{Z}^*_x|\mathcal{S}} - \mathbb{E}_{Y_\mathcal{O}|y_\mathcal{S}}\left\{R_{\widehat{Z}^*_x|\mathcal{S}\cup\mathcal{O}}\right\} \ . \tag{10}$$

The equivalence between these two strategies follows immediately from observing that the first term $\sum_{x\in\mathcal{X}\setminus\mathcal{S}} R_{\widehat{Z}^*_x|\mathcal{S}}$ in (10) remains constant with any choice of $\mathcal{O}$. Both strategies cannot be solved exactly due to the expectation term, which cannot be evaluated in closed form.

If the entropy criterion (8) is used instead, then the exploration strategy has to select the next set $\mathcal{O}$ of locations to be observed that minimize the *expected* posterior joint entropy of the labels of $Z_{\mathcal{X}\setminus(\mathcal{S}\cup\mathcal{O})}$:

$$\min_{\mathcal{O}}\ \mathbb{E}_{Z_\mathcal{O}|y_\mathcal{S}}\left\{\mathbb{H}[Z_{\mathcal{X}\setminus(\mathcal{S}\cup\mathcal{O})}|y_\mathcal{S}, Z_\mathcal{O}]\right\} \ . \tag{11}$$

This cost-minimizing strategy (11) can be reformulated as the following reward-maximizing dual strategy, which selects the next set $\mathcal{O}$ of locations with maximum label entropy to be observed:

$$\max_{\mathcal{O}}\ \mathbb{H}[Z_\mathcal{O}|y_\mathcal{S}] \ . \tag{12}$$

To show their equivalence, $\mathbb{H}[Z_{\mathcal{X}\setminus\mathcal{S}}|y_\mathcal{S}]$ (8) is first expanded using chain rule of entropy:

$$\mathbb{H}[Z_{\mathcal{X}\setminus\mathcal{S}}|y_\mathcal{S}] = \mathbb{H}[Z_\mathcal{O}|y_\mathcal{S}] + \mathbb{E}_{Z_\mathcal{O}|y_\mathcal{S}}\left\{\mathbb{H}[Z_{\mathcal{X}\setminus(\mathcal{S}\cup\mathcal{O})}|y_\mathcal{S}, Z_\mathcal{O}]\right\} \ . \tag{13}$$

From (13), since $\mathbb{H}[Z_{\mathcal{X}\setminus\mathcal{S}}|y_\mathcal{S}]$ is a constant, the choice of $\mathcal{O}$ that maximizes $\mathbb{H}[Z_\mathcal{O}|y_\mathcal{S}]$ (i.e., (12)) will also minimize $\mathbb{E}_{Z_\mathcal{O}|y_\mathcal{S}}\left\{\mathbb{H}[Z_{\mathcal{X}\setminus(\mathcal{S}\cup\mathcal{O})}|y_\mathcal{S}, Z_\mathcal{O}]\right\}$ (i.e., (11)). When $|\mathcal{O}| = k \geq 2$, both strategies cannot be solved exactly due to the entropy terms, which contain multivariate Gaussian cumulative distribution functions that cannot be evaluated in closed form.

5. DECENTRALIZED ACTIVE EXPLORATION

This section presents a novel *decentralized active robotic exploration* (DARE) strategy that can approximately achieve both cost-minimizing exploration objectives (9) and (11) (Section 4) based on the misclassification and entropy criteria, respectively. Unlike the centralized cost-minimizing and reward-maximizing exploration strategies (Section 4), the DARE strategy can be solved in closed form efficiently.

The DARE strategy for directing each of the k robots has to select the next location $x \in \mathcal{X} \setminus \mathcal{S}$ to be observed that trades off between (a) minimizing the difference between its predicted measurement $\mu_{x|\mathcal{S}}$ and the boundary threshold γ, and (b) maximizing the square root of its posterior variance $\sigma^2_{x|\mathcal{S}}$:

$$\min_x \ |\gamma - \mu_{x|\mathcal{S}}|/\sigma_{x|\mathcal{S}} \ . \tag{14}$$

Intuitively, the behavior of the DARE strategy exhibits an interesting trade-off between that of (a) boundary tracking and (b) wide-area coverage: it simultaneously tracks a hotspot region boundary that is found until it can be accurately predicted as well as searches for new hotspot region boundaries in sparsely sampled areas to be tracked.

In this paper, the domain $\mathcal{X}$ of the field is assumed to be a grid of sampling locations. The next location x to be observed by each robot is then constrained to be selected from the 4-connected neighborhood $\mathcal{N}$ of the robot's current location instead of from $\mathcal{X} \setminus \mathcal{S}$.

THEOREM 1 (TIME COMPLEXITY). *Solving the DARE strategy (14) requires* $\mathcal{O}(|\mathcal{S}|^2(|\mathcal{S}| + |\mathcal{N}|))$ *time.*

The above result reveals that the time needed to compute the DARE strategy is independent of the map resolution (i.e., domain size $|\mathcal{X}|$) and the number k of robots, thus making it practical for *in situ*, real-time active sampling.

THEOREM 2 (COMMUNICATION OVERHEAD). *Let the communication overhead be the number of broadcast messages sent by each robot over the network. Then, the asynchronous communication overhead of DARE strategy (14) is* $\mathcal{O}(1)$.

In terms of data sharing, each robot broadcasts a message to the other robots sharing its sampled observations since its last broadcast. Coordination between robots is needed only if their neighborhoods intersect: in this case, they may select the same next location to be observed. To avoid this, each robot can broadcast on the same or another message sharing its selected location to be observed next. With asynchronous (e.g., turn-based) communication, the remaining robots avoid choosing prior selected locations. With synchronous communication, for each location in conflict, the higher-numbered robot (obtained by numbering robots uniquely) uses (14) to choose new unselected location to be observed. This is iterated until the conflicts are resolved. This process is not communication-expensive as every location is in conflict with at most 4 robots in its neighborhood.

Under conditional independence assumption, the DARE strategy (14) produces the same observation paths as that of the centralized cost-minimizing strategies (9) and (11) (Section 4), as established in the result below:

THEOREM 3 (PERFORMANCE GUARANTEE). *If the unobserved measurements* $Y_{\mathcal{X}\setminus\mathcal{S}}$ *are conditionally independent given the sampled measurements* $y_{\mathcal{S}}$*, then the DARE strategy (14) is equivalent to both cost-minimizing strategies (9) and (11) based on the misclassification and entropy criteria.*

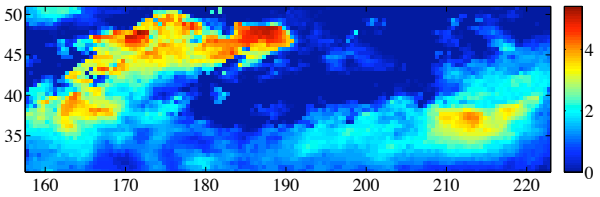


Figure 1: Temperature field bounded within lat. $30.75 - 50.75$**N and lon.** $157.75 - 222.25$**E:** γ **is set to** $3\,^\circ$C**, which results in a hotspot region in the top left and another one in the bottom right.**

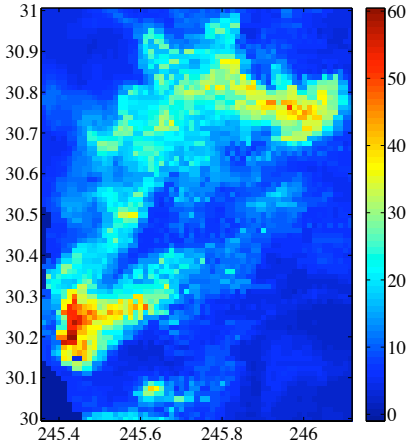


Figure 2: Plankton density field bounded within lat. $30 - 31$**N and lon.** $245.3625 - 246.1125$**E:** γ **is set to** 30 **mg/m**^{3}**, which results in a hotspot region in the top right and another one in the bottom left.**

The proof of the above result can be found in the appendix. The proof construction in fact describes how the DARE strategy (14) can be derived from either reward-maximizing dual strategy (i.e., (10) or (12)) (Section 4). A simple but important implication of this result is that the uncertainty of estimating the hotspot regions in a GP-based field (i.e., in terms of misclassification or entropy criterion) is greatest at or close to the hotspot region boundaries.

In practice, how reasonable is the conditional independence assumption? Firstly, such an assumption is often made in order to calculate the widely-used sum of posterior variances (i.e., mean-squared error) criterion (3) [2, 4]. Secondly, we conjecture that the assumption becomes less restrictive (i.e., Theorem 3 becomes more reliable) when the number $|\mathcal{S}|$ of sampled locations increases to potentially reduce the degree of violation of conditional independence, the spatial correlation between field measurements decreases, and the robots are sufficiently far apart (this last case applies only to the entropy criterion).

6. EXPERIMENTS AND DISCUSSION

This section evaluates the active exploration performance of the DARE strategy (14) empirically on 2 real-world spatial datasets off the west coast of USA: (a) August 2009 AVHRR temperature data (Fig. 1), and (b) March 2009 MODIS plankton density data (Fig. 2). These regions are discretized, respectively, into (a) 130×41 (i.e., $|\mathcal{X}| = 5330$) and (b) 61×81 (i.e., $|\mathcal{X}| = 4941$) grids of sampling locations. Each location x is, respectively, associated with (a) temperature measurement y_x in $^\circ$C, and (b) chlorophyll-a

Table 1: Comparison of active exploration strategies (WC: Wide-area Coverage, HS: Hotspot Sampling, BT: Boundary Tracking).

Exploration strategy	Behavior	Coordination type	Time complexity	Map resolution $\lvert\mathcal{X}\rvert$	No. of robots k
Maximize mutual information [3]	WC	Centralized	$\mathcal{O}(\lvert\mathcal{N}\rvert^k\lvert\mathcal{X}\rvert^2(\lvert\mathcal{X}\rvert+k^2))$	Cubic	Exponential
Minimize sum of variances [4]	WC	Centralized	$\mathcal{O}(\lvert\mathcal{N}\rvert^k\lvert\mathcal{S}\rvert^2\lvert\mathcal{X}\rvert)$	Linear	Exponential
MES [5]	WC	Centralized	$\mathcal{O}(\lvert\mathcal{N}\rvert^k\lvert\mathcal{S}\rvert^2(\lvert\mathcal{S}\rvert+k^2))$	Independent	Exponential
MES+HS [5]	WC+HS	Centralized	$\mathcal{O}(\lvert\mathcal{N}\rvert^k\lvert\mathcal{S}\rvert^2(\lvert\mathcal{S}\rvert+k^2))$	Independent	Exponential
Straddle [1]	WC+BT	Decentralized[1]	$\mathcal{O}(\lvert\mathcal{S}\rvert^2(\lvert\mathcal{S}\rvert+\lvert\mathcal{N}\rvert))$	Independent	Independent
DARE	WC+BT	Decentralized	$\mathcal{O}(\lvert\mathcal{S}\rvert^2(\lvert\mathcal{S}\rvert+\lvert\mathcal{N}\rvert))$	Independent	Independent

(chl-a) measurement y_x in mg/m^3. Using a team of $k = 2, 4, 8$ robots, each robot is tasked to, respectively, explore $1250, 625, 312$ locations in its path to sample a total of about 2500 observations. The simulated robot team is given 120 randomly selected observations as prior data before exploration. We use 2000 randomly selected observations to learn the hyperparameters (i.e., mean and covariance structure) of GP through maximum likelihood estimation [8].

6.1 Comparing Active Exploration Strategies

Since the domains $\mathcal{X}$ of both fields are considerably large, it is prohibitively expensive to compare meaningfully with the wide-area coverage strategies [3, 4] that scale poorly with increasing map resolution and are thus not practical for *in situ*, real-time active sampling. For example, it was reported in [6] that the greedy mutual information-based strategy of [3] incurred more than 62 hours to generate paths for 3 robots to sample a total of 267 observations in a grid of only $\lvert\mathcal{X}\rvert = 1424$ locations. The performance of the DARE strategy is therefore compared to that of three state-of-the-art exploration strategies whose incurred times are independent of the map resolution: (a) The decentralized[1] *straddle* strategy [1] for directing each robot selects the next location x to be observed using $\max_x 1.96\sigma_{x|\mathcal{S}} - |\gamma - \mu_{x|\mathcal{S}}|$. Similar to DARE, its exploration behavior is a trade-off between that of boundary tracking and wide-area coverage. Unlike DARE, its trade-off has to be manually adjusted using an arbitrary weight that, if set inappropriately, may produce suboptimal behavior. For example, this weight is proposed by [1] to be set to 1.96, which is empirically demonstrated later to emphasize boundary tracking more than wide-area coverage. As a result, it tends to persist in tracking boundaries that are already well-predicted before deciding to search for new ones. Subject to limited observations, it may consequently not perform as well as DARE in a field with multiple hotspot regions. Also, it is not known how or whether the value of this weight can be formally derived in order for the straddle strategy to achieve the cost-minimizing exploration objectives (9) and (11); (b) The centralized *maximum entropy sampling* (MES) strategy [5] for directing the robot team performs only wide-area coverage by selecting the next set $\mathcal{O}$ of locations with maximum entropy to be observed using $\max_{\mathcal{O}} \mathbb{H}[Y_{\mathcal{O}}|y_{\mathcal{S}}]$; (c) It can be coupled with hotspot sampling (HS) by modifying the exploration objective to $\max_{\mathcal{O}} \mathbb{H}[Y_{\mathcal{O}}|y_{\mathcal{S}}] + \sum_{x\in\mathcal{O}} \mu_{x|\mathcal{S}}$. We call this the MES+HS strategy [5]. For these centralized strategies, the joint action space is exponential in the number of robots. So, they scale poorly with increasing number of robots. Table 1 summarizes and compares the characteristics of the above-mentioned active exploration strategies; it does not include the communication overhead, which is $\mathcal{O}(1)$ for all strategies.

[1]The original straddle strategy proposed by [1] is developed for a single robot. To transform it into a decentralized multi-robot strategy, we simply execute the single-robot straddle strategy on every robot in the team.

6.2 Performance Metric

The first performance metric used to evaluate the tested strategies is the number of misclassifications

$$M(\mathcal{A}) \triangleq \sum_{x\in\mathcal{A}} L(z_x, \widehat{Z}^*_x)$$

over all locations in a given set $\mathcal{A}$ where the function L is previously defined in (4). Three cases are considered:

(a) $\mathcal{A} = \mathcal{X}$ (i.e., all locations in the domain of the field),

(b) $\mathcal{A} = \mathcal{X}'$ where

$$\mathcal{X}' = \{x \in \mathcal{X} \mid |\gamma - y_x| \leq 0.2(\max_{x'\in\mathcal{X}} y_{x'} - \min_{x'\in\mathcal{X}} y_{x'})\}$$

(i.e., all locations with measurements that are close to the boundary threshold of 30 mg/m^3 for the plankton density field and 3 °C for the temperature field), and

(c) $\mathcal{A} = \mathcal{X} \setminus \mathcal{X}'$.

We observe that $|\mathcal{X}'|$ is only about 22% of $|\mathcal{X}|$ for both fields. The second metric is the time taken to compute a strategy.

6.3 Temperature Field Data

Fig. 3 shows the results of the performance of tested strategies averaged over 5 randomly generated starting robot locations for the temperature field. In terms of the $M(\mathcal{X})$ performance, Figs. 3a−3c show that the DARE strategy quickly outperforms the MES and MES+HS strategies as the number of observations increases: their performance differences have been verified using t-tests ($\alpha = 0.1$) to be statistically significant after a total of 500, 750, and 800 observations sampled by teams of 2, 4, and 8 robots, respectively. Hence, the boundary-tracking DARE strategy reduces a greater number of misclassifications over the entire field than wide-area coverage and hotspot sampling. With more observations, the DARE strategy can also perform better than the straddle strategy: their performance differences have been verified using t-tests ($\alpha = 0.1$) to be statistically significant after a total of 500, 1000, and 1600 observations sampled by teams of 2, 4, and 8 robots, respectively. To explain this, we examine the observation paths of a team of 2 robots in one of the 5 test runs, as shown in Fig. 4. The initial performance of the DARE and straddle strategies are similar because they are both searching for hotspot region boundaries (Fig. 4a). As the number of observations increases further, DARE's performance improves over that of the straddle strategy because we observe that it directs the robots to search for new boundaries when the ones that are currently being tracked are well-predicted. In contrast, the

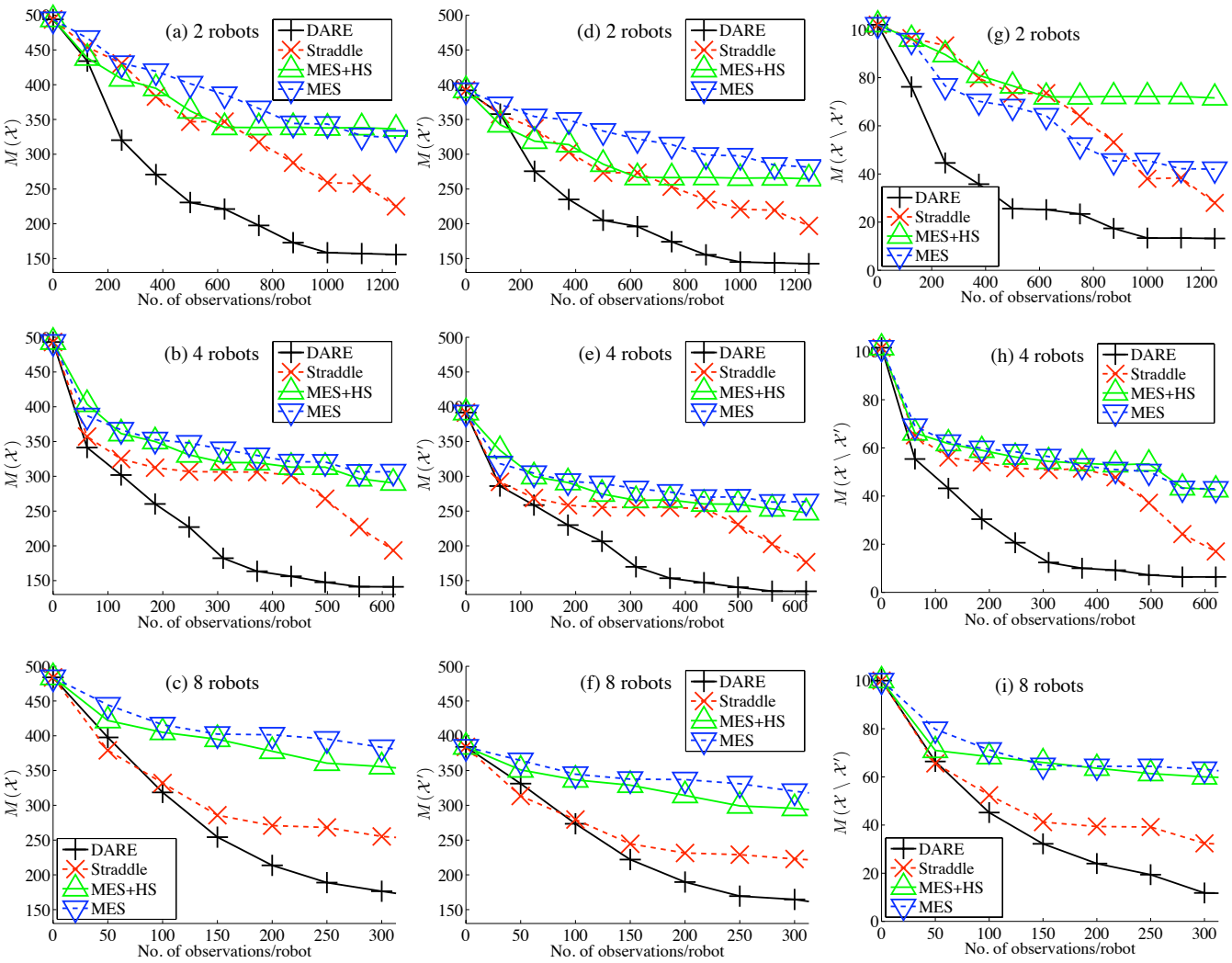


Figure 3: Graphs of (a-c) $M(\mathcal{X})$, (d-f) $M(\mathcal{X}')$, and (g-i) $M(\mathcal{X} \setminus \mathcal{X}')$ vs. no. of observations/robot for varying number of robots actively exploring the temperature field.

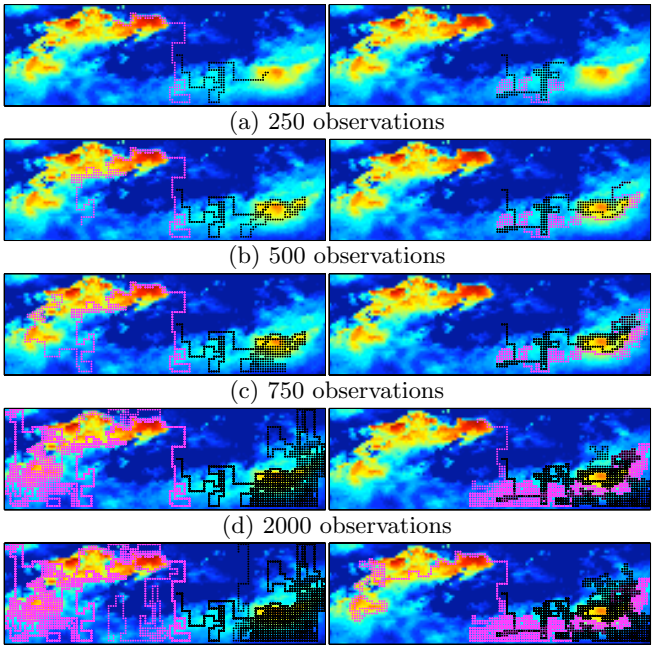
straddle strategy tends to persist in tracking boundaries that are already well-predicted before deciding to search for new ones (Figs. 4b−4e). In terms of the $M(\mathcal{X}')$ and $M(\mathcal{X} \setminus \mathcal{X}')$ performance, Figs. 3d−3i reveal that, with increasing observations, the DARE strategy also reduces a greater number of misclassifications than the other evaluated strategies whether they are over locations close to the boundaries (i.e., in $\mathcal{X}'$) or away from the boundaries (i.e., in $\mathcal{X} \setminus \mathcal{X}'$). It is interesting to note that locations close to the boundaries incur the majority of the misclassifications as compared to those away from the boundaries, which further corroborates the implication of Theorem 3 that there is higher uncertainty in labeling the locations close to the hotspot region boundaries.

6.4 Plankton Density Field Data

Fig. 5 shows the results of the performance of tested strategies averaged over 5 randomly generated starting robot locations for the plankton density field. The results are very similar to that of the temperature field (Section 6.3) except that the performance of the straddle strategy approaches that of the DARE strategy with excessive observations: their performance differences have been verified using t-tests ($\alpha = 0.1$) not to be statistically significant after a total of 2000 and 2240 observations sampled by teams of 2 and 4 robots, respectively. This is expected because the straddle strategy can track and predict the boundaries as well as the DARE strategy given a long enough exploration. However, subject to limited observations (which is more practical, as explained in Section 1), the performance of the DARE strategy is clearly superior to that of the straddle strategy.

6.5 Incurred Time

Fig. 6 shows the results of the time taken to compute the tested strategies averaged over 5 randomly generated starting robot locations for the temperature field; the results of incurred time for the plankton density field are very similar and therefore not shown here. The time differences between the DARE and straddle strategies have been verified using t-tests ($\alpha = 0.1$) not to be statistically significant, which is expected due to the same time complexities, as shown in Table 1. The time taken to compute these two decentralized boundary tracking strategies are shorter than that needed to compute the centralized wide-area coverage and



(a) 250 observations

(b) 500 observations

(c) 750 observations

(d) 2000 observations

(e) 2500 observations

Figure 4: Evolution of 2-robot observation paths produced by DARE (left column) and straddle (right column) strategies sampling a total of (a) 250, (b) 500, (c) 750, (d) 2000, and (e) 2500 observations. The robots start at locations of different lat. 37.75N and 40.75N and same lon. 192.75E.

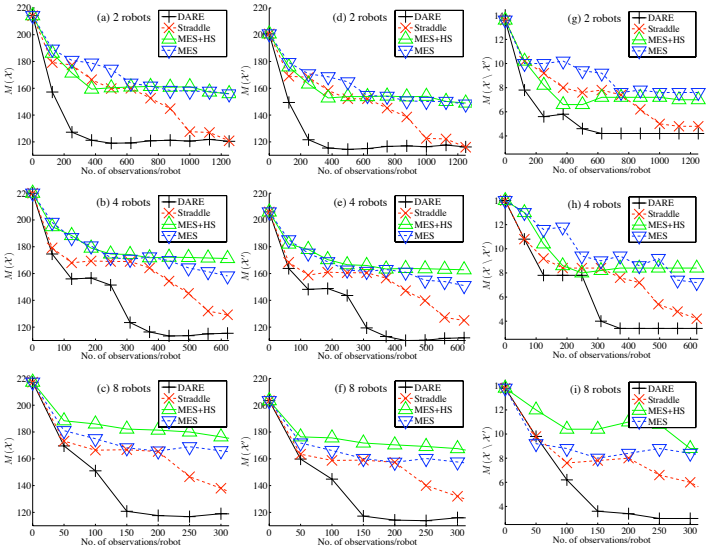


Figure 5: Graphs of (a-c) $M(\mathcal{X})$, (d-f) $M(\mathcal{X}')$, and (g-i) $M(\mathcal{X}\backslash\mathcal{X}')$ vs. no. of observations/robot for varying number of robots actively exploring the plankton density field.

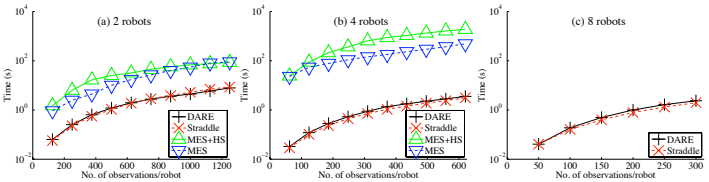


Figure 6: Graphs of incurred time vs. no. of observations/robot for varying number of robots actively exploring the temperature field.

hotspot sampling strategies (i.e., MES+HS and MES) by about one and two orders of magnitude for the cases of 2 and 4 robots, respectively. With a larger number of robots, it can be observed from Fig. 6 that the centralized wide-area coverage and hotspot sampling strategies incur significantly more time because their time complexities are exponential in the number of robots, as shown in Table 1. In contrast, the time incurred by the decentralized boundary tracking strategies do not increase because their time complexities are independent of the number of robots. Note that Fig. 6 does not show the graphs of time taken to compute the centralized strategies for the case of 8 robots because they incur significantly more time than that of 4 robots and their incurred time consequently cannot be recorded correctly due to long integer overflow in C's clock function.

7. CONCLUSION

This paper describes a decentralized active robotic exploration strategy for probabilistic classification of hotspots in a large-scale GP-based environmental field. It has the practical advantage of being significantly more time-efficient over existing state-of-the-art active exploration strategies [3, 4, 5] because its incurred time is independent of the map resolution and the number of robots. In terms of active exploration performance, we have theoretically guaranteed that, under reasonable conditional independence assumption, the DARE strategy can optimally achieve the formal cost-minimizing exploration objectives based on the misclassification and entropy criteria, both of which otherwise cannot be optimized exactly to yield closed-form solutions. We have demonstrated theoretically and empirically that the uncertainty of labeling the hotspots in a GP-based field is greatest at or close to the hotspot region boundaries. The DARE strategy is capable of exploiting this to produce an exploration behavior that formally trades off between that of boundary tracking until the hotspot region boundary can be accurately predicted and wide-area coverage to find new boundaries in sparsely sampled areas to be tracked. Empirical evaluation on real-world plankton density and temperature field data shows that, given limited observations, the DARE strategy can reduce a greater number of misclassifications than state-of-the-art active exploration strategies.

8. ACKNOWLEDGMENTS

This work was supported by MOE AcRF Tier 1 R-252-000-426-133. A portion of this work was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

9. REFERENCES

[1] B. Bryan, J. G. Schneider, R. Nichol, C. Miller, C. Genovese, and L. A. Wasserman. Active learning for identifying function threshold boundaries. In Y. Weiss, B. Schölkopf, and J. Platt, editors, *Advances in Neural Information Processing Systems 18*, pages 163–170, Cambridge, MA, 2006. MIT Press.

[2] N. A. C. Cressie. *Statistics for Spatial Data*. Wiley, NY, 2nd edition, 1993.

[3] A. Krause, A. Singh, and C. Guestrin. Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies. *JMLR*, 9:235–284, 2008.

[4] K. H. Low, J. M. Dolan, and P. Khosla. Adaptive multi-robot wide-area exploration and mapping. In *Proc. AAMAS*, pages 23–30, 2008.

[5] K. H. Low, J. M. Dolan, and P. Khosla. Information-theoretic approach to efficient adaptive path planning for mobile robotic environmental sensing. In *Proc. ICAPS*, pages 233–240, 2009.

[6] K. H. Low, J. M. Dolan, and P. Khosla. Active Markov information-theoretic path planning for robotic environmental sensing. In *Proc. AAMAS*, pages 753–760, 2011.

[7] K. H. Low, G. J. Gordon, J. M. Dolan, and P. Khosla. Adaptive sampling for multi-robot wide-area exploration. In *Proc. IEEE ICRA*, pages 755–760, 2007.

[8] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, Cambridge, MA, 2006.

[9] R. N. Smith, A. Pereira, Y. Chao, P. P. Li, D. A. Caron, B. H. Jones, and G. S. Sukhatme. Autonomous underwater vehicle trajectory design coupled with predictive ocean models: A case study. In *Proc. IEEE ICRA*, pages 4770–4777, 2010.

[10] S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics*. MIT Press, Cambridge, MA, 2005.

[11] R. Webster and M. Oliver. *Geostatistics for Environmental Scientists*. John Wiley & Sons, Inc., NY, 2nd edition, 2007.

[12] F. Zhang and N. E. Leonard. Cooperative control and filtering for cooperative exploration. *IEEE Trans. Automat. Contr.*, 55(3):650–663, 2010.

APPENDIX

Proof of Theorem 3

In Section 4, we have already shown the equivalence between the cost-minimizing and reward-maximizing strategies based on the misclassification and entropy criteria. Therefore, it suffices to prove that the DARE strategy (14) is equivalent to the reward-maximizing strategies.

Let us first prove that the reward-maximizing strategy (10) for the misclassification criterion is equivalent to the DARE strategy (14). From (10),

$$\begin{aligned}
&\max_{\mathcal{O}} \sum_{x\in\mathcal{X}\setminus\mathcal{S}} R_{\widehat{Z}^*_x|\mathcal{S}} - \mathbb{E}_{Y_\mathcal{O}|y_\mathcal{S}}\Big\{R_{\widehat{Z}^*_x|\mathcal{S}\cup\mathcal{O}}\Big\}\\
&= \max_{\mathcal{O}} \sum_{x\in\mathcal{O}} R_{\widehat{Z}^*_x|\mathcal{S}} + \sum_{x\in\mathcal{X}\setminus(\mathcal{S}\cup\mathcal{O})}\Big(R_{\widehat{Z}^*_x|\mathcal{S}} - \mathbb{E}_{Y_\mathcal{O}|y_\mathcal{S}}\Big\{R_{\widehat{Z}^*_x|\mathcal{S}\cup\mathcal{O}}\Big\}\Big)\\
&= \max_{\mathcal{O}} \sum_{x\in\mathcal{O}} R_{\widehat{Z}^*_x|\mathcal{S}} + \sum_{x\in\mathcal{X}\setminus(\mathcal{S}\cup\mathcal{O})}\Big(R_{\widehat{Z}^*_x|\mathcal{S}} - \mathbb{E}_{Y_\mathcal{O}|y_\mathcal{S}}\Big\{R_{\widehat{Z}^*_x|\mathcal{S}}\Big\}\Big)\\
&= \max_{\mathcal{O}} \sum_{x\in\mathcal{O}} R_{\widehat{Z}^*_x|\mathcal{S}}\\
&= \sum_{i=1}^{k} \max_{x_i} R_{\widehat{Z}^*_{x_i}|\mathcal{S}} \ .
\end{aligned}$$

The first equality follows from $R_{\widehat{Z}^*_x|\mathcal{S}\cup\mathcal{O}} = 0$ for $x \in \mathcal{O}$ by assuming no observation noise. The second equality is due to the conditional independence assumption that is provided as a sufficient condition in the theorem. The third equality is due to the second summation term evaluating to zero. The last equality follows from the observation that each risk term in the summation depends only on the choice of the next location x to be observed by a single different robot. Hence, we can maximize each risk term in the summation independently and in a decentralized manner to achieve the same result as that in the third equality.

$$\begin{aligned}
&\max_x R_{\widehat{Z}^*_x|\mathcal{S}}\\
&= \max_x \{\min\left(P(Z_x = 1|y_\mathcal{S}), 1 - P(Z_x = 1|y_\mathcal{S})\right)\}\\
&= \max_x \{\min\left(P(Y_x \geq \gamma|y_\mathcal{S}), 1 - P(Y_x \geq \gamma|y_\mathcal{S})\right)\}\\
&\equiv \max_x \left\{\min\left[-\mathrm{erf}\left(\frac{\gamma-\mu_{x|\mathcal{S}}}{\sigma_{x|\mathcal{S}}\sqrt{2}}\right), \mathrm{erf}\left(\frac{\gamma-\mu_{x|\mathcal{S}}}{\sigma_{x|\mathcal{S}}\sqrt{2}}\right)\right]\right\}\\
&= \max_x -\left|\ \mathrm{erf}\left(\frac{\gamma-\mu_{x|\mathcal{S}}}{\sigma_{x|\mathcal{S}}\sqrt{2}}\right)\ \right|\\
&\equiv \min_x \left|\ \mathrm{erf}\left(\frac{\gamma-\mu_{x|\mathcal{S}}}{\sigma_{x|\mathcal{S}}\sqrt{2}}\right)\ \right|\\
&\equiv \min_x \frac{|\gamma-\mu_{x|\mathcal{S}}|}{\sigma_{x|\mathcal{S}}\sqrt{2}}\\
&\equiv \min_x \frac{|\gamma-\mu_{x|\mathcal{S}}|}{\sigma_{x|\mathcal{S}}} \ .
\end{aligned}$$

The first equality follows from (6). The first equivalence is due to $P(Y_x \geq \gamma|y_\mathcal{S}) = \frac{1}{2}\left[1 - \mathrm{erf}\left(\frac{\gamma-\mu_{x|\mathcal{S}}}{\sigma_{x|\mathcal{S}}\sqrt{2}}\right)\right]$.

Now, let us prove that the reward-maximizing strategy (12) for the entropy criterion is equivalent to to the DARE strategy (14). From (12),

$$\begin{aligned}
&\max_{\mathcal{O}}\ \mathbb{H}[Z_\mathcal{O}|y_\mathcal{S}]\\
&= \max_{\mathcal{O}} \sum_{x\in\mathcal{O}} \mathbb{H}[Z_x|y_\mathcal{S}]\\
&= \sum_{i=1}^{k} \max_{x_i} \mathbb{H}[Z_{x_i}|y_\mathcal{S}] \ .
\end{aligned}$$

The first equality follows from chain rule of entropy and conditional independence assumption. The second equality follows from observing that each entropy term in the summation depends only on the choice of the next location x to be observed by a single different robot. Hence, we can maximize each entropy term in the summation independently and in a decentralized manner to achieve the same result as that in the first equality.

$$\begin{aligned}
&\max_x\ \mathbb{H}[Z_x|y_\mathcal{S}]\\
&\equiv \min_x P(Z_x = 1|y_\mathcal{S}) \log P(Z_x = 1|y_\mathcal{S}) +\\
&\qquad (1 - P(Z_x = 1|y_\mathcal{S})) \log(1 - P(Z_x = 1|y_\mathcal{S}))\\
&\equiv \min_x \left|\frac{1}{2} - P(Z_x = 1|y_\mathcal{S})\right|\\
&= \min_x \left|\frac{1}{2} - P(Y_x \geq \gamma|y_\mathcal{S})\right|\\
&\equiv \min_x \left|\ \mathrm{erf}\left(\frac{\gamma-\mu_{x|\mathcal{S}}}{\sigma_{x|\mathcal{S}}\sqrt{2}}\right)\ \right|\\
&\equiv \min_x \frac{|\gamma-\mu_{x|\mathcal{S}}|}{\sigma_{x|\mathcal{S}}\sqrt{2}}\\
&\equiv \min_x \frac{|\gamma-\mu_{x|\mathcal{S}}|}{\sigma_{x|\mathcal{S}}} \ .
\end{aligned}$$