# Statistical Models and Decision Making for Robotic Scientific Information Gathering

by

Genevieve Elaine Flaspohler

B.S.E., Computer Engineering, University of Michigan (2016)

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

Master of Science in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY AND
WOODS HOLE OCEANOGRAPHIC INSTITUTION

September 2018



Author .....
Department of Electrical Engineering and Computer Science
August 31, 2018

Certified by .....
Yogesh Girdhar
Assistant Scientist of Applied Ocean Physics and Engineering, WHOI
Thesis Supervisor

Certified by .....
Nicholas Roy
Professor of Aeronautics and Astronautics, MIT
Thesis Supervisor

Accepted by .....
Leslie A. Kolodziejski
Professor of Electrical Engineering and Computer Science
Chair, Department Committee on Graduate Students

Accepted by .....
Henrik Schmidt
Professor of Mechanical and Ocean Engineering
Chair, Joint Comittee for Applied Ocean Science and Engineering

# Statistical Models and Decision Making for Robotic Scientific Information Gathering

by

Genevieve Elaine Flaspohler

Submitted to the Department of Electrical Engineering and Computer Science
on August 31, 2018, in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical Engineering and Computer Science

## Abstract

Mobile robots and autonomous sensors have seen increasing use in scientific applications, from planetary rovers surveying for signs of life on Mars, to environmental buoys measuring and logging oceanographic conditions in coastal regions. This thesis makes contributions in both planning algorithms and model design for autonomous scientific information gathering, demonstrating how theory from machine learning, decision theory, theory of optimal experimental design, and statistical inference can be used to develop online algorithms for robotic information gathering that are robust to modeling errors, account for spatiotemporal structure in scientific data, and have probabilistic performance guarantees.

This thesis first introduces a novel sample selection algorithm for online, irrevocable sampling in data streams that have spatiotemporal structure, such as those that commonly arise in robotics and environmental monitoring. Given a limited sampling capacity, the proposed *periodic secretary algorithm* uses an information-theoretic reward function to select samples in real-time that maximally reduce posterior uncertainty in a given scientific model. Additionally, we provide a lower bound on the quality of samples selected by the periodic secretary algorithm by leveraging the submodularity of the information-theoretic reward function. Finally, we demonstrate the robustness of the proposed approach by employing the periodic secretary algorithm to select samples irrevocably from a seven-year oceanographic data stream collected at the Martha's Vineyard Coastal Observatory off the coast of Cape Cod, USA.

Secondly, we consider how scientific models can be specified in environments – such as the deep sea or deep space – where domain scientists may not have enough *a priori* knowledge to formulate a formal scientific model and hypothesis. These domains require scientific models that start with very little prior information and construct a model of the environment online as observations are gathered. We propose unsupervised machine learning as a technique for science model-learning in these environments. To this end, we introduce a hybrid Bayesian-deep learning model that learns a nonparametric topic model of a visual environment. We use this semantic visual model to identify observations that are poorly explained in the current model,

and show experimentally that these highly perplexing observations often correspond to scientifically interesting phenomena. On a marine dataset collected by the SeaBED AUV on the Hannibal Sea Mount, images of high perplexity in the learned model corresponded, for example, to a scientifically novel crab congregation in the deep sea.

The approaches presented in this thesis capture the depth and breadth of the problems facing the field of autonomous science. Developing robust autonomous systems that enhance our ability to perform exploratory science in environments such as the oceans, deep space, agricultural and disaster-relief zones will require insight and techniques from classical areas of robotics, such as motion and path planning, mapping, and localization, and from other domains, including machine learning, spatial statistics, optimization, and theory of experimental design. This thesis demonstrates how theory and practice from these diverse disciplines can be unified to address problems in autonomous scientific information gathering.

Thesis Supervisor: Yogesh Girdhar
Title: Assistant Scientist of Applied Ocean Physics and Engineering, WHOI

Thesis Supervisor: Nicholas Roy
Title: Professor of Aeronautics and Astronautics, MIT

# Acknowledgments

I am privileged to be able to think, work, and write surrounded by so many incredible people at MIT and WHOI. First and foremost, thank you to my thesis supervisors, Yogesh Girdhar and Nicholas Roy, for introducing me to robotics, for helping me to think critically about interesting problems, and for giving me an extraordinary amount of support, freedom and trust in the research process. To my labmates in the Robust Robotics Group, I could not imagine a more intelligent, hilarious, and considerate group of people to spend graduate school with.

To my parents and family, thank you for supporting my wanderlust; for being the toughest and most compassionate people I could have in my corner; and for showing me what it means to great, in every way possible. To my incredible partner, Nick, who has supported me with patience, compassion and laughter during the hard parts of graduate school, and whose intellectual curiosity, courage to challenge ideas, and drive to learn inspire me.

Finally, I would like to dedicate this thesis to my grandmother, Helen Jean Flapsohler, and my grandfather Ron Flaspohler, who took me seriously at age six when I asked how long it would take to get my Ph.D. (I have since learned that the answer is a potentially unbounded amount of time). This thesis is a step towards that Ph.D., which would not have been possible without the incredible spirit, wits, perseverance, and love that my grandparents brought to our family.

# Contents

**1 Introduction** **15**

1.1 Models for Scientific Information Gathering . . . . . 16

1.2 Planning for Scientific Information Gathering . . . . . 17

1.3 Thesis Contributions . . . . . 19

1.3.1 Irrevocable sampling in periodic data streams . . . . . 20

1.3.2 Hybrid Bayesian-deep models for visual terrains . . . . . 20

1.4 Thesis Outline . . . . . 21

**2 Technical Background and Foundational Related Work** **23**

2.1 Probabilistic Science Models . . . . . 24

2.1.1 Bayesian networks . . . . . 25

2.1.2 Bayesian nonparametric models . . . . . 26

2.2 Robot Models for Planning under Uncertainty . . . . . 32

2.2.1 Markov decision processes . . . . . 32

2.2.2 POMDPs . . . . . 33

2.3 Information Measures as Reward Functions . . . . . 35

2.3.1 Theory of optimal experimental design . . . . . 35

2.3.2 Fisher Information . . . . . 37

2.3.3 Information-theoretic reward . . . . . 38

2.4 Approximate Planning Strategies . . . . . 39

2.4.1 Open loop planning in Linear-Gaussian models . . . . . 40

2.4.2 Greedy algorithms with submodular reward . . . . . 41

**3 Irrevocable Information Gathering in Periodic Data Streams 45**

3.1 Introduction . . . 46

3.2 Related Work . . . 49

3.3 Scientific Model and Objective . . . 51

3.3.1 Gaussian process environmental model . . . 52

3.3.2 Reward functions and tradeoffs . . . 53

3.3.3 Submodular set functions . . . 54

3.4 Secretary Sampling Problem Formulation . . . 55

3.4.1 Periodic data streams . . . 55

3.4.2 Secretary sampling problem . . . 56

3.5 Periodic Secretary Sampling Algorithm . . . 57

3.6 Analysis of Algorithm Performance . . . 59

3.7 Experiments . . . 63

3.7.1 Using simulation to tune algorithm parameter $\lambda$ . . . 63

3.7.2 MVCO experiments . . . 64

3.7.3 MVCO sampling results . . . 66

3.8 Discussion . . . 68

**4 Hybrid Bayesian-Deep Topic Models for Terrain Characterization 69**

4.1 Introduction . . . 71

4.2 Related Work . . . 74

4.3 Methods . . . 75

4.3.1 Bayesian topic models . . . 76

4.3.2 Realtime spatiotemporal HDP model . . . 76

4.3.3 Convolutional autoencoder architecture and training . . . 78

4.3.4 Generating a visual vocabulary for topic models . . . 79

4.3.5 CAE feature visualization . . . 80

4.4 Experiments . . . 80

4.5 Results . . . 82

4.5.1 Mission I - seafloor terrain discovery . . . 82

4.5.2 Mission II - biological anomaly detection . . . . . 84
4.6 Discussion . . . . . 87

**5 Conclusion** **89**
5.1 Summary of Contributions . . . . . 89
5.1.1 Periodic secretary sampling with information reward . . . . . 90
5.1.2 Hybrid deep-Bayesian environmental models . . . . . 90
5.2 Directions for Future Work . . . . . 91
5.2.1 Secretary sampling problems . . . . . 91
5.2.2 Hybrid deep-Bayesian environmental models . . . . . 93
5.3 Final Remarks . . . . . 94

# List of Figures


2-1 Predictive mean and variance of a GP conditioned on varying amounts of training data . . . . . 30

3-1 Streaming irrevocable sample selection . . . . . 47
3-2 Approximately periodic data stream . . . . . 56
3-3 Periodic secretary algorithm with threshold parameter $\lambda$ . . . . . 59
3-4 Tuning parameter $\lambda$ . . . . . 64
3-5 MVCO environmental dataset . . . . . 65
3-6 Samples selected from the MVCO dataset . . . . . 66
3-7 Quantitative results on the MVCO dataset . . . . . 67

4-1 Visual topic models . . . . . 72
4-2 Architecture of visual scene understanding models . . . . . 73
4-3 Graphical model representation of ROST . . . . . 77
4-4 Network architecture for the convolutional autoencoder (CAE) used to extract low level visual features from the image datasets . . . . . 78
4-5 Results for unsupervised topic models versus hand-annotated terrain labels (b) for the Mission I dataset . . . . . 83
4-6 Results for two unsupervised topic models versus annotated labels (b) in for the Mission II dataset . . . . . 85
4-7 Correlation of perplexity score of the models (c,d,e) with annotated biological anomalies (b) for the Mission II dataset . . . . . 86

# List of Tables

4.1 Mutual information between discovered topics and annotations . . . . 84
4.2 Mutual information between perplexity and annotations . . . . . 87

# Chapter 1

# Introduction

Mobile robots and autonomous sensors are emerging as valuable tools in the human scientific endeavor, from planetary rovers surveying for signs of life on Mars, to environmental buoys measuring and logging oceanographic conditions in coastal regions. The use of autonomous agents for scientific information gathering requires consideration of the following questions:

- How should **scientific models** be formalized for use by autonomous agents?
- How should **scientific objectives** be specified for autonomous agents?
- How should autonomous agents evaluate the **utility or reward of potential actions** with respect to a scientific model and objective?
- How should autonomous agents **plan to take high-reward actions** in partially-observable environments by levering problem structure?

The first two questions relate to modeling; the second two relate to planning and decision-making. This thesis addresses the problem of scientific information gathering from both a modeling and planning perspective. The subsequent chapters demonstrate how theory from machine learning, decision theory, theory of optimal experimental design, and statistical inference can be used to develop online algorithms for scientific information gathering that are robust to modeling errors, account for the spatiotemporal structure in scientific data, and have probabilistic performance

guarantees. This chapter introduces the problem of scientific information gathering for mobile robots and autonomous sensors from both a modeling and planning perspective (Sections 1.1 and 1.2 respectively), followed by a detailed overview of the contributions presented in this thesis (Section 1.3).

## 1.1 Models for Scientific Information Gathering

When deploying robots or autonomous sensors on scientific missions, these autonomous agents must generally operate in partially-observable, stochastic environments. Effective decision-making in these environments requires a model of the underlying environment, describing how sensor observations relate to the latent scientific phenomenon of interest. We take a probabilistic perspective on the modeling problem, modeling the environmental phenomenon of interest as a stochastic process.

**Definition 1.** *(**Stochastic process**) A stochastic process is a collection of random variables indexed by a set of integers, often associated with time or space [28].*

Within this statistical model of the environment, one potential way to define a "scientific objective" is to denote a subset of the random variables in the environment as being of of scientific interest. The following applications exemplify how statistical models of environmental phenomena and scientific objectives can be formulated within the context of scientific information gathering missions:

**Example 1:** A chemical oceanographer is studying the structure of a methane plume in a coastal estuary. If the methane concentration throughout the estuary can be modeled using a Gaussian process, one scientific objective may be to uniformly reduce model uncertainty throughout the estuary, while another scientific objective may be to reduce model uncertainty only in a specific region of interest corresponding to the methane plume source.

**Example 2:** An astrogeologist is studying the geology in a Martian crater, as in Arora et al. [3]. A discrete grid may be used to represent the environment of interest,where hyper-spectral camera observations can be taken in each grid

cell to infer the geologic type of that cell. Given a geologic feature of interest, the scientific objective may be gather information about cells in the environment that are likely to contain this geologic feature.

In each example, the scientific phenomenon of interest is modeled as a stochastic process and scientific objective is defined by specifying properties of the environment that are of scientific interest e.g., methane plume source, containing a geologic feature.

## 1.2 Planning for Scientific Information Gathering

Once a scientific information gathering model and objective have been defined, the planning challenge is to understand the constraints of the autonomous robot or sensor and develop planning algorithms to sequentially select high-utility actions with respect to this mission objective in partially-observable environments. This section summarizes several canonical planning problems that appear in applications of autonomy to scientific problems.

Planning for scientific information gathering problems requires an autonomous agent to take actions sequentially in a partially-observable environment in order to satisfy an overall mission objective. The utility of potential actions with respect to a scientific objective can be quantified using a reward function. Chapters 2 and 3 demonstrate how information-theoretic reward functions can be applied to scientific information gathering to select actions that explicitly reduce the posterior model uncertainty of variables of interest. For a specific choice of reward function, many scientific information gathering planning problems can be classified into the following three canonical problem structures: sample selection, informative path planning, and secretary hiring or sampling.

Given a (discrete or continuous) $d$-dimensional environmental domain of interest $\mathcal{V} \subseteq \mathbb{R}^d$ and a set of potential sensing locations $\mathcal{S} \subseteq \mathcal{V}$:

**Sample Selection:** Sample selection problems require an autonomous agent to choose a discrete set of locations $\mathcal{A} \subseteq S$ to sample or place sensors [62, 63, 60, 104, 30]. These locations should be selected to maximize some scientific objective,

while satisfying a cost requirement. For example, a policy maker may wish to infer pollutant flow throughout a water body but only spend a limited budget on water sensors. Choosing where to place these sensors to optimally infer the pollutant flow is a sample selection problem.

**Informative Path Planning:** Informative path planning problems require a mobile autonomous agent to plan a trajectory contained within $\mathcal{S}$, such that the observations collected along the trajectory maximize some scientific objective, while satisfying a trajectory length or time/energy budget [122, 50, 52, 88, 48, 110, 74, 37, 121, 66, 49, 83, 76]. For example, an autonomous drone may need to survey the surface of a lake to estimate the size of a harmful algae bloom with a finite battery capacity. Planning a global trajectory that covers the bloom and gathers observations in regions of high plankton concentration is an informative path planning problem.

**Secretary Sampling Problems:** Like sample selection problems, secretary sampling problems [73, 26, 72] require an autonomous agent to choose a discrete set of locations $\mathcal{A} \subseteq \mathcal{S}$ in order to maximize some scientific objective. However, secretary sampling problems are constrained to operate on streams of data and furthermore impose the constraint that the agent must decide in real-time whether to collect a sample at a potential location in the data stream and cannot collect a sample at a location that was passed-over earlier in the stream. Additionally, sampling decisions are restricted to be irrevocable. If the sampling process is expensive (e.g., placing a physical sensor) or destructive (e.g., deploying a single-use water filter), the agent can only sample a finite number of times over the course of the mission, even if a previously collected sample is later revealed to be of low reward. For example, a static marine buoy may need to deploy 100 single-use water filters throughout the course of a year, such that the collected samples have the highest concentration of a harmful algae species. Selecting at which times to irrevocably deploy the water filters in real-time is a secretary sampling problem.

Although sample selection, informative path planning, and secretary sampling problems all commonly arise in scientific information gathering applications, the con-

tent of this thesis is focused on secretary sampling for scientific information gathering. The interested reader is referred to the aforementioned citations, which represent a small and incomplete subset of many contributions in these fields, for work in sample selection and informative path planning.

## 1.3 Thesis Contributions

The previous sections illustrate the sufficient components of a scientific information gathering mission: (1) a statistical model of the environment of interest as a collection of random variables, (2) a scientific objective that defines a subset of these random variables of scientific interest, (3) a reward function that quantifies the value of actions with respect to the scientific objective, and (4) a planning algorithm for choosing high-utility actions with respect to this reward function.

This thesis makes contributions in both planning algorithms and model design for scientific information gathering. First, Chapter 3 introduces a novel secretary sampling algorithm for irrevocable sampling in data streams that have spatiotemporal structure, such as those commonly arising in robotics and environmental monitoring. We demonstrate that information-theoretic reward functions can be used to select samples that maximally reduce model posterior uncertainty, and provide performance guarantees by leveraging the submodularity of the these information-theoretic reward functions. Second, Chapter 4 considers how to construct scientific models in domains such as the deep sea or deep space, when domain scientists may not have enough *a priori* knowledge to formulate a scientific model and hypothesis. These domains require scientific models that are provided with no or very little prior information, and construct a model of the environment as observations are gathered. To this end, Chapter 4 introduces a hybrid Bayesian-deep learning model for constructing nonparametric science models of unstructured visual environments.

### 1.3.1 Irrevocable sampling in periodic data streams

Chapter 3 considers the task of monitoring spatiotemporal phenomena in real-time by deploying limited sampling resources at locations of interest irrevocably and without knowledge of future observations. This task can be modeled as an instance of the classical *secretary problem*. Although this problem has been studied extensively in theoretical domains, existing algorithms require that data arrive in random order to provide performance guarantees. These algorithms will perform arbitrarily poorly on data streams such as those encountered in robotics and environmental monitoring domains, which tend to have complex spatiotemporal structure.

This chapter focuses on the problem of selecting representative samples from phenomena with periodic structure and introduces the *periodic secretary algorithm* that recovers a near-optimal sample set according to any monotone submodular reward function. The algorithm is evaluated on a seven-year environmental dataset collected at the Martha's Vineyard Coastal Observatory. We show that the periodic secretary algorithm selects phytoplankton sample locations that are nearly optimal in an information-theoretic sense for predicting phytoplankton concentrations in locations that were not directly sampled. The proposed periodic secretary algorithm can be used with theoretical performance guarantees in many real-time sensing and robotics applications for streaming, irrevocable sample selection from periodic data streams.

### 1.3.2 Hybrid Bayesian-deep models for visual terrains

Certain environments of scientific interest – such as those in the deep sea, on the surface of extraterrestrial planets, and in hazardous terrestrial environments – may be only coarsely understood *a priori* by scientific domain experts. For scientific information gathering problems in these domains, the appropriate choice of environmental model and scientific objective may be unclear. However, the emergence of low-cost cameras as a ubiquitous sensor on autonomous robots and sensing platforms provides a potential source of semantically rich information about an unfamiliar environment. Some of the most successful previous approaches to uncover latent semantic structure

directly from data have used probabilistic topic models [9, 45]. Despite the success of these models on textual data, they have not generalized as well to image data, in part because of the spatial and temporal structure that may exist in an image stream.

Chapter 4 introduces a hybrid Bayesian-deep unsupervised machine learning framework for learning environmental model structure directly from high-dimensional visual percepts. The proposed hybrid model incorporates the ability of convolutional autoencoders to discover features from images that directly encode spatial information, within a Bayesian nonparametric topic model that discovers meaningful latent patterns within visual data. By using this hybrid framework, this model overcomes the fundamental dependency of traditional topic models on rigidly hand-coded data representations, while simultaneously encoding spatial dependency in topics without additional model complexity. This model could be used in combination with an informative path planning algorithm to enable exploratory science missions, given no prior information about an environment. Experiments on a seafloor dataset collected by the SeaBED AUV marine robot show that the proposed hybrid framework outperforms current state-of-the-art approaches on the task of unsupervised seafloor terrain characterization.

## 1.4 Thesis Outline

The remainder of the thesis is organized as follows:

Chapter 2 introduces the technical background and foundational related work used throughout the thesis. The periodic secretary algorithm is presented in Chapter 3, along with a specification of the environmental model, scientific objective, and reward function used in this secretary sampling problem. Chapter 4 introduces a hybrid Bayesian-deep model for visual science understanding, and motivates this model with the problem of high-level scene understanding and mission summarization for exploratory marine robots. Finally, Chapter 5 concludes with a summary of the methods presented and suggests directions for future research within the field of autonomous scientific information gathering.

# Chapter 2

# Technical Background and Foundational Related Work

This chapter presents the technical background, notation, and concepts used throughout the thesis. The focus of this thesis is on scientific information gathering, with specific emphasis on modeling and planning problems. This chapter contains background on a variety of topics, including statistical models, decision-making under uncertainty, information-theoretic reward functions, and submodular optimization. The focus of this chapter is on high-level technical concepts and foundational related work; Chapters 3 and 4 additionally contain problem-specific related work and technical concepts.

Section 2.1 introduces strategies from statistics and machine learning for learning probabilistic models of scientific phenomena, focusing on two classes of Bayesian nonparametric models: the hierarchical Dirichlet process and the Gaussian process. Following this discussion of modeling strategies, common frameworks for planning and sequential decision-making under uncertainty are presented; Section 2.2 introduces Markov decision processes (MDPs) and partially-observable MDPs (POMDPs). Finally, Sections 2.3 and 2.4 discuss common technical challenges and solutions that arise when representing a scientific information-gathering problem as a POMDP. Section 2.3 introduces several reward functions appropriate for information-gathering problems and Section 2.4 describes planning strategies for approximately solving

MPDs and POMDPs with information reward, presenting several scenarios in which problem structure can be used to avoid the full complexity of planning under uncertainty, while maintaining performance guarantees.

Throughout this chapter, capital letters are used to denote random variables and corresponding lowercase letters are used to denote a specific value taken by the random variable from its alphabet or a non-random quantity. Boldface represents quantities which are explicitly vector-valued.

## 2.1 Probabilistic Science Models

Within robotics and autonomous sensing, there are two primary modeling challenges: modeling robot motion and dynamics, and modeling the environment the robot acts within. In the field of scientific robotics, the robot is often be assumed to have simple dynamics; however, the environment is generally complex, stochastic, dynamic, and otherwise challenging to model. Perhaps the most canonical environmental model used in robotics is the occupancy grid or voxel grid [115, 29], used to represent the 2D or 3D obstacle structure of an environment respectively. However, scientific problems often require an autonomous agent to infer properties of the environment other than occupancy. This section focuses on models that can be used to represent complex, partially observable scientific phenomena in scientific information gathering problems.

A variety of models exist for representing scientific phenomena within domain sciences and spatial statistics [22, 2, 24]. This thesis takes a probabilistic perspective on problem of modeling scientific phenomena, and focuses on Bayesian models from machine learning that allow for model learning from observed data and have explicit representations of uncertainty, which can be used to evaluate the utility of information-gathering actions with respect to their effect on model uncertainty. We utilize two nonparametric models to represent scientific phenomena: Gaussian processes (GPs) and hierarchical Dirichlet processes (HDPs). These models allow prior scientific knowledge to be incorporated into the model in the form of high-level structure, but use training data collected online to determine model complexity and refine

the model throughout the course of a mission as sensory information is collected. This mix of domain-specified prior structure and online *in situ* model refinement is desirable in many scientific applications. Section 2.1.1 first introduces a common class of causal probabilistic models known as Bayesian networks (BN). Section 2.1.2 expands upon general BN models and presents Bayesian nonparametric models by describing the GP and HDP models.

### 2.1.1 Bayesian networks

In general inference problems, we have a set of random variables $\mathcal{X} = \{X_0, \ldots, X_{N-1}\}$ that are related via a joint probability distribution $\mathbf{Pr}(X_0, \ldots, X_{N-1})$. The objective of Bayesian inference is to observe the values of some subset of these random variables $\mathcal{X}_{obs}$ and infer the likely values of unobserved, latent random variables $\mathcal{X}_{lat} = \mathcal{X} \setminus \mathcal{X}_{obs}$. By applying the rules of probability, this can be written as:

$$\mathbf{Pr}(\mathcal{X}_{lat} \mid \mathcal{X}_{obs}) = \frac{\mathbf{Pr}(\mathcal{X}_{lat}, \mathcal{X}_{obs})}{\mathbf{Pr}(\mathcal{X}_{obs})} = \frac{\mathbf{Pr}(\mathcal{X}_{lat}, \mathcal{X}_{obs})}{\sum_{\mathcal{X}_{lat}} \mathbf{Pr}(\mathcal{X}_{lat}, \mathcal{X}_{obs})}. \tag{2.1}$$

For general models, the computational complexity of exact inference is exponential in the number of latent random variables due to the summation in the normalization constant. This section explores probabilistic models that exploit conditional independence and conjugacy structure to reduce the complexity of inference and learning.

Probabilistic graphical models (PGMs) provide a general inference and estimation framework for modeling dependencies between random variables [8]. PGMs employ a graph structure of nodes and vertices to represent dependencies between random variables, where dependencies are represented by conditional probability distributions, factor tables, joint potentials, etc. As stated in the previous paragraph, in the general case inference in fully-connected PGMs – where every random variable is dependent on every other random variable – is exponentially difficult. However, prior knowledge about structure in a problem can be used to omit edges from the dependency graph (i.e., assuming conditional independence between random variables); efficient inference is possible in sparsely connected PGMs.

Bayesian networks (BNs) are a type of PGM in which the dependency structure is represented by a directed, acyclic graph. A BN consists of a graph $\mathcal{G}$ consisting of a set $V$ of nodes or vertices such that $|V| = N$ and a set of edges $\mathcal{E} \subseteq N \times N$. Each vertex represents a random variable (or a group of random variables) $X_i \in V$ and edges represent directed dependencies between random variables as an ordered tuple $(i, j)$, such that $(i, j) \in \mathcal{E}$ if random variable $X_j$ depends directly on $X_i$. This graph captures how the joint probability distribution over the random variables in $V$ can be decomposed. In a BN, the joint probability of random variables $V$ is:

$$\mathbf{Pr}(X_0, \ldots, X_{N-1}) = \prod_{i=0}^{N-1} \mathbf{Pr}(X_i \mid \text{parents}(X_i)), \tag{2.2}$$

where the parents of a node $X_i$ in the graph are nodes $X_j$ for which $(j, i) \in \mathcal{E}$. Exact inference is possible for certain BNs using inference algorithms that employ message passing such as the sum-product/max-sum algorithm [8]. However, for many models, exact inference is computationally intractable even for sparsely connected dependency structures. Approximate inference techniques, such as Monte Carlo techniques and variational inference, are discussed and employed in Chapter 4 .

Bayesian networks have seen adoption in scientific sensing problems, as they allow for a very flexible environmental model and directly encode causal structure between the random variables in the environment using the dependency graph. Arora et al. [3] model the environment for scientific information gathering as a discrete grid and use a tree-structured BN to represent the geologic type of a cell based on nearby cells and environmental sensor measurements. Candela et al. [14] use a BN for hypothesis-driven geologic mapping using a mobile robot in a space-based exploration application. Active SLAM and object localization and tracking information gathering problems often employ filtering-based probabilistic models [66, 59, 17, 18, 15, 16].

### 2.1.2 Bayesian nonparametric models

One of the most challenging aspects of scientific model design is managing the trade-off between model complexity and model generalization. More complex models with

more free parameters will generally fit observed data better then simpler models. However, these highly flexible models will tend to overfit the observed data and have worse generalization performance when presented with new data. On the other hand, models that are too simple will be unable to capture variation in the dataset; this is known as underfitting.

One popular strategy for model selection is to fit parametric models with a fixed number of parameters to a portion of the observed data, and quantify inference accuracy on the other portion of the data. This technique is known as cross-validation. By fitting parametric models with varying complexity (different numbers of parameters) and comparing the cross-validation performance of each, a model designer can choose a model that balances the underfit-overfit trade-off and performs well when presented with new data in a deployment setting. Another option is Bayesian model selection, in which a prior is placed over the different classes of models, or explicit regularization, in which more complex models incur a complexity penalty.

In this thesis, we consider an alternative approach to model design and managing model complexity known as Bayesian nonparametrics (BNP). Rather than explicitly considering parametric models of differing complexity, BNP models allow model complexity to grow as more data is observed. For example, in clustering problems, instead of explicitly defining a number of clusters, a BNP approach allows the number of inferred clusters to grow as more data points are observed. In problems for which model growth is an accurate assumption, BNPs are a powerful way to manage model complexity. In many scientific contexts, model growth is desirable: the more of an environment a robot explores or the more of a phenomenon a sensor observes, the more complex the inferred model will generally be.

### The hierarchical Dirichlet process

The hierarchical Dirichlet process (HDP) is a Bayesian nonparametric latent variable model that models clustering problems in which data are partitioned into multiple groups and cluster parameters are shared between groups. The HDP is considered a nonparametric Bayesian model because the number of clusters is not fixed and can

grow with the number of data points observed.

The most canonical application of the HDP is to the problem of topic modeling. In topic modeling problems, the observed data are a set of words divided amongst a fixed number of documents in a text corpora. The words in each document are assumed to represent exchangeable samples from a document-specific mixture over global topics or categories, such as sports, economics, or politics; this mixture is denoted $\theta_j$ for document $j$. Within each document, words are generated from these document-topic mixture weights as follows: for every word $w_{ij}$ in document $j$, a topic-label $z_{ij}$ is generated from $\theta_j$, and word $w_{ij}$ is generated from the corresponding topic-word distribution represented by $z_{ij}$. Topic models, including the HDP and the original parametric latent Dirichlet allocation, are discussed in much further detail in Blei et al. [9], Griffiths and Steyvers [45], and Teh et al. [113].

The generative model described by the HDP for text corpora is as follows:

1. Sample a global measure over topics/mixture weights: $G_0 \mid \gamma, H \sim \mathrm{DP}(\gamma, H)$.
2. For each document $j$:
   (a) Sample a measure over document-specific topic weights: $\theta_j \mid \alpha_0, G_0 \sim \mathrm{DP}(\alpha_0, G_0)$.
   (b) For each word $i \in \{1, \ldots, N^{(j)}\}$ :
      i. Sample a word-topic distribution: $z_{ij} \mid \theta_j \sim \theta_j$.
      ii. Sample a word: $w_{ij} \sim \mathrm{Mult}(z_{ij})$.

The generative model for the HDP relies on the stick-breaking construction of the Dirichlet process (DP), which generates a countably infinite number of cluster parameters $\phi$ and mixture weights $\beta, \pi$ for which the cluster weights sum to one, i.e., the cluster weights form a proper probability distribution:

$$G_0 = \sum_{k=1}^{\infty} \beta_k \delta_{\phi_k}, \text{ and,} \tag{2.3}$$

$$\theta_j = \sum_{k=1}^{\infty} \pi_{jk} \delta_{\phi_k}, \tag{2.4}$$

where the $\pi_{jk}$ are the document-specific mixture weights (generally, $\beta_k \neq \pi_{jk}$) and $\phi = (\phi_k)_{k=1}^{\infty}$ are drawn from the base distribution $H$ which is assumed to be a symmetric Dirichlet over the vocabulary simplex i.e., each $\phi_k$ is a "topic".

The nonparametric nature of the hierarchical Dirichlet process allows for infinitely many topics to be inferred in the limit as infinitely many words are observed. However, for a dataset with finitely many words, only a finite number of these topics will ever be instantiated. Due to the conjugacy of the HDP model, as well as the conditional independencies encoded by the generative model (see also the graphical model representation of the HDP in Chapter 4, Figure 4-3), approximate inference of latent topics and mixing proportions given a set of observed words partitioned into documents is possible using variational inference or MCMC methods. Chapter 4 applies a collapsed Gibbs sampling approach to infer latent parameters from a set of scientific images.

The hierarchical Dirichlet process has many desirable properties for scientific problems due to the ability to instantiate an unlimited number of semantic topics as more data is observed and because the HDP requires no supervision other then the segmentation of data into "documents" or partitions, which often arises naturally for specific problems. The HDP and Dirichlet process mixtures models have been used to perform unsupervised robotic exploration and terrain classification in a variety of scientific contexts, especially in the domain of marine robotics [105, 108, 36, 107, 42, 41, 40].

### Gaussian process models

Whereas the hierarchical Dirichlet process is a nonparametric Bayesian model for discrete data such as text corpora, Gaussian processes (GPs) [90] are a versatile nonparametric model, often used to perform continuous regression in a variety of applications and contexts [63, 101, 25, 69, 78]. Let the $d$-dimensional environmental domain of interest be represented as a compact subset $\mathcal{V} \subset \mathbb{R}^d$ and let $f : \mathcal{V} \to \mathbb{R}^m$ be an unknown function representing an $m$-dimensional continuous environmental phenomenon of interest. For the remainder of this chapter, we assume that $m = 1$ for simplicity of notation.

A robot or autonomous sensor can collect a noisy observation of the function $f$ at location $\mathbf{x}$ by deploying the appropriate sensor. The observation $Y$ collected at location $\mathbf{x}$ is assumed to is a random variable with the following form: $Y = f(\mathbf{x}) + \epsilon$, where $\epsilon \sim \mathcal{N}(0, \sigma_n^2)$ and $\sigma_n^2$ is specified by a sensor noise model.

We model the function $f$ using a Gaussian process with isometric process noise $\sigma_n^2$. A Gaussian process describes a distribution over functions, where function values at any finite number of points are restricted to have a joint Gaussian distribution:

**Definition 2.** *(**Gaussian process [90]**) A Gaussian process is an infinite collection of random variables, any finite number of which have a joint Gaussian distribution.*

The form of this joint Gaussian distribution is specified by the GP mean $m(\mathbf{x})$ and covariance $k(\mathbf{x}, \mathbf{x}')$ functions:

$$m(\mathbf{x}) = \mathbb{E}[f(\mathbf{x})], \tag{2.5}$$

$$k(\mathbf{x}, \mathbf{x}') = \mathbb{E}[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))]. \tag{2.6}$$

The resulting Gaussian process places a prior over functions $f$ and is written $f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}'))$. This predictive mean and variance are visualized in Figure 2-1.

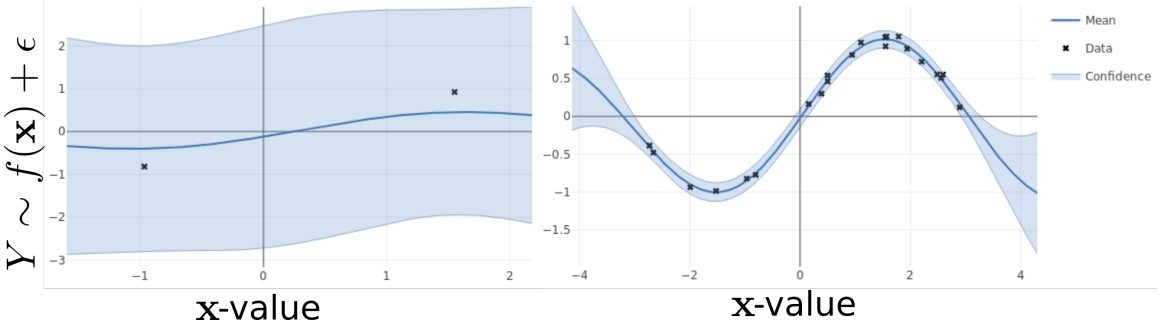


Figure 2-1: **Predictive mean and variance of a GP conditioned on varying amounts of training data:** In the left plot, the predictive mean and variance are visualized after observing two paired samples $\{\mathbf{x}_i, y_i\}$ (shown as black x's) with a squared exponential kernel. Predictive variance is shown as blue shading, representing a 95% confidence interval for functions drawn from this conditioned GP. In the right plot, more training examples have been observed (22 points). Note that the 95% confidence interval has shrunk near previous sample locations and the predictive mean is close to the observed samples. Figure created using GPy [44].

Given a training dataset of $N$ data points, consisting of locations $\mathbf{x} \in \mathcal{V}$ and process observations $y \in \mathbb{R}$, such that $\mathcal{D}_N = \{(\mathbf{x}_i, y_i) \mid i = 0, \ldots, N-1\}$, the posterior distribution at a new location $\mathbf{x}' \in \mathcal{V}$ can be computed as:

$$f(\mathbf{x}') \mid \mathcal{D}_N \sim \mathcal{N}(\mu_N(\mathbf{x}'), \sigma_N^2(\mathbf{x}')), \text{where} \tag{2.7}$$

$$\mu_N(\mathbf{x}') = \kappa_N(\mathbf{x}')^T(\mathbf{K}_N + \sigma_n^2\mathbf{I})^{-1}\mathbf{y}_N, \tag{2.8}$$

$$\sigma_N^2(\mathbf{x}') = \kappa(\mathbf{x}', \mathbf{x}') - \kappa_N(\mathbf{x}')^T(\mathbf{K}_N + \sigma_n^2\mathbf{I})^{-1}\kappa_N(\mathbf{x}'), \tag{2.9}$$

where $\mathbf{y}_N = [y_0, \ldots, y_{N-1}]^T$, $\mathbf{K}_N$ is the positive definite kernel matrix with $\mathbf{K}_N[i,j] = \kappa(\mathbf{x}_i, \mathbf{x}_j)$ for all $\mathbf{x}_i, \mathbf{x}_j \in \mathcal{D}_N$, and $\kappa_N(\mathbf{x}') = [\kappa(\mathbf{x}_0, \mathbf{x}'), \ldots, \kappa(\mathbf{x}_{N-1}, \mathbf{x}')]$.

Generally, a zero or constant-valued mean function is used; non-constant mean functions can represent time- or space-varying offsets in a regression problem. The squared-exponential and Matérn covariance functions are frequently used to model continuous environmental phenomenon with varying degrees of prior smoothness:

$$k_{\text{SE}}(\mathbf{x}, \mathbf{x}') = \sigma^2 \exp\Big(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{l^2}\Big), \tag{2.10}$$

$$k_{\text{Matérn}}(\mathbf{x}, \mathbf{x}') = \frac{2^{1-\nu}}{\Gamma(\nu)}\Big(\frac{\sqrt{2\nu}\,\|\mathbf{x} - \mathbf{x}'\|}{l}\Big)^{\nu} K_\nu\Big(\frac{\sqrt{2\nu}\,\|\mathbf{x} - \mathbf{x}'\|}{l}\Big), \tag{2.11}$$

with positive hyperparameters $\sigma, l, \nu$, where $K_\nu$ is a modified Bessel function [1]. These hyperparameters specify properties such as smoothness, function value range, and strength of correlation between points in the resulting GP prior over functions. For certain problems, it is possible to rely directly on scientific prior knowledge to set the values of these hyperparameters; more often, cross-validation or online optimization are employed to learn appropriate values [90].

Gaussian process have been applied widely in scientific sensing problems to model continuous environmental phenomena [62, 7, 46, 63, 48, 122, 114]. GPs are a powerful nonparametric model, are easy to train and perform inference with, and admit closed form solutions for a variety of useful information-theoretic measures (see Section 2.3). However, there are several limitations when using GPs to model complex environmental phenomena, including the cubic growth of training and inference time

with the number of observations and the relative difficulty encoding causal scientific structure in an intuitive manner using only the covariance and mean function.

## 2.2 Robot Models for Planning under Uncertainty

Although environmental modeling is often the primary modeling focus in scientific information-gathering problems, selecting and executing information-gathering actions in a real environment requires a scientific agent to reason about the actions available to it, as well as the immediate and long-term effect of these actions over the course of a mission. Planning for scientific missions additionally requires a model of robot dynamics and the reward of candidate actions. This section reviews standard approaches for modeling sequential decision-making problems. The environmental model and robot model are both contained within the decision-state of the autonomous agent and are connected via the reward function, through which the value of candidate actions can be evaluated with respect to both the perceived state of the robot and the state of the environment. Different specifications of reward function can induce a variety of desirable behaviors e.g., information gathering, maxima-seeking, target tracking and localization.

Many scientific information-gathering problems can be formulated as sequential decision-making problems under uncertainty. When this decision-making process has a Markovian structure, i.e., the optimal decision is independent of previous observations and actions, conditioned on the current decision-state (or belief), the most general problem formulation is a partially observable Markov decision process (POMDP). We first introduce the Markov decision process, which is a sequential decision-making framework used when the decision state is fully observable, and then show how this model changes when state uncertainty is incorporated.

### 2.2.1 Markov decision processes

A Markov decision process (MDP) is a general framework for representing sequential decision-making problems. An MDP can be represented as a tuple $(S, A, T, R, s_0)$,

where:

- $S$: The set of possible decision states $\{s_1, s_2, s_3, \dots\}$. Can generally be finite or infinite, as in the case of continuous state spaces. The decision state is a sufficient statistic for the history of actions taken by the agent i.e., the sequence of states form Markov process, where future states are independent of past states conditioned on the current state.

- $A$: The set of actions $\{a_1, a_2, a_3, \dots\}$ available to the agent. Can generally be finite or infinite, but often assumed to be finite. The action set can be time-varying, state-dependent, or static.

- $T : S \times A \times S \to \mathbb{R}$ The **transition function**, representing the probability density of being in state $s' \in S$ after executing action $a \in A$ from state $s \in S$ i.e., $T(s, a, s') = \mathbf{Pr}(S_{t+1} = s' \mid S_t = s, A_t = a)$. The transition function can model imperfect robot control and dynamic environments.

- $R : S \times A \to \mathbb{R}$ The **reward function**, representing reward of performing action $a \in A$ from state $s \in S$.

- $s_0$: The initial state $s_0$, such that $\mathbf{Pr}(S_0) = \delta_{S_0 = s_0}$.

The solution to an MDP is a policy $\pi : S \to A$, a function mapping from states to actions. The optimal policy $\pi^*$ determines the optimal action to take from any given decision-state. The optimal policy over an infinite horizon will maximize total discounted reward for discount factor $\gamma$:

$$\pi^* = \underset{\pi}{\operatorname{argmax}}\, \mathbb{E}\Big[\sum_{t=0}^{\infty} \gamma^t \cdot R(s_t, a_t) \mid s_0, \pi\Big], \tag{2.12}$$

### 2.2.2 POMDPs

The MDP formulation assumes that the decision-state of the agent is fully observable at each time step. However, in many scientific information gathering problems, either the state of the robot or the state or the environment or both could be uncertain.

Uncertainty in decision-making for mobile robots or autonomous sensors presents a serious challenge. Developing robust, efficient algorithms in uncertain environments is an active area of research in robotics and decision-making. The partially observable Markov decision process (POMDP) framework extends the MDP framework to situations where the decision-state is only partially observable, and provides a principled and general framework to model an autonomous agent acting in a partially observable, stochastic environment [54]. The POMDP framework is represented as a tuple $(S, A, T, R, Z, O, b_0)$, where $S$, $A$, and $T$ are identical to the MDP definition in Section 2.2.1, but additionally including:

- $Z$: The set of possible observations $\{z_1, z_2, z_3, \dots\}$. Can generally be finite or infinite, as in the case of continuous observation spaces.

- $O : S \times A \times Z \to \mathbb{R}$ The **observation function**, representing the probability density of observing $o$ after executing action $a$ from state $s$ i.e., $O(s, a, o) = \mathbf{Pr}(O_t = o \mid S_t = s, A_t = a)$. The observation function can model imperfect perception and sensor noise.

- $b_0$: The prior distribution on the initial state $S_0$ i.e., $b_0(s) = \mathbf{Pr}(S_0 = s)$.

In a POMDP, the optimal policy for the autonomous agent can no longer depend only on the current state of the model, because this state is uncertain; by including uncertainty in the model state, we have lost the Markov property of MDPs. However, this Markov property can be restored by considering the current belief over the state in place of the state. Importantly, in a POMDP, the optimal action depends only on the agent's current belief state. In other words, the belief state summarizes all relevant information in the history of actions and observations.

The following sections discuss two important components of defining and solving a POMDP for scientific information gathering problems: Section 2.3 discusses information-theoretic reward functions and Section 2.4 discusses approximate planning strategies for solving sequential decision-making problems once they have been formulated as an MDP or POMDP.

## 2.3 Information Measures as Reward Functions

Formulating a sequential decision-making problem as a POMDP can be straightforward. Often, the appropriate choice of state space, action space, transition function, and observation function are obvious from the properties of the physical autonomous agent. Depending on the task considered, the appropriate choice of reward function may be obvious or, as in the case of scientific missions, can be less clear. Reward shaping and design of reward signals has received considerable attention in fields such as reinforcement learning [111].

For example, if the objective of the autonomous agent is to move from a start location to a goal state while minimizing distance traveled, an appropriate reward function is fairly straightforward to specify: the agent receives some reward for reaching the goal state and receives some negative reward for every unit of distance traveled. However, if the objective of the autonomous agent is to learn a scientific model of the environment, the reward function must somehow reflect this "information reward". This section presents several metrics from literature that quantify the information content of a set of actions or measurements.

### 2.3.1 Theory of optimal experimental design

Information measures have been considered extensively in the field of optimal experimental design [86]. The field of experimental design considers the problem of estimating a vector $\mathbf{z} \in \mathbb{R}^d$ from a set of measurements or "experiments" and outcomes:

$$y_i = \mathbf{x}_i^T \mathbf{z} + w_i, \quad i = 0, \ldots, N-1 \tag{2.13}$$

where $w_i$ is measurement noise, often assumed to be drawn independently from a Gaussian with zero mean and unit variance.

The maximum likelihood estimate of $\mathbf{z}$ is given by the standard least square solution:

$$\hat{\mathbf{z}} = \Big( \sum_{i=0}^{N-1} \mathbf{x}_i \mathbf{x}_i^T \Big)^{-1} \sum_{i=0}^{N-1} y_i \mathbf{x}_i. \tag{2.14}$$

For this estimator, the estimation error $\mathbf{E} = \hat{\mathbf{z}} - \mathbf{z}$ is unbiased with covariance matrix:

$$\mathbf{E} = \Big( \sum_{i=0}^{N-1} \mathbf{x}_i \mathbf{x}_i^T \Big)^{-2}. \tag{2.15}$$

This matrix $\mathbf{E}$ characterizes the variance of the MLE estimator and therefore characterizes the "informativeness" of the selected measurements $\mathbf{x}_0, \ldots, \mathbf{x}_{N-1}$ by quantifying how effectively these measurements reduce the error variance of estimator $\hat{\mathbf{z}}$.

The goal of experimental design is to select the set of measurement $\mathbf{x}_0, \ldots, \mathbf{x}_{N-1}$ from the set of possible measurements $\mathcal{V} \subseteq \mathbb{R}^d$ to make some function of the error covariance small. The "size" of the error covariance can be measured in a variety of ways, giving rise to the so-called "alphabet-soup of experimental design criteria":

- ***D*-optimal design:** Perhaps the most widely used experimental design criteria. Minimizes the determinant or of the error covariance matrix $\mathbf{E}$ or the volume of the resulting confidence ellipsoid for a fixed confidence level.

- ***E*-optimal design:** Minimizes the norm of the error covariance matrix or the maximum eigenvlaue of $\mathbf{E}$. Geometrically, corresponds to minimizing the diameter of the confidence ellipsoid.

- ***A*-optimal design:** Minimizes the trace of the error covariance matrix.

The interested reader is referred to Boyd and Vandenberghe [12] for a more detailed treatment of experimental design, additional design criteria, and a discussion of methods for optimizing these experimental design criteria.

A number of works have used these optimal design criteria to select measurements that maximally reduce the error variance in the maximum likelihood estimator. The active SLAM community has explored a variety of design criteria for selecting measurements that are the most informative about the state of the robot [98, 59, 17, 18]. These criteria have also been applied for informative sensor selection [53, 21].

### 2.3.2 Fisher Information

Another perspective on how selected measurements or actions can affect an estimator's variance arises from considering Fisher information. We again consider the problem of estimating a vector $\mathbf{z} \in \mathbb{R}^d$ from a set of observations $\mathbf{Y} = [Y_0, \ldots, Y_{N-1}]^T$, related by an observation likelihood. An estimator $\hat{\mathbf{z}}$ maps from observations $\mathbf{Y}$ to a value of $\mathbf{z}$. This estimator can have several properties, and Section 2.3.1 describes how to select measurements that minimize the error variance when the maximum likelihood estimator is used.

In this section, we consider general admissible estimators. An admissible estimator depends only on the observed data $\mathbf{Y}$ and not on the unknown value of $\mathbf{z}$ i.e., the estimator is valid, and additionally constrained to be unbiased. Within the class of admissible estimators, the Cramér-Rao bound gives a lower bound on the error covariance $\mathbf{E}$ of any estimator $\hat{\mathbf{z}}$:

$$\mathbf{E} = \mathbb{E}_{\mathbf{Y}}\left[(\hat{\mathbf{z}}(\mathbf{Y}) - \mathbf{z})(\hat{\mathbf{z}}(\mathbf{Y}) - \mathbf{z})^T\right] \geq J_{\mathbf{Y}}(\mathbf{z})^{-1}. \tag{2.16}$$

Here, $J_{\mathbf{Y}}(\mathbf{z})$ is known as the Fisher information in $\mathbf{Y}$ about $\mathbf{z}$, and is a quantity determined by the latent vector of interest $\mathbf{z}$ and the selected measurements $\mathbf{Y}$. The Fisher information can be interpreted as a measure of "peaky-ness" of the likelihood of the measurements as a function of the latent quantity $\mathbf{z}$. The more peaky the likelihood function, the better an estimator can determine the value of $\mathbf{z}$ from observations $\mathbf{Y}$. For any admissible estimator, choosing measurements $\mathbf{Y}$ that maximize the Fisher information reduces the lower bound on the best achievable error variance.

Fisher information has been used as an information measure in robotic informative path planning and static target localization [109, 75]. Fisher information has also been used in conjunction with an RRT path planner to allow robots with complex dynamics to plan trajectories that are informative for static target localization [66, 67].

### 2.3.3 Information-theoretic reward

The final class of information measures commonly used in scientific information gathering applications are the information-theoretic reward functions, which include mutual information and conditional entropy. Unlike the experimental design criteria and the Fisher information metric, the information-theoretic measures assume a "soft" estimate of an unknown random variable $\mathbf{Z}$, i.e., a distribution $\mathbf{Pr}(\mathbf{Z} \mid \mathbf{Y})$, instead of a "hard" estimate $\hat{\mathbf{z}}(\mathbf{Y})$.

Conditional entropy measures the average log-loss of this inferred distribution $\mathbf{Pr}(\mathbf{Z} \mid \mathbf{Y})$ with respect to random variable $\mathbf{Z}$. In other words, conditional entropy measures the average uncertainty or randomness in a random variable $\mathbf{Z}$ after observing another random variable $\mathbf{Y}$:

$$H(\mathbf{Z} \mid \mathbf{Y}) \triangleq -\sum_{a,b} \mathbf{Pr}(\mathbf{Z} = a, \mathbf{Y} = b) \log \mathbf{Pr}(\mathbf{Z} = a \mid \mathbf{Y} = b). \tag{2.17}$$

Mutual information measures the average change in log-loss/entropy achieved by observing $\mathbf{Y}$, i.e., mutual information quantifies how much $\mathbf{Y}$ tells us about $\mathbf{Z}$.

$$\begin{aligned} I(\mathbf{Z}; \mathbf{Y}) &\triangleq H(\mathbf{Z}) - H(\mathbf{Z} \mid \mathbf{Y}) \\ &= \sum_{a,b} \mathbf{Pr}(\mathbf{Z} = a, \mathbf{Y} = b) \log \frac{\mathbf{Pr}(\mathbf{Z} = a, \mathbf{Y} = b)}{\mathbf{Pr}(\mathbf{Z} = a)\mathbf{Pr}(\mathbf{Y} = b)}. \end{aligned} \tag{2.18}$$

For models in which these information-theoretic measures are computable, choosing observations $\mathbf{Y}$ that minimize the conditional entropy of $\mathbf{Z}$ or maximize the mutual information between $\mathbf{Y}$ and $\mathbf{Z}$ will allow for soft estimators that minimize expected log-loss and reduce uncertainty/randomness in the resulting posterior distribution over $\mathbf{Z}$. Results such as Fano's inequality [23] relate the probability of error in guessing a discrete random variable $\mathbf{Z}$ to its conditional entropy $H(\mathbf{Z} \mid \mathbf{Y})$.

For this reason, information-theoretic measures have been employed as reward functions in a variety of information planning and scientific applications. Krause et al. [62, 7, 46, 63] have used mutual information and conditional entropy to select maximally informative locations for sensor placement and to plan informative trajectories

for robots. Mutual information has been applied for target tracking [93] and unknown map exploration [94]. Other work uses generalized entropy or mutual information for the problem of unknown map exploration and pose estimation [91, 15, 16].

Many applications of information-theoretic reward functions assume an environmental model for which information-theoretic measures are computable in closed form or easy to approximate. For general environmental models, mutual information and conditional entropy do not have closed form, because of dependence on the posterior predictive distribution. Zheng et al. [123] derive a sampling-based method for estimating mutual information in general models with performance guarantees. Arora et al. [3] apply a Monte Carlo sampling approach to estimate mutual information without performance guarantees in general Bayesian network environmental models. The utility of information-theoretic reward measures is evidenced by their widespread adoption in these diverse application areas. Chapter 3 will discuss information-theoretic reward functions and their application to scientific information-gathering missions in detail.

## 2.4 Approximate Planning Strategies

Once a sequential decision-making problem is formulated as a POMDP or MDP, solving this POMDP or MDP involves finding the optimal policy or plan, and is often nontrivial. There are a variety of planning approaches that differ in how online information about the decision-state of the autonomous agent is incorporated into planning decisions. In closed loop control, the agent computes a full policy and uses the new information gathered during execution about the state as it become available to select the optimal action in the current state. In contrast, in open loop control the agent computes a single plan instead of a full policy, consisting of a set of actions, and executes this plan regardless of what information about state is revealed to the agent online during execution. The middle ground between these extremes is open loop feedback control. As in open loop control, a plan is constructed over a finite horizon. However, it is only partially executed. After executing one or more steps of the initial

plan, the agent incorporates information received during execution into its decision state and replans. Like closed loop control, open loop feedback control is able to use new information received online during plan execution to improve performance.

Generally, closed loop control can provide optimal decision-making; however, finding the optimal closed loop policy can be computationally intractable [10, 85]. Open loop control is a computationally simpler alternative, but the performance gap between the solution of the optimal closed loop and open loop control problem can be arbitrarily large [120]. Open loop feedback control represents a compromise between the two extremes. Interestingly, for certain problems, it can be shown that open loop and closed loop solutions are equivalent, as will be seen in the following section and in Chapter 3.

Despite the difficulty of solving for the optimal closed loop policy, an extensive body of literature has explored both exact and approximate solvers for general POMDP and MDP frameworks. For MDPs, there are a variety of approaches based on dynamic programming and the optimal Bellman recursion, which solve for optimal or near optimal policies [70]. For POMDPs, offline POMDP solvers have shown promising results in closed loop control for small problems [54], and online solvers, such as those based on random Monte Carlo forward simulation [97], have recently seen compelling success by applying open loop feedback control to larger problems. These solvers are powerful tools for solving general purpose sequential decision-making problems. However, the planning problems considered in this thesis have special structure that allows for simpler approximation algorithms. In the following section, we discuss several common scenarios in which full, closed loop control can be avoided, while maintaining strong performance guarantees.

### 2.4.1 Open loop planning in Linear-Gaussian models

Although the difference in plan reward between open and closed loop controllers is generally unbounded, there are certain models for which these two planning strategies are, perhaps surprisingly, equivalent. The most canonical example of this problem structure is Linear-Gaussian systems, such as those employed in Kalman filtering [65]

and information-theoretic sampling [63]. In these Linear-Gaussian systems, when an autonomous agent collects observations and updates its belief on the decision-state, the posterior distribution is also Gaussian with a mean vector that depends on the realized observations but covariance matrix that *only depends on the location of the observations, not on their realized values.* This property of Linear-Gaussian models is exemplified in Eq. 2.9 (the posterior covariance of a Gaussian process); the posterior covariance depends on pairwise evaluations of the covariance function, which in turn depends only on the location of sample points.

For information-theoretic reward functions that depend only on the posterior covariance (information gain, entropy, uncertainty reduction, etc.), this property of Linear-Gaussian systems implies that the reward function does not depend on observations gathered online. For these problems, closed loop planners can perform no better than open loop planners [120]. A variety of previous works have utilized this property to plan optimal or near-optimal actions in Linear-Gaussian systems [63, 101].

### 2.4.2 Greedy algorithms with submodular reward

Even for Linear-Gaussian systems, solving open loop control problems can be exponentially difficult. Therefore, further work has sought approximate solutions to control problems, with performance that can be bounded with respect to the optimal open or closed loop solution. Perhaps the simplest approximate planner is a greedy, myopic planner. At each discrete time, this greedy planner selects the action that has the highest immediate reward or utility according to a reward function $R$. For most problems, the performance of this greedy algorithm is unbounded; it can be arbitrarily worse than the optimal, nonmyopic planner. However, certain reward functions have special structure which allows the greedy algorithm to be applied with bounded suboptimality guarantees. This structure is submodularity [82].

**Definition 1** (Submodularity) A set function $R : 2^{\mathcal{V}} \to \mathbb{R}$ is submodular if for every $A \subseteq B \subseteq \mathcal{V}$ and $e \in \mathcal{V} \setminus B$, $R(A \cup \{e\}) - R(A) \geq R(B \cup \{e\}) - R(B)$.

Submodularity formalizes the intuitive notion of diminishing returns: the benefit

you get from adding a new sample to a large set is less than the benefit you get from adding that new sample to a smaller subset. Another important property of set functions is monotonicity.

**Definition 2** (Monotonicity) A set function $R : 2^{\mathcal{V}} \rightarrow \mathbb{R}$ is monotone if for every $A \subseteq B \subseteq \mathcal{V}$, $R(B) \geq R(A)$ i.e., adding elements to a set will not decrease reward.

Monotone submodular reward functions have many beneficial properties: they can be minimized efficiently and near-optimal constrained maximization is possible in polynomial time. For example, sampling applications often require an algorithm to select a set of sampling points $\mathcal{A}$ from the full set of possible points $\mathcal{S}$, such that $|\mathcal{A}| = K$ and the reward $R(\mathcal{A})$ is maximized, where in general $|\mathcal{S}| \gg K$. Finding the optimal subset of points may require trying all $\mathcal{O}\left(\binom{|S|}{K}\right)$ possible subsets of size $K$. This problem can be shown to be NP$^{PP}$-complete [61]. However, the simple iterative, greedy approximation algorithm in Algorithm 1 can be shown to choose a sample set $\mathcal{A}$ such that $R(\mathcal{A}) \geq (1 - 1/e) \cdot R(\mathcal{A}^*)$, where $\mathcal{A}^*$ is the optimal set for monotone, submodular reward function $R$.

**Algorithm 1** Greedy set selection algorithm

**Input:** Reward function $R$, possible sampling locs $\mathcal{S} = \{\mathbf{x}_i\}$, sampling capacity $K$
**Output:** Sample set $\mathcal{A} \subseteq \mathcal{S}$

```
procedure GREEDY SET SELECTION
    A ← ∅
    while |A| ≤ K do
        x* = argmax_{x∈S\A} R_A(x)
        A ← A ∪ x*
    return A
```

**Property 1.** *Let $R : 2^{\mathcal{S}} \rightarrow \mathbb{R}$ be a be a monotone submodular set function, defined on subsets of domain $\mathcal{S}$. Define the greedy set selection algorithm as in Algorithm 1. The set $\mathcal{A}$ returned by a greedy set selection algorithm will have reward $R(\mathcal{A}) \geq (1 - 1/e) \cdot R(\mathcal{A}^*)$, where $\mathcal{A}^*$ is the optimal set for reward function $R$.*

*Proof.* Let $\mathcal{A}^* = \text{argmax}_{\mathcal{A}:|\mathcal{A}|\leq K} R(\mathcal{A})$. Let $\mathcal{A}_G$ be the set returned by Algorithm 1 and $\mathcal{A}_k \subseteq \mathcal{A}_G$ be the current set of $k$ observations sampled by the greedy algorithm

after $k$ iterations. Additionally, let $R_{\mathcal{A}_k}(\mathbf{x})$ be the marginal gain of adding sample $\mathbf{x}$ to a set $\mathcal{A}_k$ i.e., $R(\mathcal{A}_k \cup \mathbf{x}) - R(\mathcal{A}_k)$.

Following the general proof in [51]:

$$R(\mathcal{A}^*) \leq R(\mathcal{A}_{k-1}) + \sum_{\mathbf{x} \in \mathcal{A}^* \setminus \mathcal{A}_{k-1}} R_{\mathcal{A}_{k-1}}(\mathbf{x}) \tag{2.19}$$

$$\leq R(\mathcal{A}_{k-1}) + \sum_{\mathbf{x} \in \mathcal{A}^* \setminus \mathcal{A}_{k-1}} R(\mathcal{A}_k) - R(\mathcal{A}_{k-1}) \tag{2.20}$$

$$\leq R(\mathcal{A}_{k-1}) + K \cdot (R(\mathcal{A}_k) - R(\mathcal{A}_{k-1})). \tag{2.21}$$

The first line (2.19) follows directly from $R(\cdot)$ being a monotone submodular set function [51], the second (2.20) from the greedy property of Algorithm 1, and the third (2.21) because $|\mathcal{A}^*| \leq K$. Subtracting $K \cdot R(\mathcal{A}^*)$ from both sides:

$$R(\mathcal{A}_k) - R(\mathcal{A}^*) \geq \frac{K-1}{K}(R(\mathcal{A}_{k-1}) - R(\mathcal{A}^*)), \tag{2.22}$$

which implies by induction, with $R(\emptyset) = 0$:

$$R(\mathcal{A}_k) \geq \left(1 - \left(1 - \frac{1}{K}\right)^i\right) R(\mathcal{A}^*). \tag{2.23}$$

The proof is achieved by setting $k = K$, and using the identity $(1 - \frac{1}{K})^K \leq \frac{1}{e}$. ∎

This example demonstrates the power of submodular reward functions: by running a simple $\mathcal{O}(K \cdot |\mathcal{S}|)$ greedy algorithm, the recovered sample set has reward at least $(1 - 1/e)$ the reward of the optimal set, which requires solving an $\mathrm{NP}^{PP}$-complete combinatorial optimization problem to identify. Chapter 3 exploits this submodular structure to provide performance guarantees for informative secretary sampling problems in a similar context.

# Chapter 3

# Irrevocable Information Gathering in Periodic Data Streams

Chapter 1 introduced the problem of scientific information gathering from both a modeling and planning perspective and provided several diverse examples of scientific information gathering applications. This chapter focuses on the secretary sampling problem, as introduced in Chapter 1. We consider secretary sampling problems where the scientific phenomenon of interest can be modeled as a smooth, continuous function over a bounded domain and present a complete framework for environment and information-theoretic reward modeling in Section 3.3. Given this model, Section 3.5 introduces the *periodic secretary algorithm* for secretary sampling in spatiotemporally periodic data streams. By formulating an information-theoretic reward function and exploiting the submodular structure of this reward function, the periodic secretary algorithm performs irrevocable sample selection with performance guarantees in spatiotemporally correlated data streams, a domain in which previous secretary sampling approaches perform arbitrarily poorly.

The remainder of this chapter introduces the problem of secretary sampling in periodic data steams and provides several motivating applications (Section 3.1). Related work in secretary sampling problems is reviewed in Section 3.2. The specific scientific information gathering problem considered in this chapter is presented in Section 3.3, followed by an discussion of the appropriate choice of environmental model

and reward function. The planning problem is formalized in Section 3.4 and the proposed secretary sampling algorithm is presented in Section 3.5. Finally, this chapter concludes with theoretical analysis (Section 3.6) and experimental validation of the proposed method (Section 3.7). The work presented in this chapter has previously appeared in abbreviated form in Flaspohler et al. [35].

## 3.1 Introduction

Many interesting phenomena vary on spatial and temporal scales that are too large to monitor in their entirety. Attempting to gather information about these phenomena using limited representative samples is known as sample selection or experimental design [80]. In most problem formulations, samples are chosen to maximize some reward function while satisfying a fixed cost requirement: an autonomous underwater vehicle (AUV) may need to maximize the amount of phytoplankton in 10 collected water samples; a planetary rover may need to collect a maximally diverse set of rock samples that weigh less than 5 $kg$; a policy maker may wish to infer pollutant flow throughout a water body but only spend $\$10,000$ on water samples. Constrained sample selection problems arise in many real-world contexts, spanning domains from robotics to data mining to online auctions.

Sample selection problems can be divided into offline and streaming problems. In offline problems, potential sample locations are known ahead of time and an algorithm can make arbitrarily many passes through these locations to find the optimal placement of samples. In streaming problems, potential sample locations are revealed to the algorithm sequentially, and the algorithm must choose to collect or not collect a sample in real-time. Both the offline and streaming constrained sample selection problems are known to be at least NP-hard, but polynomial time approximation schemes exist for a variety of problem formulations [82].

One important variant of the streaming sample selection problem is the *secretary sampling problem*, which arises when an autonomous agent must choose to collect samples irrevocably in a data stream, i.e., the agent must decide in real-time whether

to collect a sample, cannot return to collect a sample at a previously rejected location, and cannot later reject a collected sample. This streaming, irrevocable-choice variant of the constrained sample selection problem arises frequently in real-time domains and is known as the secretary sampling problem because of parallels to the problem of hiring the best secretarial candidate from a stream of applicants [32].

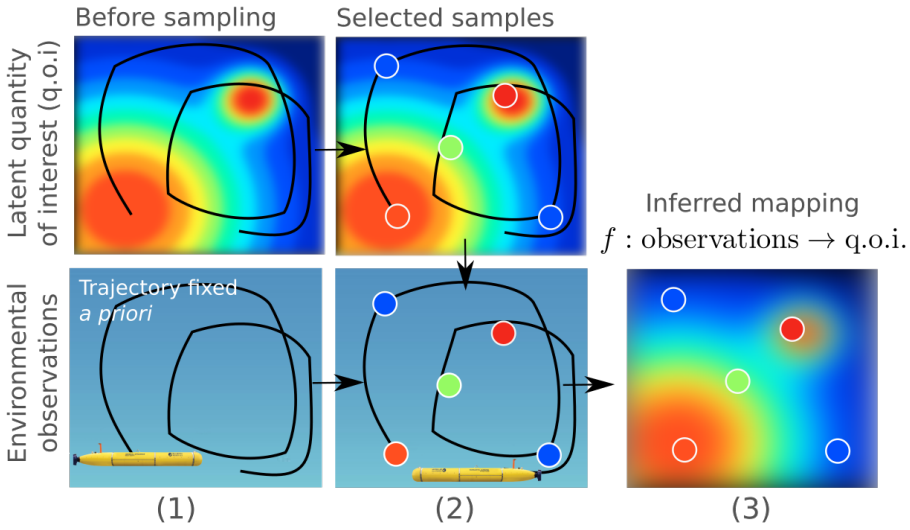


Figure 3-1: **Streaming irrevocable sample selection:** (1) In an example of streaming, irrevocable sample selection, an autonomous underwater vehicle must irrevocably collect representative water samples along a fixed trajectory at locations that are the most informative about a latent quantity of interest (q.o.i.), e.g. plankton concentrations. After observing the value of the quantity of interest at the sample locations (2), we can infer a mapping between environmental observations and the latent q.o.i. for later use (3).

For example, an AUV following a fixed trajectory through a marine environment may be equipped with $K$ single-use water samplers and need to collect the set of water samples that are the most informative about the distribution of a quantity of interest (q.o.i.) e.g. plankton species (Figure 3-1). The AUV can measure the surrounding environmental conditions and must decide to collect a plankton sample based on these environmental percepts, given a model relating these observable quantities to the latent plankton distribution. Finding the optimal set of locations to sample at along its trajectory without a prior map of the partially-observable environment is a hard problem: if the AUV collects samples too early, it will not be able to sample

the interesting locations it discovers later in the mission; if the AUV passes over interesting locations at the start of the mission, it may not see enough high quality locations later in the mission at which to collect samples.

The secretary problem has a long history and a variety of near-optimal solutions for different problem domains have been developed [32]. However, solutions to the secretary problem nearly always require that data are seen in random order. This stringent requirement is rarely met in robotics and real-time sensing domains, which produce spatially and temporally correlated data streams. In this work, we focus on data streams with periodic spatiotemporal structure. Periodic data arise commonly in environmental monitoring datasets due to natural cycles on a daily, lunar, and annual basis and in robotics tasks such as repetitive surveying. While it is easy to imagine adversarial data orderings for which a streaming, irrevocable-choice algorithm would perform arbitrarily badly, given data with known spatial or temporal structure, a carefully-designed algorithm may be able to provide performance guarantees. In this work, we introduce the *periodic secretary algorithm* to choose $K$ samples irrevocably from a data stream with periodic structure and provide a lower bound on the reward of the selected samples.

The contributions of this chapter include:

- We introduce the *periodic secretary algorithm*, which leverages spatiotemporal structure to choose samples from a periodic data stream according to any monotone submodular reward function with probabilistic performance guarantees.
- We demonstrate how information-theoretic reward functions can be leveraged for multiple-choice secretary sampling algorithms, in order to select samples that reduce posterior model uncertainty of a latent quantity of interest, as well as provide lower-bounds on sampling algorithm performance.
- We demonstrate the robustness of the periodic secretary algorithm on a data stream containing plankton observations from January 2009 to January 2016, and show that the plankton samples selected by the periodic secretary algorithm construct the most predictive model of the overall plankton concentrations.

## 3.2 Related Work

The problem of constrained sample selection has been given a thorough treatment in both the offline and streaming settings. In offline settings, previous work has explored using information-theoretic reward functions in spatiotemporally correlated data domains to select high utility samples. Nemhauser et al. [82] show that for submodular reward functions, a simple iterative greedy algorithm where the highest-reward sample given previous samples is selected at each iteration will produce a set with reward greater than $(1-1/e)$ times the reward of the optimal set (see Algorithm 1 in Chapter 2). Other works use this greedy algorithm along with Gaussian process (GP) models and information-theoretic reward functions to do offline sample selection [63] and to plan information-rich exploration paths for robots [7]. There is a rich body of literature in the spatial statistics community discussing optimal sensor placement in an offline setting [80] for a variety of reward functions and information measures.

On the other hand, secretary sampling algorithms remain largely constrained to simple reward functions and random arrival order assumptions. The *secretary problem* is a classical problem in stopping theory. In the standard problem formulation, a company wants to hire the most qualified secretarial candidate. The company interviews a stream of $N$ candidates and has to choose irrevocably after each interview whether or not to hire the candidate. When selecting a single maximal candidate, Lindley [68] provides a well known result: by observing the maximum reward sample in the first $1/e$ fraction of the stream and picking the first sample with higher reward, the highest reward sample will be selected in $1/e$ fraction of cases. It can be shown that this single secretary algorithm is the optimal strategy for this problem, given that the stream of candidates arrives in random order and a static reward function:

**Single secretary algorithm**: Given a stream of $N$ candidates, observe but do not accept the first $\lfloor N/e \rfloor$ candidates. Keep track of the maximum-reward sample for reward function $R$: $\mathbf{x}^* = \text{argmax}_{\mathbf{x} \in \mathbf{x}_0, \ldots, \mathbf{x}_{\lfloor N/e \rfloor - 1}} R(\mathbf{x})$ in the observation period. After the observation period, hire the first candidate $\mathbf{x}'$ whose score $R(\mathbf{x}') \geq R(\mathbf{x}^*)$.

This single secretary problem can be expanded to the case of hiring the $K$ most

qualified candidates, known as the multiple-choice secretary problem. In this case, the reward function must now score sets of secretarial candidates. The optimal set $\mathcal{A}^*$ for any arbitrary set function is generally NP-hard to find [58]. Babaioff et al. [4] introduce an $e$-competitive algorithm for the multiple-choice secretary problem and an alternative approach introduced by Kleinberg provides a $1/(1-5/\sqrt{k})$-competitive algorithm [57]. However, both algorithms require that data arrive in random order and "static" reward functions (i.e. the reward of a data point in the stream cannot change as new samples are selected).

For the special case of monotone submodular reward functions, Bateni et al. [5] provide a $7/(1-1/e)$-competitive algorithm known as the submodular secretary algorithm. The submodular secretary algorithm splits the data into $K$ segments of equal length (the last segment can be padded with "filler secretaries" if necessary) and runs the single secretary algorithm on each segment:

**Submodular secretary algorithm**: Given a stream of $N$ potential candidates, split the stream into $K$ equal length segments. In each of the $K$ segments, choose a single secretary using the single secretary algorithm, where the reward $R(\cdot)$ of a secretarial candidate is calculated with respect to the previously selected candidates.

The submodular secretary algorithm is applied by Luo et al. with an information-theoretic reward function to select samples irrevocably from an environmental data stream [72]. The use of information-theoretic reward functions for irrevocable sample selection was developed concurrently and independently in this work and by Luo et al. However, despite allowing flexibility in reward function, the submodular secretary algorithm and its application by Luo et al. require that data arrive in random order, a restriction that is relaxed in this work.

Kesselheim et al. [56] attempt to relax the assumption that data arrive in random order and define a class of distributions for which the assumption is violated but the performance of the standard secretary algorithm remains bounded. However, many spatiotemporally correlated data streams, including periodic data, do not satisfy even these relaxed constraints. Vardi [33] proposes a secretary algorithm for quasi-periodic data which arrive in random order. This algorithm requires that each observation

will appear exactly $m$ times in the data stream, an often unrealistic assumption in noisy data streams; indeed, this event occurs with probability zero in the continuous observation domains considered in this chapter.

Streaming, irrevocable-choice algorithms have been applied to select samples in environmental monitoring and robotics applications, even when the data streams violate random arrival order assumptions. Das et al. [26] apply the submodular secretary algorithm on-board an AUV to select $K$ water samples with the highest concentration of phytoplankton and use a GP model to predict these concentrations. However, they directly apply the submodular secretary algorithm, despite their data being spatially correlated, which could lead to arbitrarily poor sampling performance. Girdhar et al. [38] also deploy a modified multiple-choice secretary algorithm on an AUV to choose the most informative images to send back to a ground station. However, this approach is incompatible with the use of information-theoretic reward functions and requires that data arrive in random order. Recently, Manjanna et al. [73] used the submodular secretary algorithm to perform irrevocable sampling using a heterogeneous robot team, again despite the data stream having strong spatiotemporal correlation.

## 3.3 Scientific Model and Objective

In the general constrained sample selection problem, we must choose a set $\mathcal{A}$ consisting of $K$ sample locations from a finite set of possible locations in a $d$-dimensional environmental state space $\mathcal{V} \subseteq \mathbb{R}^d$, such that a reward set function $R : 2^{\mathcal{V}} \to \mathbb{R}$ is maximized:

$$\mathcal{A}^* = \text{argmax}_{\mathcal{A} \subseteq \mathcal{S} : |\mathcal{A}| = K} R(\mathcal{A}) \tag{3.1}$$

The full state space $\mathcal{V}$ is split into a set of locations where it is possible to collect samples $\mathcal{S}$ and a set where no samples can be collected $\mathcal{U} = \mathcal{V} \setminus \mathcal{S}$. The state space $\mathcal{V}$ can consist of geographic locations or locations in an environmental sensor space, e.g. temperature, salinity. In the offline setting, $\mathcal{S}$ and $\mathcal{U}$ are defined by accessibility, price, or other concerns. In the streaming setting, $\mathcal{S}$ consists of states encountered in the data stream, which are generally *a priori* unknown, necessitating algorithms

that choose sample locations reactively in the observed data stream.

### 3.3.1 Gaussian process environmental model

Our sampling objective is to select a set of $K$ sampling locations $\mathcal{A} \subseteq \mathcal{S}$ that contain the most information about a latent scientific quantity of interest that takes values in $\mathcal{V}$. Evaluating the information value of potential sampling locations requires a model of how a latent quantity of interest is correlated with environmental state. Given that the physical sensors in robotics and environmental monitoring domains are noisy, this model would ideally be probabilistic. Following [63], we use a Gaussian process model (GP), a nonparametric generalization of the multivariate Gaussian distribution (see Chapter 2). A GP model learns a regression model for the q.o.i. at new states in the data stream $\mathcal{S}$ based on a set of noisy q.o.i. samples at known states, and explicitly quantifies the uncertainty in these predictions.

Let $f : \mathcal{V} \rightarrow \mathbb{R}$ be an unknown function representing the value of a continuous environmental phenomenon of interest e.g. plankton concentration, as a function of current state $\mathbf{x} \in \mathcal{V}$, where $\mathbf{x}$ can be the physical location of a robot or the current environmental conditions such as temperature, salinity as perceived by a static sensor. Noisy observations $y$ of the function value at state $\mathbf{x}$ can be obtained by irrevocably choosing to sample in state $\mathbf{x}$, such that $y_i = f(\mathbf{x}) + \epsilon_i$ with $\epsilon_i \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_n^2)$, where $\sigma_n^2$ is determined by noise model of the environmental sensor. We represent belief on $f$ as a GP with mean $\mu(\mathbf{x})$ and covariance function $\kappa(\mathbf{x}, \mathbf{x}')$, such that $f \sim \mathcal{GP}(\mu(\mathbf{x}), \kappa(\mathbf{x}, \mathbf{x}'))$. Given a set of $t$ samples and their corresponding locations in observation space $\mathcal{D}_t = \{\mathbf{x}_i, y_i\}_{i=0}^{t-1}$, the posterior belief at a new state $\mathbf{x}' \in \mathcal{V}$ can be computed:

$$f(\mathbf{x}') \mid \mathcal{D}_t \sim \mathcal{N}(\mu_t(\mathbf{x}'), \sigma_t^2(\mathbf{x}')), \text{where} \tag{3.2}$$

$$\mu_t(\mathbf{x}') = \kappa_t(\mathbf{x}')^T (\mathbf{K}_t + \sigma_n^2 \mathbf{I})^{-1} \mathbf{y}_t, \tag{3.3}$$

$$\sigma_t^2(\mathbf{x}') = \kappa(\mathbf{x}', \mathbf{x}') - \kappa_t(\mathbf{x}')^T (\mathbf{K}_t + \sigma_n^2 \mathbf{I})^{-1} \kappa_t(\mathbf{x}'), \tag{3.4}$$

where $\mathbf{y}_t = [y_0, \ldots, y_{t-1}]^T$, $\mathbf{K}_t$ is the positive definite kernel matrix with $\mathbf{K}_t[i, j] =$

$\kappa(\mathbf{x}_i, \mathbf{x}_j)$ for all $\mathbf{x}_i, \mathbf{x}_j \in \mathcal{D}_t$, and $\kappa_t(\mathbf{x}') = [\kappa(\mathbf{x}_0, \mathbf{x}'), \ldots, \kappa(\mathbf{x}_{t-1}, \mathbf{x}')]$. Our experiments use a squared-exponential covariance function (Eq. 2.11) and a zero mean function; however our results hold for any covariance function.

### 3.3.2 Reward functions and tradeoffs

A variety of reward functions appear in the sample selection literature, including maximizing the sum of utilities of the collected samples [4], maximizing the minimum distance between samples [38, 124], maximizing the reduction in entropy $H(\cdot)$ over $\mathcal{V}$, known as the entropy criterion [63]:

$$R_H(\mathcal{A}) = -H(\mathcal{V} \setminus \mathcal{A} \mid \mathcal{A}) = H(\mathcal{A}) - H(\mathcal{V}), \tag{3.5}$$

or maximizing the mutual information $I(\cdot;\cdot)$ between sampled locations and the rest of the observation space, known as the mutual information criterion [63]:

$$R_I(\mathcal{A}) = I(\mathcal{V} \setminus \mathcal{A}; \mathcal{A}) = H(\mathcal{V} \setminus \mathcal{A}) - H(\mathcal{V} \setminus \mathcal{A} \mid \mathcal{A}). \tag{3.6}$$

The entropy and mutual information reward functions directly quantify how useful a sample will be for the task of inference about a quantity of interest that is distributed across the observation space. These information-theoretic reward functions have been widely used to decide optimal placements of sensors in the kriging and spatial statistics literature [80]. The mutual information criterion seeks to maximize the mutual information between a set of sampled locations $\mathcal{A}$ and the rest of the observation space $\mathcal{V} \setminus \mathcal{A}$. Intuitively, the mutual information criterion reflects how informative the sampled locations are about the rest of the space. However, calculating the mutual information criterion requires a model of the entire observation space $\mathcal{V}$ and generally requires $O(|\mathcal{V}|^3)$ operations to compute a single time. This can be challenging or impossible to compute in streaming contexts.

The entropy criterion seeks simply to maximize the reduction in entropy over the observation space by maximizing the entropy of the selected sample set $\mathcal{A}$. The en-

tropy criterion does not depend on knowledge of the entire observation space and can be calculated in $O(K^3)$ operations, where $K$ is maximum cardinality of the selected sample set $\mathcal{A}$. For Gaussian process models, maximizing the entropy criteria can be shown to be equivalent to maximizing the information gain of a set of sample locations with respect to the latent quantity of interest $f$ [101]:

$$R_{\text{IG}}(\mathcal{A}) = I(\mathcal{A}; f) = H(\mathcal{A}) - H(\mathcal{A} \mid f). \tag{3.7}$$

Despite recent use of the mutual information criteria for offline sensor selection [63], for streaming applications run on computationally constrained devices, the entropy or information gain criterion remain a computationally efficient alternative.

### 3.3.3 Submodular set functions

For an arbitrary sample set reward function $R : 2^{\mathcal{V}} \to \mathbb{R}$, the maximization problem in Eq. (3.1) is NP-hard for both the offline and streaming scenarios [58]. Fortunately, many commonly used reward functions, including the entropy criterion [102, 95], have special structure that allows near-optimal polynomial time approximation schemes. This structure is submodularity:

**Definition 1** (Submodularity) A set function $R : 2^{\mathcal{V}} \to \mathbb{R}$ is submodular if for every $A \subseteq B \subseteq \mathcal{V}$ and $e \in \mathcal{V} \setminus B$, $R(A \cup \{e\}) - R(A) \geq R(B \cup \{e\}) - R(B)$.

Submodularity formalizes the intuitive notion of diminishing returns: the benefit from adding a new sample to a large set is less than the benefit from adding that new sample to a smaller subset.

Another important property of set functions is monotonicity:

**Definition 2** (Monotonicity) A set function $R : 2^{\mathcal{V}} \to \mathbb{R}$ is monotone if for every $A \subseteq B \subseteq \mathcal{V}$, $R(B) \geq R(A)$, i.e. adding elements to a set will not decrease reward.

Monotone submodular reward functions have many beneficial properties: they can be minimized efficiently and near-optimal constrained maximization is possible in polynomial time. We will exploit this structure to provide performance guarantees for periodic irrevocable sample selection using the entropy criterion.

## 3.4 Secretary Sampling Problem Formulation

### 3.4.1 Periodic data streams

Let the dataset $\mathcal{S} = \{\mathbf{x}_i\} \subseteq \mathcal{V}$ be a stream of potential sampling states, such that the sensor is at state in observation space $\mathbf{x}_i \in \mathcal{V}$ at time step $i$. We assume the robot or autonomous sensor is passive, such that the sensing agent has little or no control over the states visited in the data stream. Many sensing platforms operate within this domain: static sensors are completely stationary and many mobile scientific robots have a very restricted set of motion primitives or are restricted to follow a preset trajectory due to safety constraints or lack of localization/planning capabilities. Additionally, even fully mobile scientific robots are often restricted to motions given by a primary sensing objective; any secondary objectives can only be satisfied by passively choosing when to collect samples.

For a state $\mathbf{X}_i$, let $Y_i$ be the corresponding latent quantity of interest (q.o.i.) value at time step $i$, which cannot be measured *in vivo* but can be sampled for offline analysis. We define a state data stream $\mathcal{S}$ to be approximately periodic with period $T$ and noise $\Sigma_d$ if the (possibly vector-valued) state $\mathbf{X}_i$ at index $i$ is drawn i.i.d. from a Gaussian distribution with mean equal to the state $\mathbf{x}_{i \bmod T}$ and covariance $\Sigma_d$ i.e., $\mathbf{X}_i \sim \mathcal{N}(\mathbf{x}_{i \bmod T}, \Sigma_d)$ for $i \geq T$ (Figure 3-2). This model of periodic data formalizes the intuition that states visited by a robot or autnomous sensor in a periodic data stream will be noisy copies of states visited in the initial period. Importantly, the reward, e.g. entropy, of approximately periodic states will also be approximately periodic with scalar reward noise $\sigma_u^2$ and the same period $T$.

**Observation 1.** If a stream of visited states are approximately periodic with period $T$ and noise $\Sigma_d$, then the reward of these states will also be approximately periodic with period $T$ and some reward noise $\sigma_u^2$. This follows directly because $\mathbf{x}_i = \mathbf{x}_j$ implies $R(\mathbf{x}_i) = R(\mathbf{x}_j)$ for all $\mathbf{x}_i, \mathbf{x}_j \in \mathcal{V}$ and any deterministic function $R(\cdot)$. Generally, $\Sigma_d \neq \sigma_u^2$, but $\sigma_u^2$ is a function of $\Sigma_d$. This is visualized in Figure 3-3.

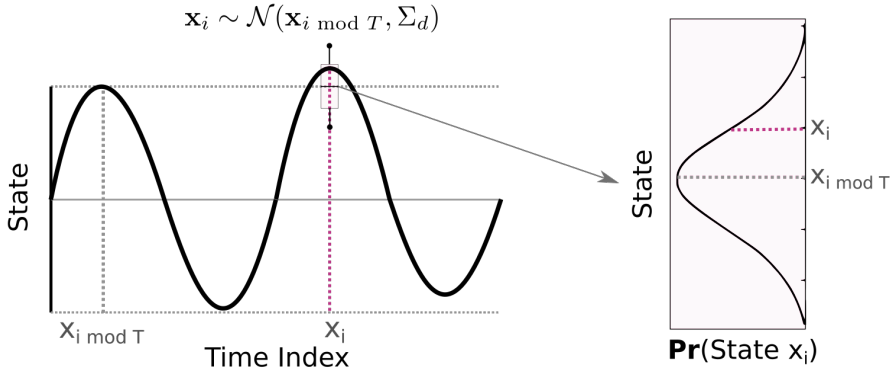


Figure 3-2: **Approximately periodic data:** Our algorithm assumes the stream of visited states are approximately periodic with period $T$ and noise $\Sigma_d$, where the state $\mathbf{X}_i$ at index $i$ is drawn i.i.d. from a Gaussian distribution with mean equal to the state $\mathbf{x}_{i \bmod T}$ and covariance $\Sigma_d$ i.e., $\mathbf{X}_i \sim \mathcal{N}(\mathbf{x}_{i \bmod T}, \Sigma_d)$

### 3.4.2 Secretary sampling problem

Let $\mathcal{A}_m \subseteq \mathcal{S}_i$ be the set of $m$ states we have sampled at, from the first $i$ possible states in the data stream. At time step $i+1$, we must irrevocably decide whether to add a sample in that state to the sample set $\mathcal{A}_m$ based on the value of the entropy criterion/reward function at that state. To calculate the entropy criterion at a potential sampling location $\mathbf{x}_{i+1}$, we must calculate the conditional entropy of the posterior distribution on the latent quantity of interest in that state, given the locations in the state space of previously collected samples. In a GP model, we can calculate the differential entropy at state $\mathbf{x}_{i+1}$ in closed form:

$$h(\mathbf{x}_{i+1} \mid \mathcal{A}_m) = \frac{d}{2}\ln(2\pi e) + \frac{1}{2}\ln(\sigma^2_{\mathbf{x}_{i+1}} \mid \mathcal{A}_m) \tag{3.8}$$

where $d$ is the dimension of $\mathcal{V}$ and $\sigma^2_{\mathbf{x}_{i+1}} \mid \mathcal{A}_m$ is the conditional variance of the GP model at point $\mathbf{x}_{i+1}$ (Eq. 3.4). Crucially, $\sigma^2_{\mathbf{x}_{i+1}} \mid \mathcal{A}_m$ depends solely on the covariance function used in the GP model and the locations of the samples $\mathcal{A}_m$ in the state space, not on the sampled quantity of interest values at these locations. This important property of GPs allows us to do streaming entropy calculations even if we are unable to observe the value $y_{i+1}$ of a sample at state $\mathbf{x}_{i+1}$ until post-processing. For our model, we use a squared exponential (SE) covariance function with maximum

likelihood parameters estimated from a previous deployment. Although our data are periodic, we do not use a periodic covariance function [74]. The periodic covariance function is used in data domains where the latent q.o.i. is periodic but the set of visited states are not; the SE covariance function is sufficient for our model because both the trajectory through state space and the latent q.o.i are assumed to be periodic with the same period.

After a sampling mission is completed and samples have been collected at various locations in state space we can, for example, use the resulting dataset $\mathcal{D} = \{(\mathbf{x}_i, y_i) \mid \mathbf{x}_i \in \mathcal{S}\}$, $|\mathcal{D}| = K$, consisting of states $\mathbf{x}_i$ and noisy quantity of interest samples at those states $y_i$, to predict the distribution of the latent quantity of interest at unsampled locations in the state space (Figure 3-1), using the analytical solution for the conditional predictive mean and variance of a GP.

## 3.5 Periodic Secretary Sampling Algorithm

Periodic phenomena occur ubiquitously in biological domains due to natural cycles on a daily, monthly, and annual basis and in repetitive robotics tasks. Given a Gaussian process model that allows us to compute the entropy reward function for states in our data stream, we propose a novel variant of the multiple-choice secretary algorithm for data with approximately periodic structure.

Assuming that the period $T$ of an approximately periodic data stream is known or can be estimated, the proposed periodic secretary algorithm consists of two stages. During the initial observation period, the first $T$ visited states in the data stream are saved into a reference set $U_R$ but no samples are collected. Then, for the remainder of the data stream, the algorithm attempts to iteratively collect a sample at the next state in the stream with the highest reward given previously selected samples. This is difficult to achieve without knowledge of the future states that will be visited in the data stream. However, for approximately periodic data streams as defined in Section 3.4, our algorithm can exploit the information it gathers during the initial observation period to make informed decisions about when to sample in the remainder

of the stream.

To select the next state from which to sample, the periodic secretary algorithm computes the reward of each previously visited state in the reference set $U_R$ and finds the state(s) with the highest reward in the reference set given previously selected samples. Then, the algorithm collects a sample at the next state in the data stream with reward greater than the maximum reward state in the reference set, minus some constant threshold parameter $\lambda$ that accounts for noise in the periodic function. In a sense that we derive explicitly in Section 3.6, we can expect to see an state of sufficient reward with high probability because our data are approximately periodic. Given this new sample, the reward of states in the reference set may have changed. We find the new maximum reward state in the reference set conditioned on the new sample set, and select to sample at the next state observed in the data stream with reward within some $\lambda$ of this maximum. This procedure repeats until $K$ samples have been collected or the end of the data stream is reached. The procedure is formalized in Algorithm 2 and depicted visually in Figure 3-3. We discuss the effect of the parameter $\lambda$ on the algorithm's performance in Section 3.7.

**Algorithm 2** Periodic secretary algorithm

**Input:** Reward function $R$, data stream $\mathcal{S} = \{\mathbf{x}_i\}$, sampling capacity $K$, data period $T$, parameter $\lambda \in \mathbb{R}$

**Output:** Sample set $\mathcal{A} \subseteq \mathcal{S}$

**procedure** PERIODIC SECRETARY ALGORITHM
    $\mathcal{A} \leftarrow \emptyset$
    $U_R \leftarrow \{R(\{\mathbf{x}_i\}),\ \text{for } i \in [0, T)$
    threshold $\leftarrow \max(U_R) - \lambda$
    **for each** $i \in [T, \ldots, N]$ **do**
        **if** $R(\{\mathbf{x}_i\} \cup \mathcal{A}) \geq$ threshold **then**
            $\mathcal{A} \leftarrow \mathcal{A} \cup \mathbf{x}_i$
            **if** $|\mathcal{A}| = K$ **then return** $\mathcal{A}$
            $U_R \leftarrow \{R(\{\mathbf{x}_i\} \cup \mathcal{A})\},\ \text{for } i \in [0, T)$
            threshold $\leftarrow \max(U_R) - \lambda$
    **return** $\mathcal{A}$

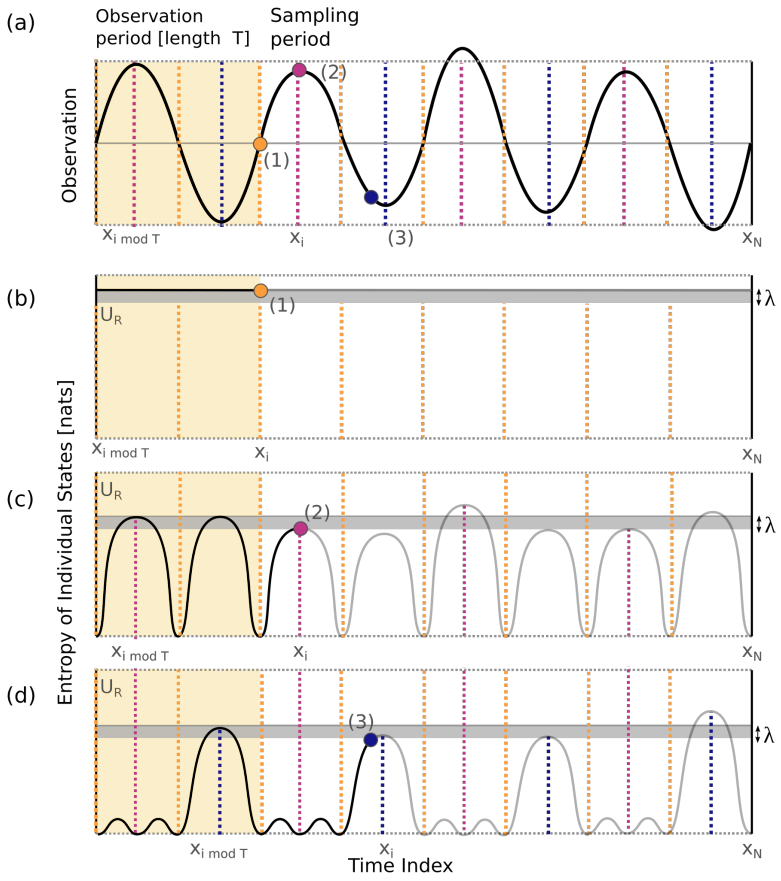


Figure 3-3: **Periodic secretary algorithm with threshold parameter $\lambda$: (a)** An approximately periodic data stream with known period. Three samples are selected using the periodic secretary algorithm; the sample selection process is visualized in (b-d). **(b)** Previous states are in black; unknown future states are in grey. When the algorithm begins and before any samples have been selected, every subsequent state has equal entropy [reward], hence the algorithm chooses the first state after the reference set $U_R$ as the first location to sample. **(c)** Given the first sample (1), the reward function is approximately periodic. The algorithm then samples in the first state with entropy [reward] $\geq$ the maximum entropy state in the reference set $U_R$ minus $\lambda$ (2). **(d)** Given samples at (1,2), the next state with entropy [reward] $\geq$ the maximum entropy state in $U_R$ minus $\lambda$ occurs at (3).

## 3.6 Analysis of Algorithm Performance

In this section, we analyze the performance of the periodic secretary algorithm as a function of the variables in our model: the reward noise $\sigma_u^2$, the number of periods in the data $\lfloor N/T \rfloor$, the number of samples selected $K$, and threshold parameter $\lambda$. We show that when the number of periods in the dataset is large compared to the

number of samples selected, the gap between the performance of the periodic secretary algorithm and the optimal solution grows slowly with the length of the dataset as $O(\sqrt{\log \lfloor N/T \rfloor})$. When the number of samples is much larger then the number of periods, however, our lower bound decreases quickly as $K$ grows, as $O(\lfloor N/T \rfloor / K)$. Although the algorithm will likely outperform this bound for specific reward functions, this is the tightest bound we could derive for general reward functions and is commensurate with bounds provided by e.g. the submodular secretary algorithm [5]. These conclusions follow directly from Theorem 1, proven at the end of this section. However, we first provide the following three useful lemmas.

Let $\mathcal{A}^*$ be the optimal sample set according to Eq. (3.1) and $\mathcal{A}$ be the set returned by the periodic secretary algorithm. We refer to the first $T$ states visited in the stream as the reference set $U_R$. Let $\mathcal{A}_m \subseteq \mathcal{S}$ be the current set of $m$ states where samples have been collected by the algorithm and $R_{\mathcal{A}_m}(\mathbf{x})$ be the marginal gain of adding a sample at state $\mathbf{x}$ to set $\mathcal{A}_m$, i.e., $R(\mathcal{A}_m \cup \mathbf{x}) - R(\mathcal{A}_m)$.

**Lemma 1.** In each iteration of the periodic secretary algorithm, the reward of the sampling state selected by the periodic secretary algorithm $\mathbf{x}_s^*$ from approximately periodic data of length $N$ with period $T$ and reward noise $\sigma_u^2$, given a previously selected sample set $\mathcal{A}_m$, is lower bounded by:

$$R_{\mathcal{A}_m}(\mathbf{x}_s^*) \geq \mathbb{E}[R_{\mathcal{A}_m}(\mathbf{X}^*)] - \left(\lambda + \sqrt{2\sigma_u^2 T^2 \log \left\lfloor \frac{N}{T} \right\rfloor}\right), \tag{3.9}$$

where $\mathbf{X}^*$ is the state with globally maximum reward.

Lemma 1 bounds how suboptimal the sampling location selected by the periodic secretary algorithm can be compared to the globally optimal sampling location, given the current sample set $\mathcal{A}_m$.

**Proof of Lemma 1.** Given that the maximum reward state in the reference set $\mathbf{x}_r^*$ occurs at index $i$, we compute the expected difference in reward between $\mathbf{x}_r^*$ and the maximum reward sample in the entire data stream $\mathbf{X}^*$ when $\mathbf{X}^*$ occurs at index $i+nT$ for some $n$, $0 \geq n \geq \left\lfloor \frac{N}{T} \right\rfloor$. Because our data are approximately periodic, we know that $R_{\mathcal{A}_m}(\mathbf{X}_{i+nT}) \sim \mathcal{N}(R_{\mathcal{A}_m}(\mathbf{x}_r^*), \sigma_u^2)$ for $n = \{0, \ldots, \left\lfloor \frac{N}{T} \right\rfloor\}$ and the global maximum

$\mathbf{X}^* = \max\{\mathbf{X}_{i+nT} \mid n = 0, \ldots, \lfloor \frac{N}{T} \rfloor\}$ i.e., the maximum of $\lfloor \frac{N}{T} \rfloor$ i.i.d. draws from a normal distribution with mean $R_{\mathcal{A}_m}(\mathbf{x}_r^*)$ and variance $\sigma_u^2$. Therefore, the expected difference between $R_{\mathcal{A}_m}(\mathbf{X}^*)$ and $R_{\mathcal{A}_m}(\mathbf{x}_r^*)$ conditioned on the event that $\mathbf{X}^*$ occurs at index $i + nT$ is no larger than the expectation of the maximum of $\lfloor \frac{N}{T} \rfloor$ samples drawn from a mean-zero Gaussian [55]:

$$\mathbb{E}[R_{\mathcal{A}_m}(\mathbf{X}^*) - R_{\mathcal{A}_m}(\mathbf{x}_r^*)] \leq \sqrt{2\sigma_u^2 \log \left\lfloor \frac{N}{T} \right\rfloor}. \tag{3.10}$$

The expected difference between the global maximum value in the stream and the maximum in the reference set will be maximized when the maximum occurs at index $i + nT$. Therefore, we can bound the unconditional expected difference between the global maximum and the maximum in the reference set as $T$ times Eq. 3.10. The final form of Lemma 1 arises from the observation that the sample the algorithm selects $\mathbf{x}_s^*$ will have expected reward $R_{\mathcal{A}_m}(\mathbf{x}_s^*) = R_{\mathcal{A}_m}(\mathbf{x}_r^*) - \lambda$.

**Lemma 2.** A set $\mathcal{A}$ of $K$ samples chosen according to the periodic secretary algorithm will have reward:

$$\mathbb{E}[R(\mathcal{A})] \geq \left(1 - \frac{1}{e}\right)\left(R(\mathcal{A}^*) - K \cdot \left(\lambda + \sqrt{2\sigma_u^2 \log \left\lfloor \frac{N}{T} \right\rfloor}\right)\right). \tag{3.11}$$

Lemma 2 states that a set of $K$ samples chosen with suboptimality bounded as in Lemma 1 also has bounded suboptimality.

**Proof of Lemma 2.** Following the general proof in [51]:

$$R(\mathcal{A}^*) \leq R(\mathcal{A}_{m-1}) + \sum_{\mathbf{x} \in \mathcal{A}^* \setminus \mathcal{A}_{m-1}} R_{\mathcal{A}_{m-1}}(\mathbf{x}) \tag{3.12}$$

$$\leq R(\mathcal{A}_{m-1}) + \sum_{\mathbf{x} \in \mathcal{A}^* \setminus \mathcal{A}_{m-1}} R(\mathcal{A}_m) - R(\mathcal{A}_{m-1}) + c \tag{3.13}$$

$$\leq R(\mathcal{A}_{m-1}) + k \cdot (R(\mathcal{A}_m) - R(\mathcal{A}_{m-1}) + c), \tag{3.14}$$

where $c = \lambda + \sqrt{2\sigma_u^2 \log \lfloor \frac{N}{T} \rfloor}$. The first line (3.12) follows directly from $R(\cdot)$ being a monotone submodular set function [51], the second (3.13) from Lemma 1, and the

third (3.14) because $|\mathcal{A}^*| \leq K$. Subtracting $K \cdot R(\mathcal{A}^*)$ from both sides:

$$R(\mathcal{A}_m) - R(\mathcal{A}^*) \geq \frac{K-1}{K}(R(\mathcal{A}_{m-1}) - R(\mathcal{A}^*)) - c, \tag{3.15}$$

which implies by induction, with $R(\emptyset) = 0$:

$$R(\mathcal{A}_m) \geq \left(1 - \left(1 - \frac{1}{K}\right)^m\right)\Big(R(\mathcal{A}^*) - K \cdot c\Big). \tag{3.16}$$

Lemma 2 is achieved by setting $m = K$, and using the identity $(1 - \frac{1}{K})^K \leq \frac{1}{e}$.

Lemma 2 assumes that the periodic secretary algorithm succeeds in sampling $K$ times, as will be the case when the reward noise $\sigma_u^2$ is small and the length of the data stream is large. However, given a finite data stream of length $N$, it is possible to fail to select all $K$ samples.

**Lemma 3.** In an approximately periodic data stream with period $T$ and reward noise $\sigma_u^2$ of length $N$, the expected number of samples selected by the periodic secretary algorithm is:

$$\mathbb{E}[\#Success] \geq \min\left(K, Q(-\lambda/\sigma_u^2)\left\lfloor\frac{N}{T}\right\rfloor\right). \tag{3.17}$$

**Proof of Lemma 3.** The probability of encountering a state in period $n$ of the data which meets or exceeds the reward threshold for a given iteration of the periodic secretary algorithm and is therefore sampled is:

$$\begin{aligned}\mathbf{Pr}(Success) \geq& \mathbf{Pr}\Big(R(\mathbf{X}_{i+nT}) \geq R(\mathbf{x}_r^*) - \lambda\Big)\\ \geq& Q(-\lambda/\sigma_u^2),\end{aligned} \tag{3.18}$$

where $Q(\cdot)$ is the standard Gaussian tail probability. In a data stream of length $N$, there are $\lfloor\frac{N}{T}\rfloor$ total periods and the expected number of successes is the number of periods multiplied by the probability of success in each period.

**Theorem 1.** Given a sample set $\mathcal{A}$ selected by the periodic secretary algorithm from a data stream of length $N$ that is approximately periodic with period $T$ and reward

noise $\sigma_u^2$, the expected reward of $\mathcal{A}$ is less than the reward of the optimal set $\mathcal{A}^*$ by a factor which depends the number of samples selected $K$ and parameter $\lambda$:

$$\begin{aligned}\mathbb{E}[R(\mathcal{A})] \geq & \frac{\min(k,\, Q(-\lambda/\sigma_u^2)\left\lfloor \frac{N}{T} \right\rfloor)}{K} \cdot \left(1 - \frac{1}{e}\right) \\ & \left(R(\mathcal{A}^*) - K \cdot \left(\lambda + \sqrt{2\sigma_u^2 T^2 \log \left\lfloor \frac{N}{T} \right\rfloor}\right)\right),\end{aligned} \tag{3.19}$$

where $Q(\cdot)$ is the standard Gaussian tail probability.

**Proof.** In Lemma 2, we showed that a set of $K$ samples selected using the periodic secretary algorithm has bounded suboptimality. In practice, for finite data streams, it is possible successfully sample less than $K$ times. Lemma 3 derives the expected number of samples the periodic secretary algorithm will select in a data stream of length $N$. Combining Lemma 2 and 3 with the observation that for a monotone submodular function, the value of the first $a$ samples of $\mathcal{A}$ have reward of at least $\left\lfloor \frac{a}{K} \right\rfloor R(\mathcal{A})$, $a \leq K$, the expected reward of set $\mathcal{A}$ is given by Theorem 1. ∎

## 3.7 Experiments

### 3.7.1 Using simulation to tune algorithm parameter $\lambda$

The submodular secretary algorithm has one tunable parameter $\lambda$ that mediates the trade-off between selecting more, lower quality samples and selecting fewer, higher quality samples in a noisy data stream. Generally, for large $\lambda$, the expected number of samples selected will grow to $K$, but the reward of the selected samples will decrease. Smaller $\lambda$ will cause the samples in $\mathcal{A}$ to be closer to their optimal reward values, but the algorithm may fail to select all $K$ samples in a noisy, short data stream. Generally, $\lambda$ should be tuned to maximize Eq. (3.19) based on the noise parameters of the periodic phenomena and the length of the data stream. We believe that it may be possible to do this maximization in closed form, but leave this as an open question for future work. It is also possible to tune $\lambda$ empirically by simulating data drawn from the periodic phenomena using the known period and periodic noise values and

then selecting the $\lambda$ which produces the largest average reward across these simulated data streams. We demonstrate this process using samples drawn from an arbitrary approximately periodic function for nine different values of $\lambda$ in Figure 3-4.

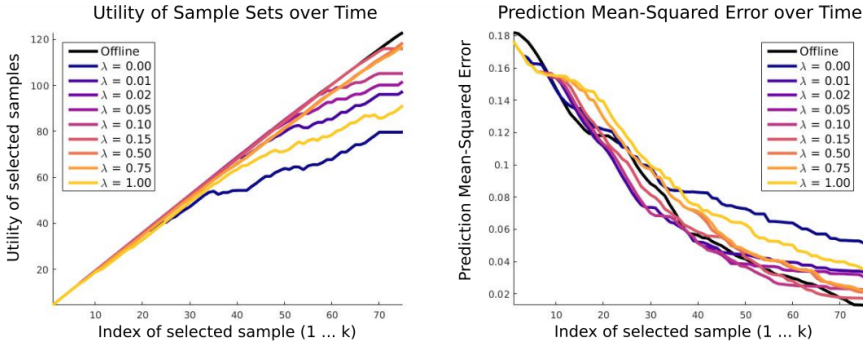


Figure 3-4: **Tuning parameter $\lambda$:** The reward and resulting model quality – as measured by prediction mean-squared error on a held-out test set – of samples sets selected using the periodic secretary algorithm on the data stream $\mathbf{x}_t = \sin(2\pi t) + \sin(3\pi t)$ and periodic noise $\sigma_d^2 = 0.35$ for nine different values of $\lambda$, $\lfloor N/T \rfloor = 10$, and $K = 75$ with the entropy criterion. For small $\lambda$, the algorithm chooses high reward samples, but is unable to successfully sample $K$ times. For medium $\lambda$, the algorithm selects $K$ samples with reward very near that of the offline upper bound. For large $\lambda$ the algorithm successfully samples $K$ times, but the samples are of low reward. For this dataset, $\lambda$ should be set to 0.50 for best performance.

### 3.7.2 MVCO experiments

We apply the periodic secretary algorithm with the entropy reward function to select water samples from a stream of potential sampling states observed by a marine sensor on the Martha's Vineyard Coastal Observatory from January 2009 to January 2016 [84]. This stationary marine sensor is equipped with $K$ single-use water samplers. The scientific objective is to collect water samples in environmental states that give the best understanding of the seasonal dynamics of the plankton species *Guinardia flaccida*.

The prevalence of this plankton species is known to vary with time of year (it is a winter blooming plankton) and water temperature (during warm winters, the species tends to be more numerous than during cold winters). However, the sensor is unable to measure the plankton present in the water stream in real-time. Instead, the

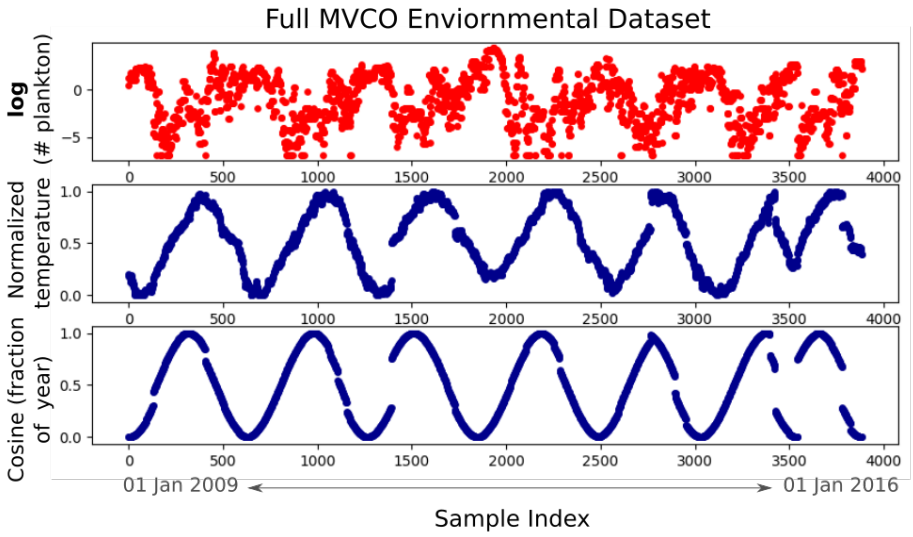


Figure 3-5: **MVCO environmental dataset:** The environmental dataset collected on the Martha's Vineyard Coastal Observatory from 01 January 2009 to 01 January 2016, averaged over half-day segments. The platform was equipped with the IFCB device [84], which allowed ground truth *Guinardia flaccida* concentrations to be measured (red). Only the periodic environmental data (blue) are available to the sample selection algorithms.

sensor can measure the temperature of the surrounding water and the day of year, and must decide to collect a sample based on these environmental covariates. In this stationary setting, the sensor is not choosing sample locations in geographic space. Instead, throughout its deployment, the sensor will observe a stream of points in this environmental space, and must choose to take water samples in the environmental conditions which are the most informative about the plankton species of interest. This seven-year dataset and ground truth *Guinardia flaccida* counts (unknown to the algorithm) are shown in Figure 3-5.

Given that these environmental data are known to be periodic on an annual basis, we apply the periodic secretary algorithm to select 84 samples (equivalent to 12 samples per year for seven years using a scheduled sampler) from the data stream using the entropy criterion with a GP model. We also select sample locations using the submodular secretary algorithm [5], a scheduled sampling algorithm commonly used in practical sensing deployments (sampling every $\frac{N}{k}$ samples), and random sampling as baselines. We use the offline greedy algorithm [82] to provide an upper bound.

### 3.7.3 MVCO sampling results

The selected samples for the periodic secretary algorithm and baseline algorithms are shown in Figure 3-6 along with the complete dataset colored by the ground-truth plankton counts. The periodic secretary algorithm selects samples which provide the most dense coverage of the observation space. The quality of plankton count predictions in unknown environmental conditions will depend on having sampled a nearby point in {temperature, cos(fraction of year)} space. Large gaps in the sampled locations will cause lower entropy reduction and poorer predictions at those locations; these gaps are evident in the submodular, scheduled and random sampling strategies.

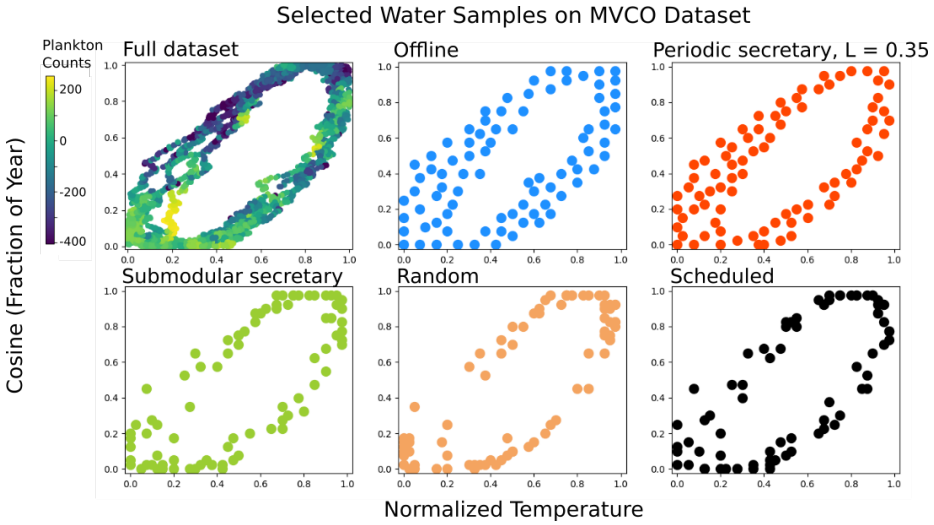


Figure 3-6: **Samples selected from the MVCO dataset:** The {temperature, cos(fraction of year)} samples selected from the full dataset (upper left, colored by the ground-truth plankton counts). The periodic secretary algorithm chooses samples which provide the most dense coverage of the environmental observation space. The quality of predictions in unknown environmental conditions will depend on having sampled a nearby point in {temperature, cos(fraction of year)} space. Gaps in the sample coverage, such as those seen in the bottom three plots, will cause large uncertainty and poor predictions of plankton counts in those regions.

To quantify this result, we compare the entropy reduction achieved by samples selected using the periodic secretary algorithm to samples selected by the baselines and the offline upper bound (the entropy reduction should be maximized). The mean reward and one standard deviation values for each algorithm are shown in Figure 3-7(a) for 50 random permutations of the yearly data in the MVCO dataset. For

small sample sets, all six algorithms produce similar results. After selecting around 30 samples, the periodic secretary algorithms begin to surpass the other streaming algorithms. The submodular secretary algorithm, which represents the current state-of-the-art in streaming, irrevocable sample selection for information-theoretic reward functions, never does significantly better than a scheduled algorithm. After selecting 70 samples, the periodic secretary algorithm with poorly tuned $\lambda$ reaches the end of the stream without selecting all $K = 84$ samples. The periodic secretary algorithm with well-tuned $\lambda$ stays close to the upper bound set by the offline algorithm.

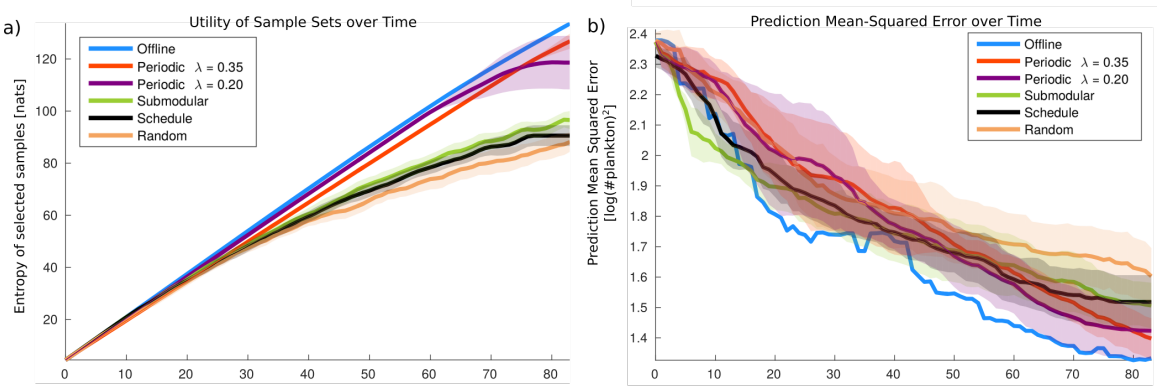


Figure 3-7: **Quantitative results on the MVCO dataset:** The mean value and standard deviation across the 50 runs of the periodic secretary algorithm on random permutations of the yearly data in the MVCO dataset. (a) The entropy reduction achieved as each of the $K = 84$ total samples are selected. The periodic secretary algorithm achieves the highest entropy reduction among the streaming algorithms. However, the algorithm with poorly tuned $\lambda$ reaches the end of the stream without selecting all samples. The periodic secretary algorithm with well-tuned $\lambda$ stays very close to the upper bound set by the offline greedy algorithm. (b) Using the selected samples, plankton counts at unknown locations are predicted on a held-out test set. The prediction mean-squared error decrease as samples are selected and is minimized using the periodic secretary algorithm with well tuned $\lambda$.

Figure 3-7(a) demonstrates that samples selected by the periodic secretary algorithm achieve the highest entropy reduction across the environmental observation space. Intuitively, this means that we can use these samples to do inference about plankton concentrations in unknown environmental conditions. To test this assumption, we quantify how well the representative samples selected by each algorithm can be used to predict *Guinardia flaccida* concentrations on a held out test set of

{temperature, cos(fraction of year)} environmental conditions (the prediction mean-squared error should be minimized). Figure 3-7(b) shows that on average the model learned using sample sets selected by the periodic secretary algorithm produce more accurate predictions of plankton counts then all other streaming algorithms. Note that choosing points according to the entropy criterion is a good strategy from an information-theoretic perspective when trying to reduce prediction error, but higher entropy reduction will not necessary directly equate to lower mean-squared prediction error for a specific dataset. This is why there are places in Figure 3-7(b) where an algorithm with lower entropy reduction achieves lower prediction mean-squared error.

## 3.8 Discussion

The periodic secretary algorithm is a robust and versatile tool that can be applied in a variety of applications. Many real-world periodic data streams can be considered approximately periodic, given that the period-to-period variation is Gaussian distributed. The algorithm is robust to noisy estimates of the period length, requiring only that the algorithm observes one full period of the data, and does not depend on knowledge of the data stream length, making it suitable for continuous monitoring. Our work extends previous results in information-theoretic sample selection and adapts secretary algorithms to data domains that produce periodic spatially and/or temporally correlated data streams, such as robotics and environmental monitoring. Although we focus on periodic phenomena, techniques similar to those presented here could be used to provide performance bounds for irrevocable sample selection from data streams with other types of spatiotemporal structure. We hope that this work will serve as a foundation for developing secretary algorithms that can be applied to these interesting data domains.

# Chapter 4

# Hybrid Bayesian-Deep Topic Models for Terrain Characterization

Chapter 3 presented a full formulation – including a model, scientific objective, reward function, and planning algorithm – for secretary sampling problems in periodic data streams. The periodic secretary algorithm used a Gaussian process to perform regression between between sensory percepts and a latent quantity of scientific interest. This environmental model allowed an autonomous science agent to select sampling actions that explicitly reduced uncertainty in this model by optimizing the entropy criterion. In Chapter 3, the scientific model and objective were assumed to be fairly straightforward to specify; the challenging aspects of the scientific information-gathering problem were specifying a reward function given the scientific objective and performing irrevocable sampling in a streaming setting.

In this chapter, we consider problems for which the specification of a good scientific model and objective is in itself challenging. The modeling problem presented in Chapter 3 was to directly learn a mapping between low dimensional environmental percepts and a low dimensional latent quantity of interest. This is a supervised learning problem: given a set of paired training examples of environmental percepts and quantity of interest samples, learn a Gaussian process regression model that relates the two. However, science missions in environments where domain scientists have little *a priori* knowledge about the environment – such as the surface of extraterrestrial

planets or the deep sea – often do not permit the formulation of a straightforward supervised learning problem. In these domains, the scientific objective may be instead to explore an unknown environment and to discover new phenomena, based only on sensory percepts about the environment. We consider the scientific information-gathering problem in these scenarios, in which the modeling problem is challenging for two primary reasons:

- The environmental percepts are high-dimensional image data. Cameras are emerging as a ubiquitous, low-cost sensor on a variety of autonomous platforms such as autonomous underwater vesicles (AUVs). Images are information rich and contain valuable semantic and scientific information. However, the dimensionality and complexity of image data can cause significant challenges for traditional modeling approaches.

- Little or no prior information is available about the environment, making it infeasible to directly specify a scientific quantity of interest and supervised learning problem. Instead, the scientific objective is one of exploration and discovery: learn about the environment online directly from perceptual information.

The remainder of this chapter introduces a hybrid deep-Bayesian model for unsupervised visual scene understanding and terrain characterization for scientific robots. This hybrid method uses deep convolutional autoencoders to learn low-dimensional feature embeddings for complex image data and unsupervised Bayesian nonparametric models to learn an environment's semantic structure directly from visual data. Although this chapter does not explore how this model can be used within a scientific planning framework, we propose the use of model perplexity as a measure of the information content of a scene and evaluate how well measures of perplexity map to human intuition about scientifically interesting scenes.

The modeling problem is introduced within the context of autonomous deep sea exploration in Section 4.1, followed by a review of related literature in Section 4.2. The hybrid approach employing Bayesian topic models and deep convolutional autoencoders is presented in Section 4.3. Finally, the combined hybrid approach is

evaluated on imagery from two different marine missions collected by the SeaBED AUV at the Hannibal Sea Mount off the coast of Panama (Section 4.4), followed by a discussion of the results (Section 4.5). The work presented in this chapter has appeared in abbreviated form in Flaspohler et al. [34].

## 4.1 Introduction

The benthic deep sea, the largest two-dimensional habitat on earth, is difficult to study and vastly unexplored. Autonomous underwater vehicles (AUVs) are filling observational gaps by collecting large datasets consisting of multiple sensor modalities, including seafloor imagery. This chapter presents a novel unsupervised machine learning technique to discover and visualize structure in image datasets, enabling concise mission summarization and equipping exploratory robots with the capacity to describe their environment semantically, a precursor to adaptive real-time exploration. Although the focus of this chapter is the underwater domain, the proposed approach is applicable to any domain where there exists large volumes of unstructured image sequence data that would typically require human analysis, such as remote sensing and long term monitoring.

Some of the most successful models for discovering structure within discrete data without supervision are Bayesian topic models, such as the latent Dirichlet allocation (LDA) [9] and its non-parametric extension, the Hierarchical Dirichlet process (HDP) [112]. Initially applied to text corpora, the modeling assumptions made by LDA and HDP allow them to discover useful latent structure that often corresponds to cohesive, human-understandable topics [19]. This property of topic models led to impressive results in areas such as text clustering [9], corpora summarization, and recommender systems [64].

The success of Bayesian topic models for semantic understanding of text documents led to their adaptation to computer vision applications; this application is visualized in Figure 4-1. By replacing text words with discrete visual features, LDA and HDP models can be applied directly to image data [11, 31, 119]. The most pop-

ular discretization of an image into visual words employs standard features such as SIFT [71], SURF [6], or Oriented BRIEF (ORB) [92]. Visual topic models have been used successfully in robotics applications for for unsupervised scene understanding [106] and adaptive mission planning [42].

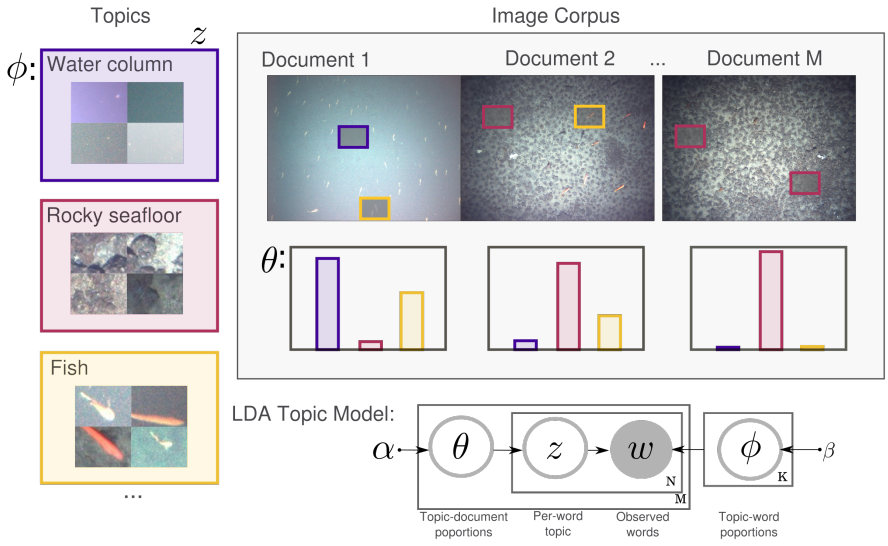


Figure 4-1: **Visual topic models**: Topic models can be applied directly to visual corpora by considering each image to be a "document", which is a collection of "visual words". In the LDA model, given observed visual words that are separated into documents, inference techniques can be used to recover the topic-word distribution $\phi$ and the topic-image distributions $\theta$, as well as the per-word topic assignments $z$. The graphical model representation of the LDA is presented on the bottom of the figure.

However, the modeling capacity of topic models is fundamentally limited by the expressive power of the observed "words". Hand-crafted image features capture low-level patterns based on local image gradients. In contrast, deep neural models are often able to learn more complex, domain-specific features. Several papers have leveraged this property of neural networks to build more expressive models of textual data [79], [96]. In this chapter, we make the natural extension to unsupervised feature discovery for image data; the proposed model architecture is visualized in Figure 4-2.

Much like textual data, image data show strong spatial correlations. These correlations are ignored by the simplifying bag of words (BOW) assumption made in most Bayesian topic models. Ideally, data features could encode these spatial correlations directly. Convolutional autoencoders (CAE) [77] preserve spatial relationships

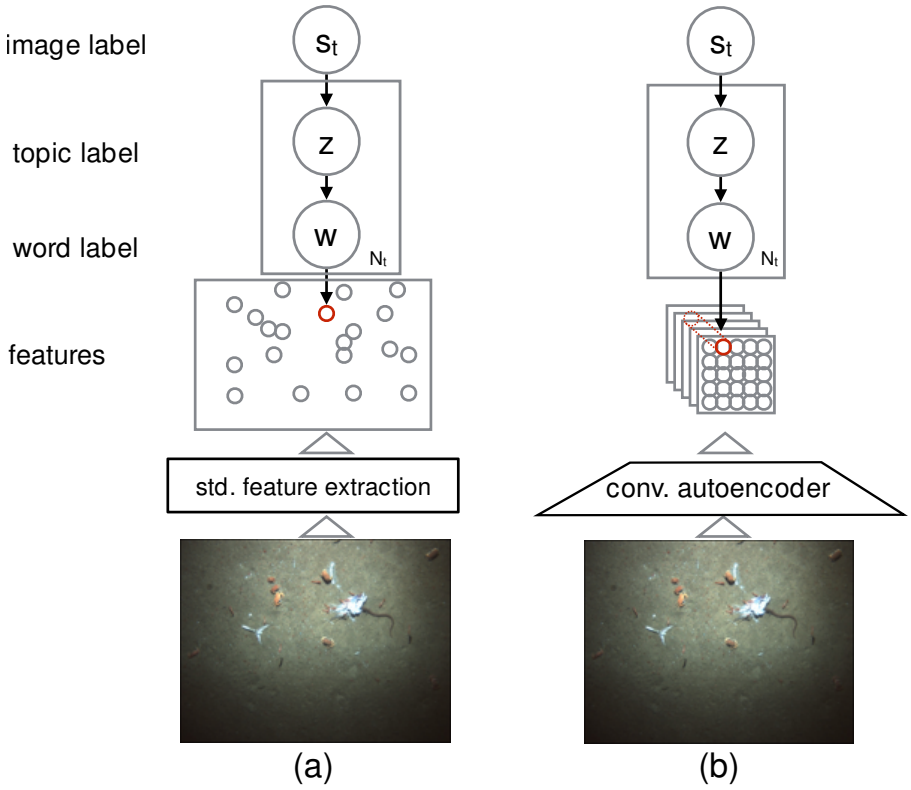


Figure 4-2: **Architecture of visual scene understanding models**: The two scene modeling techniques evaluated in this chapter are depicted. Here $s_t$ is the image label, and $z$ is the topic label of a visual word $w$ in the input image. (a) A baseline spatiotemporal topic model using standard computer vision features as input. (b) The proposed spatiotemporal topic model using convolutional autoencoder-based features.

in data and hence are a powerful method for discovering useful features for image data. However, these features have not yet been incorporated into a topic modeling framework, in part because of the challenges of designing a CAE network architecture that produces useful features within a topic modeling context.

This work presents a CAE architecture that discovers feature representations directly from an image dataset and applies those features within an HDP-based topic modeling framework to discover cohesive visual topics. We explore the performance of this hybrid HDP-CAE model on the motivating application of autonomy and mission summarization for exploratory marine robots. We evaluate how well the topics discovered by the hybrid HDP-CAE model correspond to biologically distinct seafloor terrains and compare the hybrid HDP-CAE model to an HDP model using standard

image features. Finally, we quantify the performance of the hybrid model on the secondary task of identifying anomalous images within a dataset. The probabilistic anomaly detection enabled by Bayesian topic models can inform more effective mission planning for exploratory marine robots and is more broadly useful for data summarization and visualization.

All models are evaluated on a realistic dataset that an individual biologist or data scientist could collect and wish to analyze, consisting of less than 4,000 seafloor images collected in-situ by a marine robot. Even in this small-data domain, we demonstrate that state-of-the-art performance can be achieved by applying neural feature discovery and nonparametric topic modeling to the task of unsupervised seafloor terrain characterization.

## 4.2 Related Work

Recent efforts have leveraged the power of neural models to discover data features within the context of topic modeling. Many, however, continue to rely on predefined data features at some level. Frameworks that overcome this dependence have difficulty incorporating custom features within a completely unsupervised Bayesian topic model.

For textual data, Mikolvo et al. [79] propose the reverse of the architecture we present here; an LDA model is used to produce a contextual feature vector that is input into a recurrent neural network for contextually-aware language modeling. While powerful, this model does not address the model's dependence on standard word features and employs a BOW assumption. In [96], the BOW assumption is relaxed. Instead, a convolution operation maps variable length text sequences into a low-dimensional latent space. Unlike the work presented here, simple distance-based clustering is applied to discover semantically similar documents in place of a Bayesian topic model.

For image data, Wan et al. [118] introduce a hybrid neural-Bayesian topic model based on a Deep Boltzmann Machine (DBM). As in this work, the feature represen-

tation discovered by the DBM is fed directly into an HDP topic model to discover visual topics. However, instead of discovering features directly from image data, SIFT features are extracted from the image and the neural network learns an image representation based on these features, thereby not reducing the dependence on human-designed features.

The work most similar to our own is presented in [116]. The Hierarchical-Deep model introduced uses an HDP to learn priors over the activations of a DBM. In this way, the model is able to learn generic features from image data that enable learning image classes from very few examples. Each image in the model is annotated with a lower level class and the HDP discovers a hierarchy over these low-level classes. While suitable for the goal of one-shot learning, this supervision limits the generality of the Hierarchical-Deep model to a purely unsupervised problem. The model is also not convolutional, limiting the utility of the learned features. Convolutional autoencoders are directly able to model spatial correlations in image data and therefore are more suited to discover useful image representations.

Additionally, none of the aforementioned works evaluate their models on the small or medium-sized datasets that are prevalent in unsupervised learning applications. Instead, they use large (4 million+) standard image datasets [116] or 2D toy, simulated images [118].

Other works have incorporated neural feature learning for robotics applications outside of a topic modeling context. Naseer et al. [81] use up-convolutional networks to discover latent feature representations for the task of segmenting images. Rao et al. [89] use an autoencoder to learn features for classification of marine images. However, both of these works require human annotation and thus are not applicable in an unsupervised setting.

## 4.3 Methods

In the following sections, we provide a brief review of topic models and then discuss the two major components of the proposed hybrid HDP-CAE model: 1) a spatiotemporal

HDP topic model and 2) a pipeline for discretizing an image into visual words using a convolutional autoencoder.

### 4.3.1 Bayesian topic models

Topic models [9, 45] seek to uncover semantic structure in a corpus of discrete data, segmented into documents. Topic models propose that each observed word in a document is generated by a latent topic and each document in a corpus has its own probability distribution over topics. Using word co-occurrences and distribution sparsity priors, the distribution over topics $Z_i$ for each word $W_i$ can be inferred. Under this model, the probability of the $i$th word $W_i$ in document $d$ can computed by marginalizing over the latent topics:

$$\mathbf{Pr}(W_i = w_i \mid d) = \sum_{k=1}^{K} \mathbf{Pr}(W_i = w_i | Z_i = k)\mathbf{Pr}(Z_i = k|d), \tag{4.1}$$

where $K$ is total number of topics, $\mathbf{Pr}(W_i = w_i|Z_i = k)$ is the probability of word $i$ under topic $k$, and $\mathbf{Pr}(Z_i = k|d)$ is the probability of topic $k$ in document $d$.

### 4.3.2 Realtime spatiotemporal HDP model

Traditional topic modeling frameworks treat each word in a document as exchangeable. We instead adapt the ROST HDP model presented in [42], which relaxes the BOW assumption and explicitly models the correlation between spatiotemporal neighborhoods in a continuous image stream. ROST uses a Dirichlet process to model the growth in number of topics with the size and complexity of the data. A biased Gibbs sampler [39] enables online computation of the posterior distribution over topics for observed visual words.

The ROST model factors the probability of observing the visual word $W_i$ at location $x$ and time $t$ in terms of the topic label variables $Z_i$. This factorization is

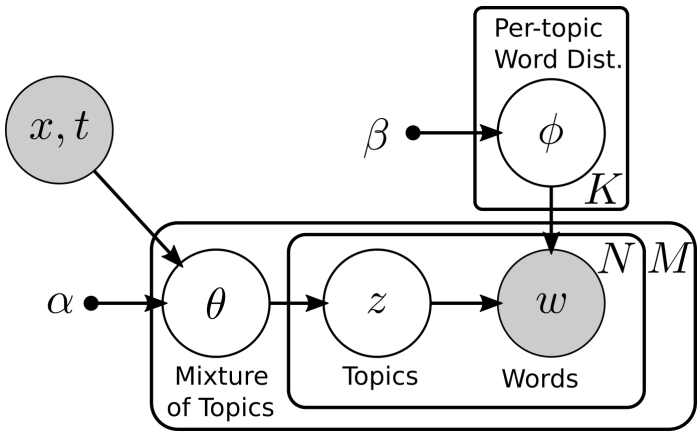


Figure 4-3: **Graphical model representation of ROST**: The graphical model representing the Realtime Online Spatiotemporal topic modeling framework (ROST). In this unsupervised Bayesian nonparametric model, visual words $w$ are observed at spatiotemporal location $\{x, t\}$ and Gibbs sampling is used to approximate the distribution over the latent parameters $\theta$, $\phi$, and $z$ for given hyperparameters $\alpha$, $\beta$.

visualized in the ROST graphical model in Figure 4-3

$$\mathbf{Pr}(W_i = w_i | x, t) = \sum_{k=1}^{K} \mathbf{Pr}(W_i w_i | Z_i = k) \mathbf{Pr}(Z_i = k | x, t). \tag{4.2}$$

The distribution $\mathbf{Pr}(W_i = w_i | Z_i = k)$ is invariant to the spatiotemporal location of the observation, while $\mathbf{Pr}(Z_i = k | x, t)$ models the distribution of topic labels in the spatiotemporal neighborhood of location $(x, t)$. $K$ is total number of topics that have at least at least one or more words assigned to them, plus one more to encode the possibility of creating a new topic for word $w_i$. ROST uses the Dirichlet distribution to model $\mathbf{Pr}(W|Z)$, allowing for control of the sparseness of the topic model, whereas $\mathbf{Pr}(Z|x, t)$ is modeled using the Chinese Restaurant Process [113], removing the need to predetermine the total number of unique topics. We refer the reader to Girdhar et al. [42] for a full description of the ROST topic model and inference procedure used to recover topic assignments.

### 4.3.3 Convolutional autoencoder architecture and training

To extract a discrete list of words from an image, we exploit the ability of neural models to discover useful abstract representations of data without supervision. We train a CAE architecture following the encoder/decoder paradigm described in [77]. The input image is first transformed into a lower dimensional bottleneck layer using successive convolution operations and rectified linear unit (ReLU) activations and then expanded back to its original size using a deconvolution operation with tied weight matrices. We call the channels in the bottleneck layer the *latent channel activations* (LCA). This network architecture is visualized in Figure 4-4.

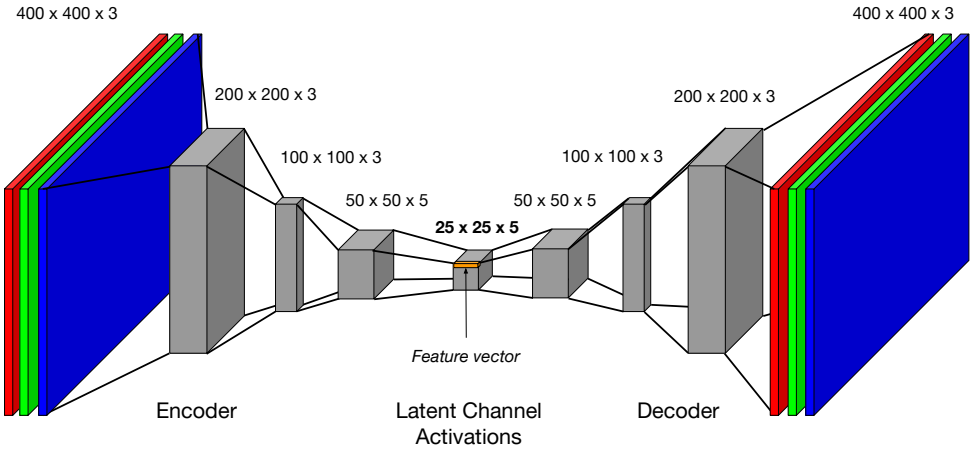


Figure 4-4: **Network architecture for the convolutional autoencoder (CAE) used to extract low level visual features from the image datasets**: Network specific parameters were set as: training epochs (400 epochs), output channels (ordered by encoding layer 3-3-3-5-5 channels), stride (2), and convolutional filter size (ordered by encoding layer 10-10-3-3 pixels),

The squared error between the original image and the image reconstruction provides an unsupervised loss function which allows the weight matrices for each layer to be learned. The network is trained using stochastic gradient descent with L2 regularization on the weight matrices. Because our goal is to treat the nodes in the bottleneck layer as non-overlapping features of the image, the neuron redundancy encouraged by dropout regularization is actually undesirable [103], so we do not include dropout. In this unsupervised setting, the entire dataset is used to train and test the model,

so generalization is much less of a concern then in supervised learning problems. In this chapter, all models are also trained without max-pooling/unpooling layers; dimensionality reduction is achieved using a stride greater than one and overlapping convolutional filter windows. During experimentation, we found that the inclusion of max-pooling and unpooling layers decreased the expressive power of the LCA, contrary to [77], so the final models were purely convolutional with ReLU nonlinearity.

The CAE network architecture used to produce the results in this chapter is shown in Figure 4-4. The network consists of four encoding layers and four associated decoding layers. Each sequential encoding layer increases the number of output channels while decreasing the height and width of each individual channel, following a pyramid architecture. The architecture-specific parameters, such as number of training epochs (400 epochs), output channels (ordered by encoding layer 3-3-3-5-5 channels), stride (2), and convolutional filter size (ordered by encoding layer 10-10-3-3 pixels), were determined empirically. After training, we remove the decoding layers of the network and use the LCA to generate low dimensional image features.

### 4.3.4 Generating a visual vocabulary for topic models

HDPs require discrete data drawn from a vocabulary $\mathcal{V}$. To produce a vocabulary for CAE features, the CAE is trained on several example ocean mission datasets and the LCA for each image are extracted as described in Section 4.3.3. This $25 \times 25 \times 5$ tensor is segmented into 625 feature vectors of length 5 by taking slices across LCA channels, as shown in Figure 4-4. These features are then clustered using the $k$-means algorithm into $|\mathcal{V}|$ clusters, where $|\mathcal{V}|$ is the desired vocabulary size. The centroid of each cluster represents a visual vocabulary word. Because the LCA are low dimensional ($5 \times 1$ pixels using the architecture in Figure 4-4), as compared to 128-dimensional SIFT/SURF features, this clustering is relatively efficient.

Given a new image, visual features are extracted and mapped to the visual word $v_i \in \mathcal{V}$ corresponding to the nearest neighbor in the space of cluster centroids.

### 4.3.5 CAE feature visualization

The LCA discovered by the CAE model correspond to a low-dimensional, abstract representation of the image. In later sections we will apply these features within a topic modeling framework and attempt to visualize the properties of the latent channels constructed in this manner.

To quantify the strength of a latent channel's response to a particular input image, we consider the magnitude of each latent channel at a particular pixel location $p_{ij}$ in the LCA. For each of the $5 \times 5$ pixels, we assign the pixel $p_{ij}$ to the channel with the maximum activation at location $(i, j)$. The magnitude of a latent channel M is equal to number of pixels for which it had the maximal value. This approach produces a clearer segmentation between channels than directly plotting channel magnitudes. Different channels have different baseline activations. To compare between channels, we normalize each channel's activation between their minimum value and maximum value before plotting.

## 4.4 Experiments

We evaluate our hybrid HDP-CAE model against a ROST HDP baseline using SURF [6] and ORB [92] features. The two models are visualized in Figure 4-2. We apply each model to image streams from two marine robot missions collected by the SeaBED AUV at the Hannibal Sea Mount, Panama [87] and present experimental results.

Mission I contains 1,117 images sampled every four seconds from a downwards facing camera mounted to the bottom of the robot. During Mission I, the robot passes over several seafloor terrains, including images of the water column, a rocky seafloor, and a porous sandy seafloor. This mission tests the model's ability to discover visually distinct terrain types.

Mission II consists of 2,296 images sampled in a similar manner. Mission II contains mostly images of a sandy seafloor, interrupted several times by large, biologically interesting phenomena, such as crab congregations, seafloor carnage, and geothermal vents. In addition to terrain discovery, this mission also tests the model's ability

to accurately identify anomalous, scientifically interesting images within a mission dataset.

To evaluate how well the image topic labels discovered by the unsupervised HDP-CAE model correspond to visually meaningful seafloor terrains, we hand-labeled each image in both missions with one of thirteen possible terrain labels, including: 'water column', 'sparse boulders', 'smooth sand', 'biological congregation', etc. These labels are used exclusively for model evaluation.

We apply the HDP-ROST model described in Section 4.3.2 to the two mission datasets, using standard image features and CAE-derived LCA features for the standard HDP model and the hybrid HDP-CAE model respectively. New images are incorporated into the model in a streaming fashion; visual words are extracted from a new image and added to the model at regular intervals (200 ms). The ROST hyperparameters for Mission I and Mission II respectively were set to maximize mutual information between discovered topics and hand annotated labels: $\alpha = 0.1, 0.1; \beta = 25, 50; \gamma = 10^{-7}, 10^{-7}$. After Gibbs sampling, we have an approximation of the posterior over topics $z_i$ for an observed visual world $w_i$, $P(z_i|w_i = v)$. The predicted topic label for each visual word is assigned as the *maximum a posteriori* (MAP) topic label given by the posterior, and the predicted scene label $s_t$ for each image is calculated as the majority consensus of the visual words in the image.

Taking the MAP is a standard way of reducing a probabilistic distribution over a latent variable to a single point estimate. However, by using only the MAP topic label for each image, we are not allowing for important visual constructs to be represented by a mixture of topics. This approximation may be suitable for our experimental domain. The majority of images in our marine datasets consist of a homogeneous visual terrain, and an ideal topic model would discover topics rich enough to have nearly a one-to-one correspondence with semantically distinct visual constructs. Having to build a heuristic on top of a topic model to extract meaningful topics from mixtures of the discovered topics adds an unnecessary layer of complication to the model.

## 4.5 Results

To quantify the accuracy of the topic labels discovered by each of our models, we use normalized mutual information between the annotated topic distribution and the computed topic distribution. Mutual information captures the reduction in entropy of a random variable X after observing random variable Y (Eq. 4.3). A normalized mutual information score of one indicates that X and Y are completely dependent (i.e. the discovered topics are completely correlated with the true labels), whereas a mutual information score of zero indicates independence (i.e. the discovered topics are unrelated to the true labels).

$$\begin{aligned} I(X,Y) &= H(X) - H(X|Y), \\ &= \sum_{x,y} P(x,y) \log \frac{P(x,y)}{P(x)P(y)}. \end{aligned} \tag{4.3}$$

### 4.5.1 Mission I - seafloor terrain discovery

Mission I tests the model's ability to uncover topics corresponding to meaningful visual terrains. Table 4.1 shows that the topics discovered by the hybrid HDP-CAE model are highly predictive of the ground-truth seafloor terrains. The raw topic distribution (before MAP reduction) for the two models is plotted in Figure 4-5 along with example images from the major terrain types. To generate the plots in Figure 4-5, visual words are extracted from an image at time $t$ and assigned a topic label $z_i$ as described in Section 4.3. The proportion of words in the image at time $t$ assigned to each topic label is shown on the y-axis, where different topics are represented by colors. Colors are unrelated across plots.

Although the hybrid HDP-CAE model differs from the human annotated terrains by, for example, not modeling the transient topic at (3), the major terrain transitions are captured faithfully. The rocky seafloor terrain that dominates at (4) and then partially appears again at (6) is assigned to the same topic. The moment that the robot first observes the seafloor through the water column in (3) is captured as a

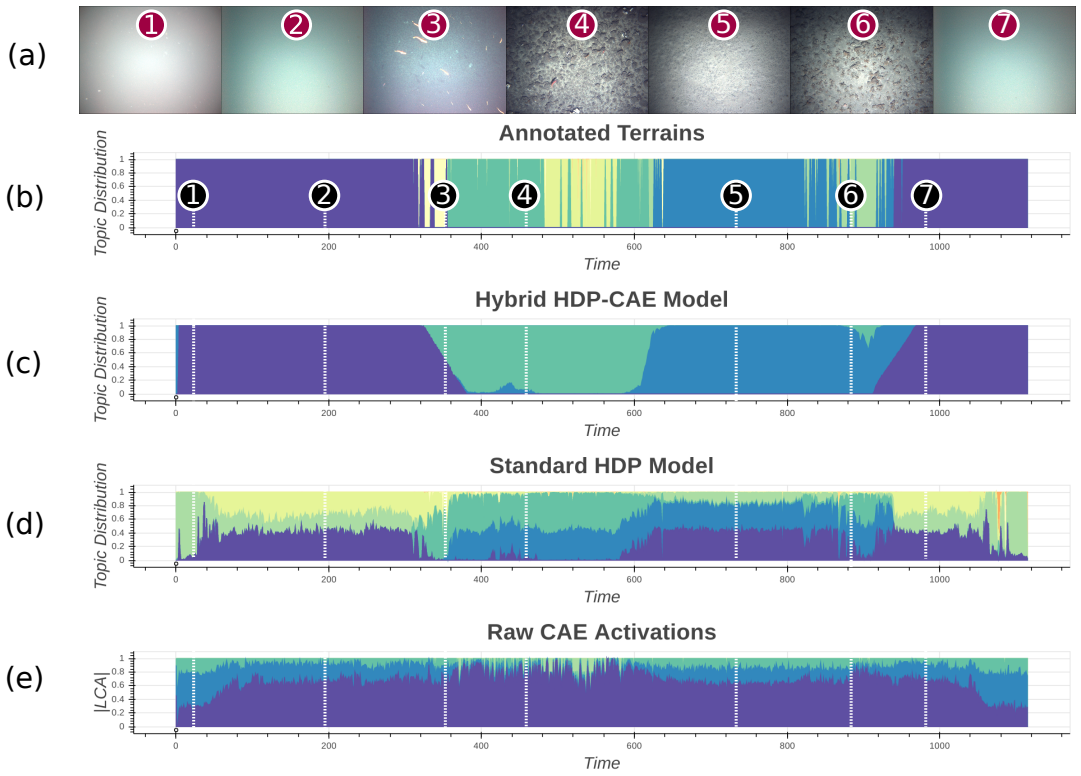


Figure 4-5: **Results for unsupervised topic models versus hand-annotated terrain labels (b) for the Mission I dataset**: Example images from the dataset are shown in (a). To generate plots (c,d), visual words are extracted from an image at time $t$ and assigned a topic label $z_i$ as described in the text. The proportion of words in the image at time $t$ assigned to each topic label is shown on the y-axis, where different topics are represented by colors. Colors are unrelated across plots. The hybrid HDP-CAE model (c), using more abstract features, is able to define topics that correspond more directly to useful visual phenomena than the HDP model using standard image features (d). The learned feature representation is visualized as described in (e).

mixture of the 'water column' topic (indigo) and the 'rocky seafloor' topic (green). The consistency of the hybrid HDP-CAE topic model with human annotated terrain labels is quantified using mutual information between topics and annotated terrains as described in Section 4.5; results are shown in Table 4.1.

Despite its low mutual information scores, the standard HDP models does capture some of the major terrain transitions as changes in the topic distributions. However, it is not clear how to distill this information into meaningful topics. Because the hybrid HDP-CAE model uses much fewer, more abstract features, it is able to define

| | Model | $I(X,Y)$ |
|---|---|---|
| Mission I | Standard HDP | 0.185 |
| | **Hybrid HDP-CAE** | **0.535** |
| Mission II | Standard HDP | 0.123 |
| | **Hybrid HDP-CAE** | **0.441** |

Table 4.1: Mutual information between discovered topics and annotations

topics that correspond more directly to useful visual phenomena.

### 4.5.2 Mission II - biological anomaly detection

Characterizing seafloor terrains is a vital task for an exploratory marine robot. Another complementary skill is the ability to identify images that are anomalous under the robot's current model of the world and flag these as interesting, one of the behaviors demonstrated by the standard HDP model presented in [42]. The Mission II dataset was designed to test both of these abilities.

Comparing mutual information between terrain labels in Table 4.1, the hybrid HDP-CAE model again outperforms the baseline model. Figure 4-6 shows the raw topic distribution for the two models, along with example images from the major terrain types. Mission II is a more visually homogeneous dataset, consisting almost entirely of sandy seafloor images. The hybrid HDP-CAE model discovers segmentations within the sandy seafloor topic that the human annotators do not, but otherwise captures major terrain transitions. However, the hybrid HDP-CAE model does not capture some of the more transient topics, such as the crustaceans at (7).

To quantify this result further, we introduce the notion of perplexity. Because HDP is a probabilistic model, it is straightforward to quantify the average word perplexity (Eq. 4.4) of a new image $X_t$ under the model.

$$Perplexity(X_t) = \exp\left(-\frac{\sum_{w \in W_t} \log p(w|X_t)}{|W_t|}\right), \tag{4.4}$$

where the set $W_t$ consists of the visual words in $X_t$. High perplexity indicates that the image is not well modeled by the topic model, whereas low perplexity indicates an

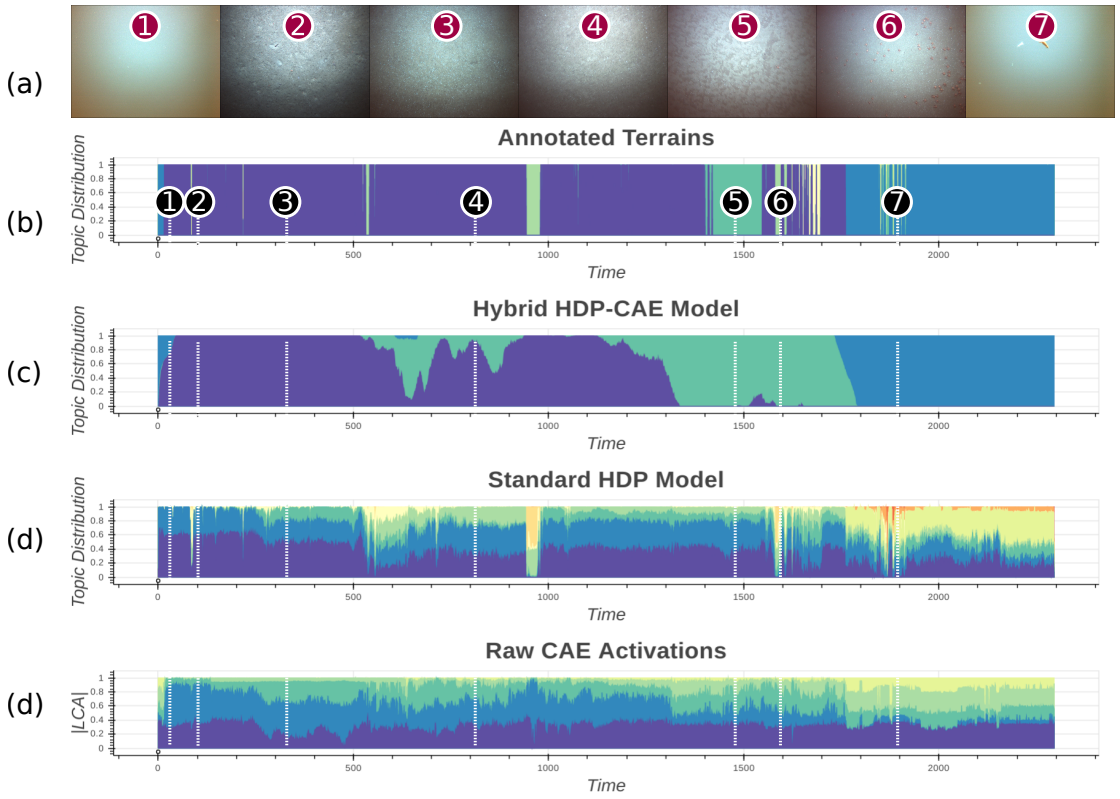


Figure 4-6: **Results for two unsupervised topic models versus annotated labels (b) in for the Mission II dataset**: Example images from the dataset are shown in (a). The hybrid HDP-CAE model (c) again outperforms the HDP model using standard image features (d). However, the hybrid model fails recognize some of the more transient topics, such as the crustacean swarm at (7). The learned feature representation is visualized as described in the text in (e).

image that is well explained by the topic model. Figure 4-7 compares the perplexity response of each model when presented with biologically interesting images; the hybrid model does not have an obvious increase in perplexity when presented with the images of seafloor carnage (2), crab congregations (3), or submerged tree (4).

To compute how well each model's perplexity score corresponds to some interesting visual phenomena, we annotated each image in the dataset as either high, medium, or low 'scientific interest' and computed the mutual information between the annotated and computed perplexity. Although perplexity scores do not necessarily correspond well with human intuition about semantic coherency [19], perplexity has been used successfully for anomaly detection in previous work [42]. The CAE is not a proba-

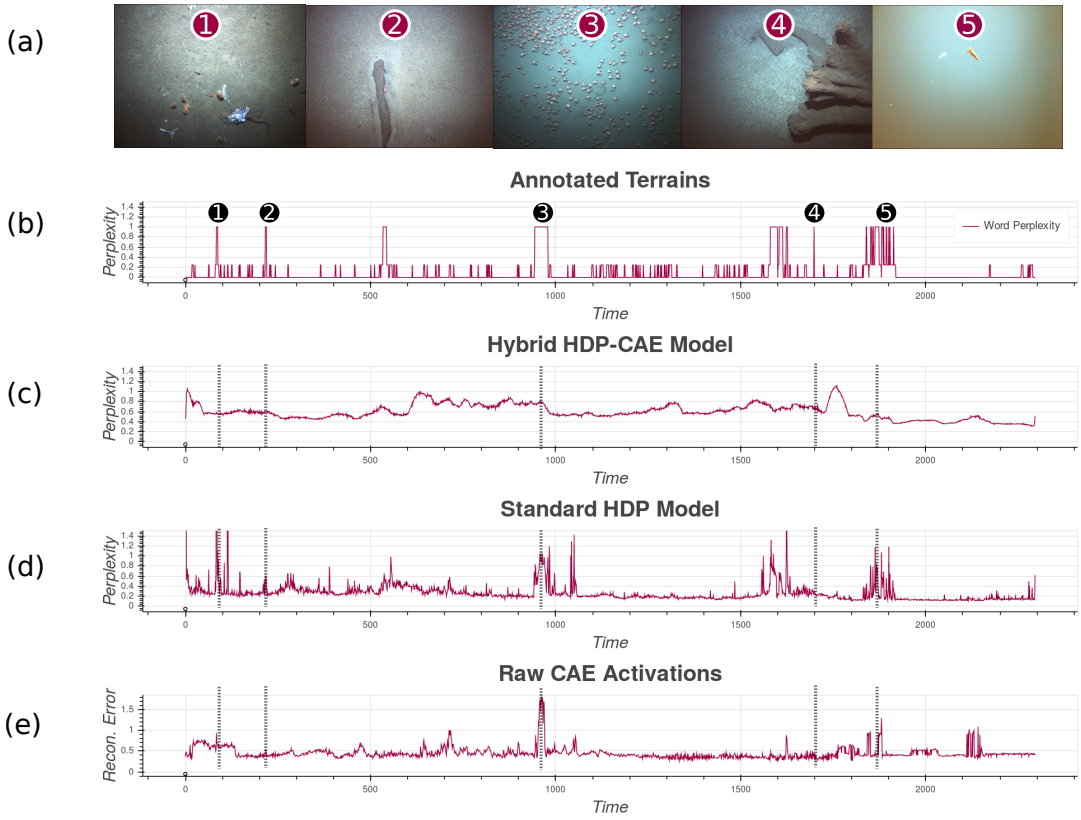


Figure 4-7: **Correlation of perplexity score of the models (c,d,e) with annotated biological anomalies (b) for the Mission II dataset**: Example images of biological anomalies are shown in (a). Each image in the dataset was labeled with high, medium, or low perplexity (b). Although all three models do have differential responses in areas of high perplexity, the HDP model using standard features (d) outperforms the hybrid HDP-CAE model (c) and raw reconstruction error from the CAE (e).

bilistic model, so there is no well-defined operation to compute perplexity. We instead use squared image reconstruction error as a proxy for perplexity. The mean $\mu$ and standard deviation $\sigma^2$ of each model's positive perplexity distribution are used to bin the perplexity into low ($0 \leq x \leq \mu + \sigma^2$), medium($\mu + \sigma^2 \leq x \leq \mu + 2\sigma^2$), and high ($x > \mu + 2\sigma^2$) perplexity. The results of this analysis are shown in Table 4.5.2. Although all three models do have differential responses in areas of high annotated perplexity, the standard HDP model's perplexity has higher mutual information with annotated perplexity. We will address this discrepancy in Section 4.6.

| | Model | $I(X,Y)$ |
|---|---|---|
| Mission II | **Standard HDP** | **0.153** |
| | Hybrid HDP-CAE | 0.006 |
| | Raw CAE | 0.033 |

Table 4.2: Mutual information between perplexity and annotations

## 4.6 Discussion

The proposed hybrid HDP-CAE model significantly outperformed alternative models on the task of seafloor terrain discovery. The hybrid model, however, did not perform as well on the secondary task of anomaly detection, as quantified by image perplexity. Our hypothesis is that this limitation stems from the inherently imbalanced nature of anomalies within a dataset. In our anomaly detection experiments, the CAE was trained on over 2000 images of sandy seafloor and only 80 images of crab congregations. Many other anomalous events, such as the seafloor carnage in Figure 4-7(a) appear for even shorter spans. Neural models have been shown to struggle when presented with imbalanced training data [20]. There may not be enough images of anomalous biological events for the CAE to learn a meaningful feature representation. Although the CAE image reconstruction error, plotted in Figure 4-7(d), does capture the inability of the features to represent the anomalous images, our current hybrid HDP-CAE model does not incorporate this uncertainty within the topic modeling stage. An interesting extension of this work would be to use image reconstruction error directly as a metric of feature quality. Alternatively, there are methods within the machine learning community for dealing with imbalanced datasets that could improve CAE training [47, 13].

For the specific biological anomalies tested here, such as the crab congregations or seafloor carnage, it may be difficult to outperform standard image features, which are designed to detect areas of high image gradient. However, there are other reasons to prefer a CAE-based anomaly detector. Interesting anomalies may not always manifest themselves as complex visual structure; a smooth sandy seafloor is anomalous within a rocky mission. Standard image features may struggle to represent these visually

simple anomalies. Additionally, CAE-based anomaly detectors can use reconstruction error to not only detect when an image is anomalous, but also which part of the image is particularly difficult to resolve. The ability to spatially localize anomalies could be very useful in robot scene understanding and real-time planning.

Another important extension to the hybrid HDP-CAE model presented in this work is the adaptation of convolutional feature discovery for realtime, streaming applications. Bayesian nonparametric models are well suited for the life-long learning required in streaming and robotics applications; this is one compelling reason to use them over purely neural models. However, for simplicity, the CAE-based feature discovery training in this work was done offline on complete datasets. Exploring methods for efficient, life-long training of convolutional models is an important area of future work for applying hybrid HDP-CAE models to realtime applications.

# Chapter 5

# Conclusion

The preceding chapters have presented novel approaches to both planning and environmental modeling within scientific-information gathering problems. First, the problem of secretary sampling in periodic data streams was explored. A novel sensor planning algorithm was introduced that utilized standard environmental modeling techniques and information-theoretic reward functions to perform secretary sampling with bounded suboptimality. Second, the problem of environmental modeling was addressed for data domains with high-dimensional discrete structure, such as images, and for which very little information was available *a priori* about the structure of the data stream. For these environments, a novel unsupervised learning model was introduced that combined deep convolutional autoencoders to learn feature embeddings directly from camera percepts, with a Bayesian nonparametric model of semantic structure in a visual data stream.

## 5.1 Summary of Contributions

The following sections summarize the planning and modeling approaches introduced in this thesis, towards the goal of enabling autonomous science missions using robots and autonomous sensors.

### 5.1.1 Periodic secretary sampling with information reward

Chapter 3 presented a novel algorithm for online, irrevocable sample selection from periodic phenomena. The proposed periodic secretary algorithm was shown to select sample sets according to any monotone submodular set function with bounded suboptimality. The periodic secretary algorithm is a robust and versatile tool that can be applied in a variety of applications. This work extended previous results in information-theoretic sample selection and adapted secretary algorithms to data domains that produce periodic spatially and/or temporally correlated data streams, such as robotics and environmental monitoring. Although the focus of Chapter 3 was on periodic phenomena, techniques similar to those presented here could be used to provide performance bounds for irrevocable sample selection from data streams with other types of spatiotemporal structure. We hope that this work will serve as a foundation for developing secretary algorithms that can be applied to these interesting data domains.

### 5.1.2 Hybrid deep-Bayesian environmental models

Bayesian topic models have achieved impressive performance by learning both model parameters and useful structure directly from data. However, these nonparametric models still fundamentally rely on predefined feature representations of data. Chapter 4 presented a novel model that overcame this limitation using convolutional autoencoders, allowing unsupervised discovery of *both* a feature representation and thematic structure in image data. The proposed hybrid model incorporated a convolutional autoencoder for data-driven feature discovery within a Bayesian topic modeling framework. Chapter 4 applied this model to the problem of high-level scene understanding and mission visualization for exploratory marine robots. On complex mission datasets, the hybrid model discovered a rich latent visual structure that has over four times the mutual information with biologically meaningful seafloor terrains when compared to a Bayesian nonparametric topic model with standard, hand-designed features. This work defined a paradigm for including the ability of unsupervised neu-

ral models to discover useful, low-dimensional data representations within a Bayesian nonparametric topic modeling framework and demonstrated state-of-the art performance on a challenging problem from the marine robotics community.

## 5.2 Directions for Future Work

This section outlines several promising directions for future work and potential extensions of the work presented in the preceding chapters.

### 5.2.1 Secretary sampling problems

Chapter 3 presented a modeling and planning framework for secretary sampling in periodic data streams. Several extensions immediately suggest themselves to improve the flexibly and robustness of the algorithm in a variety of scientific information gathering problems, as outlined in the following sections.

#### Motion planning for secretary sampling problems

The secretary sampling algorithm presented in Chapter 3 focused on reactive, opportunistic sampling. This is the only planning strategy available for some autonomous platforms, such as static sensing nodes that can only make sampling decisions by reacting to the observed data stream. However, other autonomous platforms such as mobile robots have the ability to plan trajectories and thereby modify the observed data stream. Although the problem of trajectory and motion planning for reward maximization has been studied extensively, the secretary-style constraint of online, irrevocable sample selection changes the paradigm and will require solutions that explicitly consider these new constraints in the path planning problem. Some preliminary work has studied this problem in the context of heterogeneous multi-robot teams [73]. Dividing the problem into separate path planning and secretary sampling modules is a promising method that allows techniques from both informative path planning and secretary sampling literature to be applied directly, at the cost of requiring two robots to perform a sampling task.

### Incorporating flexible spatiotemporal structure

In general, the random arrival order assumption made in many secretary sampling applications is overly-restrictive, and very rarely holds in robotics and environmental monitoring domains. Initial work by Kesselheim et al. [56] is an interesting step towards considering how much structure can exist in a data stream such that standard secretary algorithms can be applied with performance guarantees. However, in robotics and environmental monitoring, the amount of structure in a data stream is often a problem constraint. Chapter 3 therefore took an alternative perspective: given a data stream with known spatiotemporal structure, we consider how secretary-style algorithms can be adapted to perform sampling with performance guarantees. However, given the difficulty of considering data streams with arbitrary spatiotemporal structure, Chapter 3 was constrained to consider only periodic spatiotemporal structure with known period and noise parameters.

One immediate extension to the work presented here would be learn parameters of the periodic data stream directly from an observed data stream, which would make the periodic secretary algorithm more flexible and robust in real data domains. On the other hand, many applications of secretary sampling in robotics or environmental monitoring will require consideration of data streams with non-periodic spatiotemporal structure. For these problems, it is interesting to understand what structure in the data stream can be modeled, and for which kinds of structure secretary sampling algorithms will perform well. If nothing is known about the data stream and a random arrival order assumption cannot be made, designing secretary sampling algorithms that perform better than random sampling may be impossible. However, modeling data streams with non-periodic spatiotemporal structure and designing effective, general purpose frameworks for secretary sampling and algorithm analysis is an interesting and essential area of future work for this problem area.

### 5.2.2 Hybrid deep-Bayesian environmental models

Chapter 4 explored hybrid deep learning-Bayesian models for visual scene understanding in unknown, unstructured environemnts. The approach leveraged recent progress in deep feature learning with convoluational autoencoders and spatiotemporal topic models and demonstrated the effectiveness of the technique on real data from a marine robot mission in the deep sea. Potential extensions to this preliminary implementation are detailed in the following sections.

#### Motion planning for perplexity maximization

The approach presented in Chapter 4 is concerned with how unsupervised learning can be used to discover latent thematic structure in image data streams from robotic missions. Unlike Chapter 3, the work in this chapter did not provide an integrated modeling and planning approach to robotic scientific information gathering. In order to apply this model within a scientific mission, the hybrid HDP-CAE model must first be incorporated within a POMDP or decision-making framework that considers the physical constraints of the autonomous agent and allows for planning. Chapter 4 introduced image or word perplexity as an interesting reward function for encouraging information-seeking behavior when using hybrid HDP-CAE to model a visual environment. Alternative information-theoretic reward functions are possible for this Bayesian model, such as those discussed for information planning in supervised topic models [99].

#### Alternative feature learning methods

Although Chapter 4 employs a convoluational autoencoder to learn feature embeddings for the discrete words in a topic modeling framework, it would be easy to swap an alternative feature learner into the pipeline. There has been interesting recent work in feature learning with variational autoencoders (VAEs) [100] and generative adversarial networks (GANs) [43]. An exploration of the efficacy of these techniques for the problem of feature learning for topic models could be enlightening. Generally,

the approach outlined in Chapter 4 learned a real-valued feature embedding for image data, which then had to be clustered to provide a set of discrete vocabulary words that are input into a topic modeling pipeline. Other methods have been considered for directly learning categorical labels or discrete embeddings [117], which could be effective when applied to this class of problem.

#### Multi-robot scene understanding

Multi-robot exploration is an very active area of research within the scientific robot community. In many application domains, such as marine robotics, transmitting high-dimensional raw percepts such as images between robots on a multi-robot team may be impossible. Topic models and other high-level semantic science models provide a potential mechanism for representing and sharing the information available in the raw sensory stream using a more abstract, compressed data representation. Preliminary studies have explored some immediate issues that arise when distributing unsupervised models, in which the lack of labels for topics or categories makes coming to model consensus challenging [27]. Forming and maintaining a global model of an unknown environment for exploration by a multi-robot team is a challenging open problem, and will likely require consideration from both a modeling and planning perspective.

## 5.3 Final Remarks

This thesis aims to provide a holistic perspective on scientific information gathering for autonomous robots and sensors. Chapter 1 began by formulating four key questions involved in enabling autonomous science, reiterated here: 1) How should scientific models be formalized for use by autonomous agents? 2) How should scientific objectives be specified for autonomous agents? 3) How should autonomous agents evaluate the utility of potential actions with respect to a scientific model and objective?, and 4) How should autonomous agents plan to take high-utility actions in a partially observable environment by levering problem structure? Although these

questions are difficult to answer universally, the preceding chapters proposed a sufficient set of answers that allowed for increased robustness, flexibility, and performance of scientific information gathering techniques in challenging environments.

Chapters 1 and 2 provided a high-level framework for formalizing the four questions related to scientific information-gathering by introducing probabilistic environmental models, general models for planning under uncertainty, and discussing the interplay between environmental models and planning problems via the specification of an information-theoretic reward function. Then, Chapter 3 provided a complete framework for scientific information gathering in secretary sampling problems, from the specification of an environmental model and reward function, to the development of a novel secretary sampling algorithm and algorithmic analysis for planning informative actions within this environmental model. Finally, Chapter 4 considered problems for which environmental modeling can be challenging: the percepts are high-dimensional image data and very little or nothing is known about the structure of the data stream *a priori*.

The approaches in Chapters 3 and 4 illustrate the depth and breadth of the problems facing the field of autonomous science. The directions for future work outlined in Section 5.2 further expose how classical problems in robotics – e.g., motion and path planning, multi-robot coordination, map building and localization – and problems from other related areas – e.g., machine learning, spatial statistics, optimization, and theory of experimental design – can play a role in addressing aspects of autonomous scientific information gathering. Developing robust autonomous systems that enhance our ability to perform exploratory science in diverse application areas, such as the oceans, the surface of extraterrestrial planets, and in agricultural and disaster-relief zones, will require insight and techniques such as those presented in this thesis, which unify theory and practice from these diverse disciplines.

# Bibliography


[1] Milton Abramowitz and Irene A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York, 1964.

[2] Cedric Archambeau, Dan Cornford, Manfred Opper, and John Shawe-Taylor. Gaussian process approximations of stochastic differential equations. *Proceedings of Machine Learning Research*, June 2007.

[3] Akash Arora, P. Michael Furlong, Robert Fitch, Salah Sukkarieh, and Terrence Fong. Multi-modal active perception for information gathering in science missions. *CoRR*, 2017.

[4] Moshe Babaioff, Nicole Immorlica, David Kempe, and Robert Kleinberg. A Knapsack Secretary Problem with Applications. *Approximation, randomization, and combinatorial optimization. Algorithms and techniques*, 2007.

[5] Mohammadhossein Bateni, Mohammadtaghi Hajiaghayi, and Morteza Zadimoghaddam. Submodular Secretary Problem and Extensions. *APPROX-RANDOM*, 6302, 2010.

[6] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. SURF: Speeded Up Robust Features. In *European Conference on Computer Vision*, 2006.

[7] Jonathan Binney, Andreas Krause, and Gaurav S Sukhatme. Informative Path Planning for an Autonomous Underwater Vehicle. In *Robotics and automation (ICRA), 2010 IEEE international conference on*, 2010.

[8] Christopher M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag, Berlin, Heidelberg, 2006.

[9] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3:993–1022, March 2003.

[10] Vincent D Blondel and John N Tsitsiklis. A survey of computational complexity results in systems and control. *Automatica*, 36(9):1249–1274, 2000.

[11] Anna Bosch, Andrew Zisserman, and Xavier Muñoz. Scene classification via pLSA. In *European Conference on Computer Vision*, 2006.

[12] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.

[13] Mateusz Buda, Atsuto Maki, and Maciej A Mazurowski. A systematic study of the class imbalance problem in convolutional neural networks. *arXiv preprint arXiv:1710.05381*, 2017.

[14] Alberto Candela, David Thompson, Eldar Noe Dobrea, and David Wettergreen. Planetary robotic exploration driven by science hypotheses for geologic mapping. In *Intelligent Robots and Systems (IROS), 2017 IEEE/RSJ International Conference on*, pages 3811–3818. IEEE, 2017.

[15] Luca Carlone, Jingjing Du, Miguel Kaouk Ng, Basilio Bona, and Marina Indri. Active SLAM and exploration with particle filters using Kullback-Leibler divergence. *Journal of Intelligent & Robotic Systems*, 75(2):291–311, 2014.

[16] Henry Carrillo, Philip Dames, Vijay Kumar, and José A Castellanos. Autonomous robotic exploration using a utility function based on Rényi's general theory of entropy. *Autonomous Robots*, 42(2):235–256, 2018.

[17] Henry Carrillo, Yasir Latif, Maria L Rodriguez-Arevalo, José Neira, and José A Castellanos. On the monotonicity of optimality criteria during exploration in active SLAM. In *Robotics and Automation (ICRA), 2015 IEEE International Conference on*, pages 1476–1483. IEEE, 2015.

[18] Henry Carrillo, Ian Reid, and José A Castellanos. On the comparison of uncertainty criteria for active SLAM. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 2080–2087. IEEE, 2012.

[19] Jonathan Chang, Jordan Boyd-Graber, Sean Gerrish, Chong Wang, and David M. Blei. Reading Tea Leaves: How Humans Interpret Topic Models. In *Advances in Neural Information Processing Systems*, pages 1–9, 2009.

[20] Nitesh V Chawla, Nathalie Japkowicz, and Aleksander Kotcz. Special issue on learning from imbalanced data sets. *ACM Sigkdd Explorations Newsletter*, 6(1):1–6, 2004.

[21] Sundeep Prabhakar Chepuri and Geert Leus. Sparsity-promoting sensor selection for non-linear measurement models. *IEEE Trans. Signal Processing*, 63(3):684–698, 2015.

[22] Jean-Paul Chiles and Pierre Delfiner. *Geostatistics: modeling spatial uncertainty*. Wiley, New York, 1999.

[23] Thomas M Cover and Joy A Thomas. *Elements of information theory*. John Wiley & Sons, 2012.

[24] Noel Cressie. Statistics for spatial data. *Terra Nova*, 4(5):613–617, 1992.

[25] Patrick Dallaire, Stéphane Ross, and Brahim Chaib-draa. GP-POMDP : Bayesian Reinforcement Learning in Continuous POMDPs with Gaussian Processes. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2604–2609, 2009.

[26] Jnaneshwar Das, Frédéric Py, Julio B J Harvey, John P Ryan, Alyssa Gellene, Rishi Graham, David A Caron, Kanna Rajan, and Gaurav S Sukhatme. Data - driven robotic sampling for marine ecosystem monitoring. *The International Journal of Robotics Research*, 34, 2015.

[27] Kevin Doherty, Genevieve Flaspohler, Nicholas Roy, and Yogesh Girdhar. Approximate distributed spatiotemporal topic models for multi-robot terrain characterization. In *Intelligent Robots and Systems (IROS), 2018 IEEE/RSJ International Conference on*. IEEE, 2018.

[28] Joseph L Doob and Joseph L Doob. *Stochastic processes*. Wiley New York, 1953.

[29] Alberto Elfes. Using occupancy grids for mobile robot perception and navigation. *Computer*, 6:46–57, 1989.

[30] Valerii Vadimovich Fedorov. *Theory of optimal experiments*. Elsevier, 1972.

[31] Li Fei-Fei and Pietro Perona. A Bayesian Hierarchical Model for Learning Natural Scene Categories. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. IEEE, 2005.

[32] Thomas S Ferguson. Who Solved the Secretary Problem? *Statistical Science*, 4(3), 1989.

[33] Steve Fisk. The Secretary. *Mathematics Magazine*, 61(2), 1988.

[34] Genevieve Flaspohler, Nicholas Roy, and Yogesh Girdhar. Feature discovery and visualization of robot mission data using convolutional autoencoders and Bayesian nonparametric topic models. In *Intelligent Robots and Systems (IROS), 2017 IEEE/RSJ International Conference on*. IEEE, 2017.

[35] Genevieve Flaspohler, Nicholas Roy, and Yogesh Girdhar. Near-optimal irrevocable sample selection for periodic data streams with applications to marine robotics. In *Robotics and Automation (ICRA), 2018 IEEE International Conference on*. IEEE, 2018.

[36] Ariell Friedman, Daniel Steinberg, Oscar Pizarro, and Stefan B Williams. Active learning using a variational dirichlet process model for pre-clustering and classification of underwater stereo imagery. In *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*, pages 1533–1539. IEEE, 2011.

[37] Enric Galceran, Ricard Campos, Narcis Palomeras, Marc Carreras, and Pere Ridao. Coverage path planning with realtime replanning for inspection of 3D underwater structures. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6586–6591. IEEE, May 2014.

[38] Yogesh Girdhar and Gregory Dudek. Online Navigation Summaries. In *IEEE International Conference on Robotics and Automation*, 2010.

[39] Yogesh Girdhar and Gregory Dudek. Gibbs Sampling Strategies for Semantic Perception of Streaming Video Data. *ArXiv e-prints*, page 7, 2015.

[40] Yogesh Girdhar, Philippe Giguère, and Gregory Dudek. Autonomous Adaptive Underwater Exploration using Online Topic Modelling. In *International Symposium on Experimental Robotics (ISER)*, 2012.

[41] Yogesh Girdhar, Philippe Giguere, and Gregory Dudek. Autonomous adaptive exploration using realtime online spatiotemporal topic modeling. *The International Journal of Robotics Research*, 33(4):645–657, 11 2013.

[42] Yogesh Girdhar, Walter Cho, Matthew Campbell, Jesus Pineda, Elizabeth Clarke, and Hanumant Singh. Anomaly detection in unstructured environments using Bayesian nonparametric scene modeling. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2651–2656. IEEE, 5 2016.

[43] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.

[44] GPy. GPy: A gaussian process framework in python, 2012.

[45] Thomas L Griffiths and Mark Steyvers. Finding scientific topics. *Proceedings of the National academy of Sciences*, 101:5228–5235, 2004.

[46] Carlos Guestrin, Andreas Krause, and Ajit Paul Singh. Near-optimal sensor placements in Gaussian processes. In *Proceedings of the 22nd international conference on Machine learning - ICML '05*, 2005.

[47] Haibo He and Edwardo A Garcia. Learning from imbalanced data. *IEEE Transactions on Knowledge & Data Engineering*, (9):1263–1284, 2008.

[48] Gregory Hitz, Enric Galceran, Marie-Ève Garneau, François Pomerleau, and Roland Siegwart. Adaptive continuous-space informative path planning for online environmental monitoring. *Journal of Field Robotics*, 34(8):1427–1449, December 2017.

[49] Geoffrey A. Hollinger and Gaurav S. Sukhatme. Sampling-based robotic information gathering algorithms. *The International Journal of Robotics Research*, 33(9):1271–1287, August 2014.

[50] Vadiraj Hombal, Arthur Sanderson, and D. Richard Blidberg. Multiscale adaptive sampling in environmental robotics. In *2010 IEEE Conference on Multisensor Fusion and Integration*, pages 80–87. IEEE, September 2010.

[51] Thibaut Horel. Notes on Greedy Algorithms for Submodular Maximization. 2016.

[52] Syed Talha Jawaid and Stephen L. Smith. Informative path planning as a maximum traveling salesman problem with submodular rewards. *Discrete Applied Mathematics*, 186:112–127, May 2015.

[53] Siddharth Joshi and Stephen Boyd. Sensor selection via convex optimization. *IEEE Transactions on Signal Processing*, 57(2):451–462, 2009.

[54] Leslie Pack Kaelbling, Michael L Littman, and Anthony R Cassandra. Planning and acting in partially observable stochastic domains. *Artificial intelligence*, 101(1-2):99–134, 1998.

[55] Gautam Kamath. Bounds on the Expectation of the Maximum of Samples from a Gaussian.

[56] Thomas Kesselheim, Robert Kleinberg, and Rad Niazadeh. Secretary Problems with Non-Uniform Arrival Order. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, 2015.

[57] Robert Kleinberg. A Multiple-Choice Secretary Algorithm with Applications to Online Auctions. In *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, 2005.

[58] C.-W. Ko, Jon Lee, and Maurice Queyranne. An Exact Algorithm for Maximum Entropy Sampling. *Operations Research*, 43(4), August 1995.

[59] Thomas Kollar and Nicholas Roy. Trajectory optimization using reinforcement learning for map exploration. *The International Journal of Robotics Research*, 27(2):175–196, 2008.

[60] Andreas Krause and Carlos Guestrin. Near-optimal nonmyopic value of information in graphical models. In *Proceedings of the Twenty-First Conference on Uncertainty in Artificial Intelligence*, UAI'05, pages 324–331, Arlington, Virginia, United States, 2005. AUAI Press.

[61] Andreas Krause and Carlos Guestrin. Optimal value of information in graphical models. *Journal of Artificial Intelligence Research*, 35:557–591, 2009.

[62] Andreas Krause, Carlos Guestrin, Anupam Gupta, and Jon Kleinberg. Near-optimal sensor placements: Maximizing information while minimizing communication cost. In *Proceedings of the 5th international conference on Information processing in sensor networks*, pages 2–10. ACM, 2006.

[63] Andreas Krause, Ajit Singh, and Carlos Guestrin. Near-Optimal Sensor Placements in Gaussian Processes: Theory, Efficient Algorithms and Empirical Studies. *Journal of Machine Learning Research*, 9, 2008.

[64] Ralf Krestel, Peter Fankhauser, and Wolfgang Nejdl. Latent dirichlet allocation for tag recommendation. In *Proceedings of the third ACM conference on Recommender systems*, pages 61–68. ACM, 2009.

[65] John J Leonard and Hugh F Durrant-Whyte. Mobile robot localization by tracking geometric beacons. *IEEE Transactions on robotics and Automation*, 7(3):376–382, 1991.

[66] Daniel Levine, Brandon Luders, and Jonathan How. Information-rich path planning with general constraints using rapidly-exploring random trees. In *AIAA Infotech@ Aerospace 2010*, page 3360. 2010.

[67] Daniel Levine, Brandon Luders, and Jonathan P How. Information-theoretic motion planning for constrained sensor networks. *Journal of Aerospace Information Systems*, 10(10):476–496, 2013.

[68] D. V. Lindley. Dynamic Programming and Decision Theory. *Applied Statistics*, 10(1), 1961.

[69] Chun Kai Ling, Kian Hsiang Low, and Patrick Jaillet. Gaussian process planning with lipschitz continuous reward functions: Towards unifying bayesian optimization, active learning, and beyond. In *AAAI*, pages 1860–1866, 2016.

[70] William S Lovejoy. A survey of algorithmic methods for partially observed markov decision processes. *Annals of Operations Research*, 28(1):47–65, 1991.

[71] David G Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2), 2004.

[72] Wenhao Luo, Changjoo Nam, and Katia Sycara. Online decision making for stream-based robotic sampling via submodular optimization. In *2017 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, pages 118–123. IEEE, November 2017.

[73] Sandeep Manjanna, Alberto Quattrini Li, Ryan N. Smith, Ioannis M. Rekleitis, and Gregory Dudek. Heterogeneous multirobot system for exploration and strategic water sampling. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, 2018.

[74] Roman Marchant and Fabio Ramos. Bayesian Optimisation for informative continuous path planning. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6136–6143. IEEE, 2014.

[75] Sonia Martínez and Francesco Bullo. Optimal sensor placement and motion coordination for target tracking. *Automatica*, 42(4):661–668, 2006.

[76] Ruben Martinez-Cantin, Nando de Freitas, Eric Brochu, José Castellanos, and Arnaud Doucet. A Bayesian exploration-exploitation approach for optimal online sensing and planning with a visually guided mobile robot. *Autonomous Robots*, 27(2):93–103, August 2009.

[77] Jonathan Masci, Ueli Meier, Dan Cireşan, and Jürgen Schmidhuber. Stacked convolutional auto-encoders for hierarchical feature extraction. In *International Conference on Artificial Neural Networks*, pages 52–59. Springer, 2011.

[78] Mitchell McIntire, Daniel Ratner, and Stefano Ermon. Sparse gaussian processes for bayesian optimization. In *UAI*, 2016.

[79] Tomas Mikolov and Geoffrey Zweig. Context dependent recurrent neural network language model. *SLT*, 12(234-239):8, 2012.

[80] Werner G Müller. *Collecting spatial data: optimum design of experiments for random fields*. Springer Science & Business Media, 2007.

[81] Tayyab Naseer, Gabriel L Oliveira, Thomas Brox, and Wolfram Burgard. Semantics-aware visual localization under challenging perceptual conditions. In *Robotics and Automation (ICRA), 2017 IEEE International Conference on*, pages 2614–2620. IEEE, 2017.

[82] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher. An analysis of approximations for maximizing submodular set functions - I. *Mathematical Programming*, 14(1), 1978.

[83] Joseph L. Nguyen, Nicholas R.J. Lawrance, and Salah Sukkarieh. Nonmyopic planning for long-term information gathering with an aerial glider. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6573–6578. IEEE, May 2014.

[84] Robert J. Olson and Heidi M. Sosik. A submersible imaging-in-flow instrument to analyze nano- and microplankton: Imaging FlowCytobot. *Limnology and Oceanography: Methods*, June 2007.

[85] Christos H Papadimitriou and John N Tsitsiklis. The complexity of markov decision processes. *Mathematics of operations research*, 12(3):441–450, 1987.

[86] Andrej Pázman. *Foundations of optimum experimental design*, volume 14. Springer, 1986.

[87] Jesús Pineda, Walter Cho, Victoria Starczak, Annette F Govindarajan, Héctor M Guzman, Yogesh Girdhar, Rusty C Holleman, James Churchill, Hanumant Singh, and David K Ralston. A crab swarm at an ecological hotspot: patchiness and population density from auv observations at a coastal, tropical seamount. *PeerJ*, 4, 2016.

[88] Marija Popović, Gregory Hitz, Juan Nieto, Inkyu Sa, Roland Siegwart, and Enric Galceran. Online informative path planning for active classification using UAVs. In *Robotics and Automation (ICRA), 2017 IEEE International Conference on*, pages 5753–5758. IEEE, 2017.

[89] Dushyant Rao, Mark De Deuge, Navid Nourani-Vatani, Bertrand Douillard, Stefan B. Williams, and Oscar Pizarro. Multimodal learning for autonomous underwater vehicles from visual and bathymetric data. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3819–3825. IEEE, may 2014.

[90] Rasmussen and Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[91] Nicholas Roy, Wolfram Burgard, Dieter Fox, and Sebastian Thrun. Coastal navigation-mobile robot navigation with uncertainty in dynamic environments. In *Robotics and Automation, 1999. Proceedings. 1999 IEEE International Conference on*, volume 1, pages 35–40. IEEE, 1999.

[92] Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. ORB: An efficient alternative to SIFT or SURF. In *2011 International Conference on Computer Vision*, pages 2564–2571, Barcelona, 11 2011. IEEE.

[93] Allison Ryan and J Karl Hedrick. Particle filter based information-theoretic active sensing. *Robotics and Autonomous Systems*, 58(5):574–584, 2010.

[94] Mac Schwager, Philip Dames, Daniela Rus, and Vijay Kumar. A multi-robot control policy for information gathering in the presence of unknown hazards. In *Robotics research*, pages 455–472. Springer, 2017.

[95] Dravyansh Sharma, Ashish Kapoor, and Amit Deshpande. On greedy maximization of entropy. In *International Conference on Machine Learning*, pages 1330–1338, 2015.

[96] Yelong Shen, Xiaodong He, Jianfeng Gao, Li Deng, and Grégoire Mesnil. A Latent Semantic Model with Convolutional-Pooling Structure for Information Retrieval. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management - CIKM '14*, pages 101–110, New York, New York, USA, 2014. ACM Press.

[97] David Silver and Joel Veness. Monte-Carlo planning in large POMDPs. In *Advances in neural information processing systems*, pages 2164–2172, 2010.

[98] Robert Sim and Nicholas Roy. Global A-optimal robot exploration in SLAM. In *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*, pages 661–666. IEEE, 2005.

[99] Vadim Smolyakov, Jason Pacheco, and John W Fisher III. Information planning for text data. *arXiv preprint arXiv:1802.03360*, 2018.

[100] Kihyuk Sohn, Honglak Lee, and Xinchen Yan. Learning structured output representation using deep conditional generative models. In *Advances in Neural Information Processing Systems*, pages 3483–3491, 2015.

[101] Niranjan Srinivas, Andreas Krause, Sham Kakade, and Matthias Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. In *Proceedings of the 27th International Conference on International Conference on Machine Learning*, ICML'10, pages 1015–1022. Omnipress, 2010.

[102] Niranjan Srinivas, Andreas Krause, Sham M. Kakade, and Matthias W. Seeger. Information-theoretic regret bounds for Gaussian process optimization in the bandit setting. *IEEE Transactions on Information Theory*, 58(5), may 2012.

[103] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, 15:1929–1958, 2014.

[104] Vaibhav Srivastava, Kurt Plarre, and Francesco Bullo. Adaptive sensor selection in sequential hypothesis testing. In *IEEE Conference on Decision and Control and European Control Conference*, pages 6284–6289. IEEE, December 2011.

[105] D Steinberg, Ariell Friedman, Oscar Pizarro, and Stefan B Williams. A bayesian nonparametric approach to clustering data from underwater robotic surveys. In *International Symposium on Robotics Research*, volume 28, pages 1–16, 2011.

[106] Daniel M Steinberg, Oscar Pizarro, and Stefan B Williams. Hierarchical bayesian models for unsupervised scene understanding. *Computer Vision and Image Understanding*, 131:128–144, 2015.

[107] Daniel M Steinberg, Oscar Pizarro, Stefan B Williams, and Michael V Jakuba. Dirichlet process mixture models for autonomous habitat classification. In *OCEANS 2010 IEEE-Sydney*, pages 1–7. IEEE, 2010.

[108] Daniel M Steinberg, Stefan B Williams, Oscar Pizarro, and Michael V Jakuba. Towards autonomous habitat classification using gaussian mixture models. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 4424–4431. IEEE, 2010.

[109] Ethan Stump, Vijay Kumar, Ben Grocholsky, and Pedro M Shiroma. Control for localization of targets using range-only sensors. *The International Journal of Robotics Research*, 28(6):743–757, 2009.

[110] Wen Sun, Niteesh Sood, Debadeepta Dey, Gireeja Ranade, Siddharth Prakash, and Ashish Kapoor. No-regret replanning under uncertainty. In *Robotics and Automation (ICRA), 2017 IEEE International Conference on*, pages 6420–6427. IEEE, 2017.

[111] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 1998.

[112] Yee Whye Teh and Michael I. Jordan. Hierarchical Bayesian Nonparametric Models with Applications. *Bayesian nonparametrics*, pages 158–207, 2010.

[113] Yee Whye Teh, Michael I Jordan, Matthew J Beal, and David M Blei. Hierarchical Dirichlet Processes. *Journal of the American Statistical Association*, 101(476):1566–1581, December 2006.

[114] David R Thompson, David S Wettergreen, and Francisco J Calderóon Peralta. Autonomous science during large-scale robotic survey. *Journal of Field Robotics*, 28(4):542–564, 2011.

[115] Sebastian Thrun and Arno Bücken. Integrating grid-based and topological maps for mobile robot navigation. In *Proceedings of the National Conference on Artificial Intelligence*, pages 944–951, 1996.

[116] Antonio Torralba, Joshua B Tenenbaum, and Ruslan R Salakhutdinov. Learning to learn with compound hd models. In *Advances in Neural Information Processing Systems*, pages 2061–2069, 2011.

[117] Aaron van den Oord, Oriol Vinyals, et al. Neural discrete representation learning. In *Advances in Neural Information Processing Systems*, pages 6306–6315, 2017.

[118] Li Wan, Leo Zhu, and Rob Fergus. A hybrid neural network-latent topic model. In *Artificial Intelligence and Statistics*, pages 1287–1294, 2012.

[119] Xiaogang Wang and Eric Grimson. Spatial Latent Dirichlet Allocation. In *Advances in Neural Information Processing Systems*, volume 20, page 1577–1584, 2007.

[120] Jason L Williams. *Information theoretic sensor management*. PhD thesis, Massachusetts Institute of Technology, 2007.

[121] N. Yilmaz, Constantinos Evangelinos, Pierre F. J. Lermusiaux, and N. M. Patrikalakis. Path Planning of Autonomous Underwater Vehicles for Adaptive Sampling Using Mixed Integer Linear Programming. *IEEE Journal of Oceanic Engineering*, 33(4):522–537, October 2008.

[122] Bin Zhang and Gaurav S. Sukhatme. Adaptive Sampling for Estimating a Scalar Field using a Robotic Boat and a Sensor Network. In *Proceedings 2007 IEEE International Conference on Robotics and Automation*, pages 3673–3680. IEEE, April 2007.

[123] Sue Zheng, Jason Pacheco, and John Fisher. A robust approach to sequential information theoretic planning. In *International Conference on Machine Learning*, pages 5936–5944, 2018.

[124] Yan Zhu and Eamonn Keogh. Irrevocable-choice algorithms for sampling from a stream. *Data Mining and Knowledge Discovery*, 30(5), 2016.