

Adaptive Science Operations in Deep Space Missions using Robust Precomputed Autonomy

Grace Kim^{1*}, Hailey Warner^{1*}, Duncan Eddy¹, Evan Astle²,
Zachary Booth², Edward Balaban², Mykel J. Kochenderfer¹

*Equal contribution.

Abstract—Robust autonomy in deep space science mission operations is essential where communication delays prevent real-time ground control from addressing challenges of environmental uncertainty. This is problematic for astrobiological missions, which must navigate strict contamination constraints while coordinating multi-instrument measurements to interpret transient, time-sensitive biosignatures. To address these challenges, we present an offline autonomy framework that performs adaptive, resource-aware sequencing of scientific instruments under uncertainty. It integrates a Bayesian network, which probabilistically models the likelihood of observing different biotic signatures from a sample, with a partially observable Markov decision process (POMDP) to produce adaptive instrument sequences. As a case study, we apply our methodology to the Enceladus Orbilander’s proposed Life Detection Suite (LDS) of science instruments. We demonstrate how to design Bayesian networks that accurately model the likelihood of detecting biological signatures in collected samples and display how POMDP reward function tuning can be used to optimize competing objectives such as reducing sample state uncertainty and maximizing detection sensitivity. Our method’s performance is evaluated against the recorded Enceladus Orbilander Concept of Operations (ConOps) for binary classifier accuracy and efficiency. These results establish our method as a significant step toward realizing robust deep space autonomy.

I. INTRODUCTION

Scientific exploration in extreme environments—such as planetary surfaces, subsurface oceans, or deep space—demands systems capable of operating reliably under harsh and uncertain conditions. Spacecraft may experience power loss, communication blackouts, component failures, and other off-nominal events inhibiting mission success. To mitigate these risks, operators have traditionally opted for conservative approaches to manage spacecraft operations, with persistent human-in-the-loop monitoring or use of fixed operation schedules. However, this practice is growing increasingly unsustainable as upcoming astrobiology deep space missions such as NASA’s Europa Clipper [1] and ESA’s JUICE [2] face over hour-long communication delays and blackouts [3]. Although some biosignatures, like amino acids, can have delayed analysis as they remain stable in extreme environments, cell membranes can degrade through cell lysis within hours to days once sampled [4], [5]. Testing delays exceeding this window could compromise microscopy and lead to false negatives with

degraded samples. This challenge stresses the need for autonomous systems capable of making timely, reliable decisions onboard to maximize science return.

Advances in onboard computing and scientific instrumentation have opened new opportunities to incorporate autonomy into science operations. Researchers have already demonstrated that autonomy can improve efficiency in real-time data processing, navigation, trajectory planning, and sampling [6], as seen with the Perseverance rover’s autonomous navigation and target selection on Mars [7] and the Deep-space Autonomous Robotic Explorer’s onboard trajectory optimization for asteroid reconnaissance [8]. However, its use in high-stakes missions such as astrobiological discovery remains limited, as the cost of false positives, contamination, or sample loss has been unacceptably high. Enabling adoption of advanced autonomy in these contexts requires efficient yet verifiable and transparent techniques. By extensively validating precomputed autonomous solutions prior to deployment, mission operators can gain trust in system behavior and rely on its ability to consistently meet mission requirements.

To address these challenges, we propose an offline policy generation framework that allows for pre-launch verification while optimizing science return. We model the instrument operation problem as a partially observable Markov decision process (POMDP), which captures the uncertainty and sequential nature of decision-making in deep space missions. To efficiently represent the complex and correlated sensor observations, we use a Bayesian network to define the observation function. This approach encodes probability distributions over sample characteristics and incorporates expert understanding of biosignatures, making the observation space computationally tractable. Our methodology also offers a principled, quantitative means to model and prioritize scientific objectives, reducing ambiguity. We then solve the resulting POMDP using an offline solver, which precomputes an optimal policy that can be verified thoroughly before onboard, real-time deployment. We apply this approach to the Enceladus Orbilander mission concept, automating its life detection instrument suite. To evaluate our method, we compare its performance against the existing Orbilander concept of operations, using metrics such as life detection false positive rates, false negative rates, and operational efficiency. Our results demonstrate significant improvements over this baseline.

II. RELATED WORK

A. Autonomy in Space

Deep space missions have successfully deployed autonomy in spacecraft operations in the past 20 years, primarily in

[1] Department of Aeronautics and Astronautics, Stanford University, Stanford, CA, 94305, USA gkim65@stanford.edu hlwarner@stanford.edu deddy@stanford.edu mykel@stanford.edu

[2] Intelligent Systems Division, NASA Ames Research Center, Moffett Field, CA, 94035, USA david.e.astle@nasa.gov zachary.m.booth@nasa.gov edward.balaban@nasa.gov

the context of station keeping and collision avoidance. In 1999, NASA's Deep Space 1 (DS1) mission marked the first demonstration of autonomous image-based localization, fault diagnosis, and system management during a flyby of asteroid 9969 Braille and comet Borrelly [9]. Dragonfly and Europa Clipper, currently the two most high-profile and advanced astrobiology missions, navigate semi-autonomously. Both missions also employ autonomy to assist with instrument data processing and fault detection for mission operations [10], [11].

Autonomy frameworks to optimize science return have also been developed for the Europa Lander mission, which would face similar communication delays and engineering challenges to the Enceladus Orbilander. Wagner et al. designed a hierarchical utility model for the Europa Lander to maximize science return [12]. However, this specific architecture is unable to generalize to unforeseen events such as an actuator fault, as the utility of each measurement and action must be hard-coded into the hierarchical model. Techniques like reinforcement learning can overcome this with a sufficiently descriptive reward function.

B. Bayesian Networks and POMDPs

Science planning in astrobiology missions requires reasoning under uncertainty, where many interdependent variables must be assessed and acted on with limited information. Bayesian networks, a type of probabilistic graphical model, are well-suited to quantify this uncertainty through the graph's compact encoding of conditional dependencies between variables. These networks have been widely applied in the context of systems health monitoring, from chemical process fault detection to coastal erosion prediction [13], [14]. Bayesian networks are especially fitting for astrobiological life detection, which involves reasoning over the presence or absence of complex biosignatures. For example, environmental pH may alter the likelihood of observing certain biosignatures while not affecting others.

POMDPs are a type of mathematical framework used in sequential decision-making under uncertainty, where planning is needed despite incomplete information. POMDPs are often used in model-based reinforcement learning problems like autonomous spacecraft stationkeeping [15], [16] and robust science planning during communication blackouts [17]. Unlike simpler reinforcement learning methods that only optimize immediate rewards or need large amounts of training, POMDPs can be used to produce verifiable offline policies when real time planning is too time-consuming or risky. In this work, we integrate Bayesian networks and POMDPs into a single autonomy framework to support advanced deep space science operations.

III. METHODOLOGY

The fundamental challenges of autonomy in deep space mission operations can be addressed with POMDPs. These challenges include incomplete state information from sensor constraints, stochastic environmental dynamics, and the need to perform sequential decision-making under uncertainty. However, model formulations for POMDPs can grow quickly

in complexity due to the high-dimensional space of possible observations. To address this issue, we use a Bayesian network to structure the POMDP observation space, which compactly encodes conditional dependence between measurements. In the following subsections, we first detail the integration of Bayesian networks into the POMDP observation model. We then outline the POMDP formulation, including definitions of state, action, and observation spaces. Finally, we describe the process of deriving verifiable, optimal policies (i.e. state-action mappings) using offline solvers.

A. Bayesian Networks

We apply Bayesian networks to represent complex conditional dependencies in potential life detection instrument measurements. Such dependencies frequently arise in deep space missions focused on astrobiological discovery, where multiple sensors may find correlated observations of the same underlying phenomena. Bayesian networks are an effective tool for compactly representing these joint distributions and conditional dependencies between random variables. Structurally, networks take on the form of a directed, acyclic graph (DAG). We define a random variable X_1 to be dependent on X_2 if $P(X_1 | X_2) \neq P(X_1)$, and we represent this dependency by a directed edge in the Bayesian network.

Capturing variable dependencies with a DAG can significantly reduce the dimensionality of the observation space. Rather than requiring $m^n - 1$ parameters to describe a joint distribution among n variables with m possible values each, a quantity that grows exponentially, Bayesian networks reduce the number of required parameters to $\sum_{i=1}^n m_i(k_i - 1)$. This formulation ensures that, while the number of parameters needed for each variable grows exponentially with the number of parents, the total number of parameters across the network grows linearly with n , the number of variables, assuming bounded parent sets. This is made possible by factoring the joint distribution through the following chain rule

$$P(x_{1:n}) = \prod_{i=1}^n P(x_i | \text{parent}(x_i)) \quad (1)$$

Equation (1) which implies that the joint probability distribution over random variables is equivalent to the product of all conditional probability distributions. For example, a Bayesian network with structure $A \rightarrow B \rightarrow C$ can be fully described by $P(A, B, C) = P(A) \cdot P(B|A) \cdot P(C|B)$.

The structure and probability distributions of a Bayesian network can be designed by domain experts or learned from data [18]. A network for life detection must largely be expert-designed due to epistemic uncertainty about deep space environments and extraterrestrial life. Here we choose a *causal graph* structure, meaning parents directly influence the observations their child nodes produce. Causal graph structures tend to have sparser connections between network nodes, which mimics the sparsity found in graph structures in real-world applications [19].

B. Partially Observable Markov Decision Process

A partially observable Markov decision process (POMDP) is a mathematical framework for modeling sequential decision-

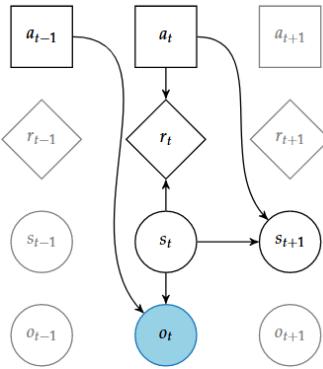


Fig. 1. Sequential decision network illustrating the POMDP formulation [18].

making problems under uncertainty [18]. In a POMDP, an agent (i.e. a spacecraft) interacts with an environment characterized by a set of hidden states \mathcal{S} . At any timestep, the agent exists in some true state $s \in \mathcal{S}$, but cannot directly observe it. Instead, the agent receives a noisy or incomplete observation $o \in \mathcal{O}$, generated according to the observation function $O(o | s, a)$, which specifies the likelihood of observing o given the current state s and action a . Using o at each time step, the agent selects an action $a \in \mathcal{A}$ from the set of possible actions. The state then updates to $s' \in \mathcal{S}$ according to the transition function $T(s' | s, a)$. This function returns the likelihood of entering the next state s' given the current state s and action a . After transitioning, the agent receives a reward $R(s, a)$, which quantifies the immediate benefit of that state and action. A POMDP \mathcal{P} is then defined by the tuple

$$\mathcal{P} = (\mathcal{S}, \mathcal{A}, \mathcal{O}, T, O, R) \quad (2)$$

which describes state, action, and observation spaces along with transition, observation, and reward functions. A visual representation of the sequential decision framework of a POMDP is shown in Figure 1.

To solve a POMDP, the agent maintains a belief B , a probability distribution over possible states, to track its uncertainty about the true state at each timestep. The agent updates its belief at each timestep based on its action and observation. Various solvers attempt to maximize the agent's expected cumulative reward given the agent's evolving belief. These mappings of an agent's belief to an optimal action are referred to as *policies*, which a variety of available solvers can compute. In this work, we use the Successive Approximations of the Reachable Space under Optimal Policies (SARSOP) algorithm [20] to precompute a policy that can be thoroughly validated and tested prior to deployment. SARSOP approximates the optimal value function by only sampling reachable belief points under optimal policies. This method significantly improves efficiency over exact methods and is well-suited for problems with large or continuous state spaces.

IV. ENCELADUS ORBILANDER CASE STUDY

To illustrate the practical application of our methodology, we present a case study focused on the Enceladus Orbilander mission concept, with particular attention to autonomous

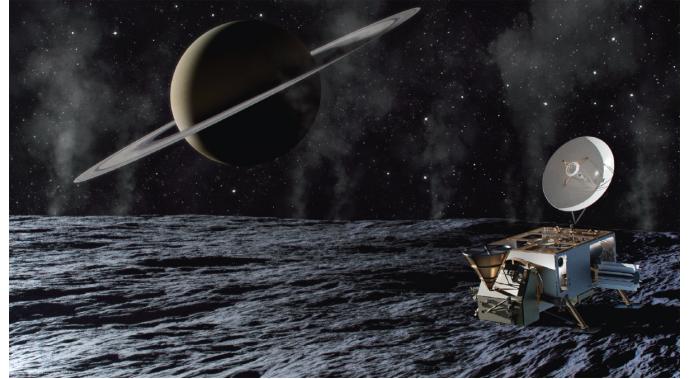


Fig. 2. Artist's impression of Orbilander on the surface of Enceladus sometime in the 2050s [21].

control of the Life Detection Suite (LDS) instruments. We provide a brief mission overview of the Enceladus Orbilander and describe the design of a Bayesian network that models LDS instrument measurements. We then formulate a POMDP that uses these observations to support efficient autonomous science operations for life detection.

A. Mission Overview

The Enceladus Orbilander, developed in partnership with NASA and APL, is a flagship mission concept designed to search for evidence of life and assess the habitability of Enceladus, a Saturnian moon known for its subsurface ocean and ice plumes. The mission architecture includes both orbiter and lander phases, enabling comprehensive characterization of biosignatures ranging from polyelectrolytes, cells, biotic amino acids, to other indirect metrics like pH. The Orbilander is equipped with a sophisticated life detection suite of six instruments: two mass spectrometers, an electrochemical sensor array, an organic analyzer, a microscope, and a nanopore sequencer. During the orbiter phase, the spacecraft is expected to collect between 1 and 20 samples, depending on plume activity and variability [21].

The Orbilander faces significant operational challenges, including stringent power and communication constraints. The 12 hour orbital period restricts the duration and frequency of ground communication windows, limiting opportunities for real-time decision-making. Moreover, the radioisotope thermoelectric generators (RTG) are projected to lose 13% of their battery capacity by Saturn orbit insertion and up to 22% by landing, potentially limiting instrument use and mission lifespan [21]. These factors underscore the importance of robust, precomputed autonomy to ensure efficient and adaptive science operations even during off-nominal scenarios. We formulate a POMDP of the Enceladus Orbilander LDS operations to address these concerns, computing a policy that adaptively and autonomously sequences science instruments. A Bayesian network integrated into the POMDP models dependencies among candidate biosignatures and measurement uncertainty. In the following section, we detail the structure and role of the Bayesian network used to model the observations of our POMDP decision framework.

TABLE I
SELECTED BIOSIGNATURES AND HABITABILITY METRICS

| Name | Characteristic | Domain | Parent Nodes |
|----------|------------------------------------|-----------------|--------------|
| C_0 | Life | {0,1} | - |
| C_1 | Polyelectrolyte Presence | {0,1} | C_0 |
| C_2 | Cell Membrane Presence | {0,1} | C_0 |
| C_3 | Autofluorescence | {0,1} | C_0 |
| C_4 | Molecular Assembly Index ≥ 15 | {0,1} | C_0 |
| C_5 | Biotic Amino Acid Diversity | {0, 1, ..., 22} | C_0 |
| C_6 | L:R Chirality Ratio Percentage | [0, 100] | C_0 |
| C_7 | Salinity Percentage | [0, 100] | C_2 |
| C_8 | CHNOPS Abundance Percentage | [0, 100] | C_4, C_5 |
| C_9 | pH | [0, 14] | C_1, C_5 |
| C_{10} | Redox Potential [V] | [-0.5, 0] | C_5 |

B. Bayesian Network Design

Before we introduce the life detection suite POMDP, we first present a Bayesian network designed to model dependencies between biosignatures and environmental metrics relevant to habitability of Enceladus. The network enables us to rigorously quantify how different observed characteristics inform the likelihood of a sample having biotic or abiotic properties while condensing the observation space for computational efficiency. We select sample characteristic nodes according to the measurement capabilities of the Orbilander life detection instrument suite. Table I outlines the full list of relevant sample characteristics and direct biosignatures.

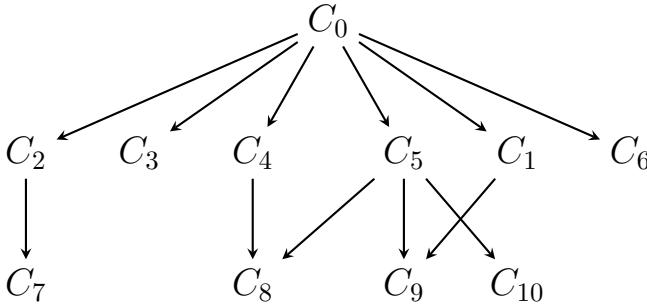


Fig. 3. Selected Bayesian network structure. Though many more characteristics and dependencies may exist, we present a simplified model for clarity. Descriptions of all nodes can be found in Table I.

These sample characteristics are formed into a Bayesian network in Figure 3, outlining how the existence of life C_0 directly informs the presence of various biosignatures in a sample. Some nodes, C_1 through C_5 , represent strong direct biosignatures, such as the presence of polyelectrolytes or cells, while others reflect weaker, more ambiguous indicators of life. Biosignature signals should arise only if life is present, which we depict as a causal, conditional relationship in our Bayesian network. For example, measured pH C_9 depends on amino acid abundance C_5 and polyelectrolytes C_1 , which in turn depend on whether life is present. This leads to several tiers of connections in the Bayesian network structure, building a causal graph with life as the root node that affects all biosignature children nodes. As discussed in Section III-A, using the Bayesian network structure leverages conditional

dependencies that minimize model parameter count and complexity while preserving all essential relationships.

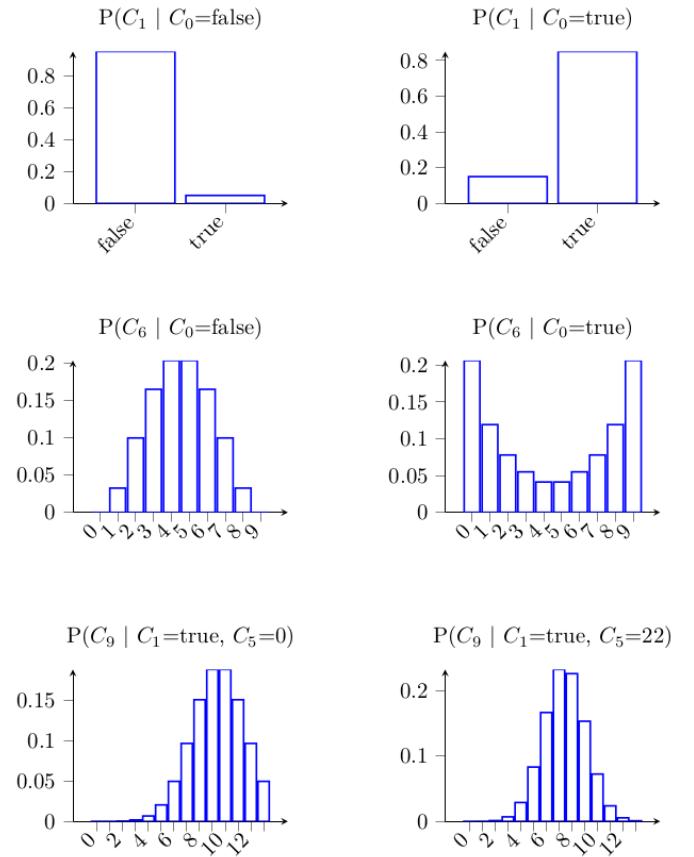


Fig. 4. Selected conditional probability distributions, each corresponding to a node and its set of parents within the Bayesian network. Direct biosignatures are functions of life and habitability metrics are functions of biosignatures, according to the causal graph convention. For a complete visualization of all distributions, see Appendix.

For each biosignature listed in Table I, we define conditional probability distributions (CPDs) representing the likelihood of observing a biosignature when life is present or absent in the sample. These CPDs are shown in Figure 4 and Figure 9. Each distribution is designed using Figure 3 and insights from existing astrobiology literature, though expert opinion may vary [4], [22], [23]. Due to the lack of empirical data of biosignatures from icy moons like Enceladus, the Bayesian Network's CPDs must be manually crafted. We follow the strategies for addressing the challenges of biosignature CPD design from Marshall et al. [23]. As future missions collect more observational data from deep space environments, both structure and CPD parameter learning can be performed to design Bayesian networks that better represent the statistical patterns of biosignature evidence.

In general, an effective biosignature is one for which the probability of detecting certain observations is significantly higher when life is present than with strictly abiotic processes; that is, $P(\text{observation} | \text{life}) \gg P(\text{observation} | \text{abiotic})$. In addition, a biotic signal must be strong enough to surpass the detection threshold of each instrument. For the purposes of

this analysis, we adopt an optimistic prior—serving as our null hypothesis— $P(\text{life}) < 10^{-5}$. To enable exact Bayesian inference and reduce the dimensionality of the observation space, we discretize the CPDs using histogram binning. For continuous variables, approximate inference techniques would be required. In the next section, we describe how this Bayesian network can represent POMDP observations.

C. Life Detection Suite POMDP Formulation

The scope of this POMDP models instrument operations for the Enceladus Orbilander life detection suite (LDS), a subsystem within the Orbilander spacecraft. This focus minimizes complexity within the state, action, and observation space.

To model the LDS as a POMDP, we introduce state vector s . The first element, s_V , denotes the volume of sample available for LDS subsystem analysis. Samples are passively collected in a funnel during the spacecraft's orbital phase [21]. This value is fully observable, meaning that observations $o_V = s_V$. The second state element, s_L , represents the sample's true life state, i.e. whether the sample is biotic or abiotic. s_L is only partially observable via instrument measurements, which produce noisy observations o_L of the characteristics outlined in Table I. In a fully independent observation model, where all m characteristics can assume n values, the observation space would scale very poorly ($\mathcal{O}(n^m)$). However, the Bayesian network defined in Figure 3 encodes dependencies between observations into a factored form. This improves scaling while preserving measurement interpretability and causality.

The POMDP allows the following actions to be taken after observing state s at each timestep. These are

$$\mathcal{A} = \begin{cases} a_1, \dots, a_6 & \text{Use instrument } a_i \quad \forall i \in \{1, \dots, 6\} \\ a_7 & \text{Accumulate sample volume (idle)} \\ a_8 & \text{Declare abiotic} \\ a_9 & \text{Declare biotic} \end{cases} \quad (3)$$

where instrument actions $a_1 \dots a_6$ can only be performed if enough sample volume is available for that instrument. For example, the Orbilander's nanopore requires far more sample volume than other instruments like the electrochemical sensor array. Each instrument action $a_1 \dots a_6$ produces an observation o of the current state s . Each instrument returns measurements listed in Table II simultaneously. For example, HRMS measures four sample characteristics C_5, C_7, C_8, C_{10} at once.

TABLE II
ORBILANDER LIFE DETECTION INSTRUMENTS

| Action | Instrument Name | Measurements |
|--------|--|-------------------------|
| a_1 | High-Resolution Mass Spectrometer (HRMS) | C_5, C_7, C_8, C_{10} |
| a_2 | Separation Mass Spectrometer (SMS) | C_5, C_6 |
| a_3 | Microfluidics Device (μ CE-LIF) | C_5, C_6 |
| a_4 | Electrochemical Sensor Array (ESA) | C_7, C_8 |
| a_5 | Microscope | C_2, C_3 |
| a_6 | Nanopore | C_1 |

The last two actions, a_8 and a_9 , are terminal actions. a_8 is used to declare a sample abiotic if insufficient signs of life are

observed, while a_9 is used to declare that life has been found in a particular sample.

Actions were designed in this way to match constraints of the Orbilander sample preparation chamber, which only prepares one sample at a time. This batching behavior implies that with each new sample collection, the underlying life state of the sample, s_L , may change. The effects of each action a_i on the sample life state s_L and volume s_V are captured by the POMDP transition function $T(s, a)$. As previously stated, given a current state-action pair, $T(s, a)$ defines a probability distribution over possible next states s' . The following equation describes the transition logic.

$$T(s, a) = \begin{cases} s'_L \sim P(\text{life}) & \text{if } a = a_7 \\ s'_V = s_V + v_{acc} & \text{if } a = a_7 \\ s'_L = s_L & \text{if } a \in \{a_1, \dots, a_6\} \\ s'_V = s_V - v_{use}(a) & \text{if } a \in \{a_1, \dots, a_6\} \\ s'_L, s'_V = \text{End State} & \text{if } a \in \{a_8, a_9\} \end{cases} \quad (4)$$

Three main types of transitions occur with sample pairs (s, a) : sample accumulation a_7 , instrument use a_1, \dots, a_6 , and declaration of the sample's biotic state a_8, a_9 . For sample accumulation, the spacecraft accumulates more sample s_V based on the Enceladus surface plume fallout rates v_{acc} , a time-varying signal. Every accumulation period has a small likelihood $P(\text{life})$ of altering the sample's life state s_L . Instrument usage actions do not affect the sample's biotic state, and only the sample volume s_V gets reduced according to the amount of sample consumed by instruments a_1, \dots, a_6 . The terminal state of the POMDP is reached when an action is taken to declare the current sample as abiotic a_8 , or biotic a_9 .

The reward function is designed to capture several competing objectives within instrument operations. In the context of life detection, a central tradeoff arises between reducing sample uncertainty and maximizing sensitivity to biosignatures: the former may lead to false negatives, and the latter to false positives. To balance the consequences of each, we introduce a design parameter λ to the reward function, which scales the cost ratio between false conclusions and sufficient information gathering. Similarly, τ is a risk aversion parameter to penalize false negatives. The full reward equation is

$$R(s, a) = \begin{cases} 0 & \text{Correctly declare biotic} \\ -\tau \cdot \lambda & \text{Correctly declare abiotic} \\ -\lambda & \text{Incorrect declaration} \\ (1 - \lambda) \frac{s_V}{s_V^{max}} & \text{Running instrument} \\ -\infty & \text{Infeasible actions} \end{cases} \quad (5)$$

where s_V^{max} represents the maximum sample volume capacity in the LDS sample preparation chamber.

To encourage the agent to search for life for as long as possible, no penalty is associated with declaring life in a sample (the ideal outcome). All other actions have associated penalties, even when the agent correctly declares samples abiotic, as donated with τ . This reflects the cost associated with a potential false negative. False conclusions about a sample's biology incur a heavier penalty following parameter

λ to motivate the agent to make accurate declarations. Infeasible actions—running an instrument with insufficient sample volume, for example—are given a near-infinite penalty. Valid instrument runs are penalized according to how much sample volume has been consumed, scaled by $(1 - \lambda)$. By tuning $\lambda \in [0, 1]$, preferences can be set on which objective the agent should focus on. If the agent should prioritize reducing uncertainty about the underlying sample’s life state, then λ should be large and approach 1. On the other hand, if the agent should react quickly to the samples it has access to (near end-of-life, for example), the λ parameter should instead approach 0. We provide experiments in the following section to determine the best λ factor to balance both objectives appropriately.

We further outline additional constants in the Life Detection Suite POMDP in Table III.

TABLE III
STATIC PARAMETERS FOR LIFE DETECTION POMDP

| Parameter | Variable | Value |
|--------------------------|----------------|-------|
| Number of instruments | n | 6 |
| Sample chamber capacity | s_V^{max} | 100% |
| Sample fallout rate | v_{acc} | 10 |
| Instrument sample usage: | v_{use} | |
| HRMS | $v_{use}(a_1)$ | 1% |
| SMS | $v_{use}(a_2)$ | 6% |
| μ CE-LIF | $v_{use}(a_3)$ | 2% |
| ESA | $v_{use}(a_4)$ | 3% |
| Microscope | $v_{use}(a_5)$ | 1% |
| Nanopore | $v_{use}(a_6)$ | 89% |

D. Concept of Operations Algorithm

The baseline we compare to our methodology is the Concept of Operations outlined in the Enceladus Orbilander’s mission planning document [21]. We provide the full ConOps in Algorithm 1. The ConOps employs a structured, cyclic, multi-phase algorithm to systematically assess the presence of life. In each cycle, the algorithm accumulates samples up to a specified maximum volume needed to run all instruments. Once sufficient sample volume is collected, the algorithm sequentially activates each scientific instrument to analyze the sample, recording the resulting observations. After each observation, the algorithm updates the belief in life based on the observations derived from the instrument measurements. The decision phase then compares the updated belief against the predefined thresholds of when to declare life. If the belief exceeds the life detection threshold, life is declared detected; if it falls below the dead threshold, the absence of life is declared. This threshold-driven approach, while systematic, imposes a very rigid structure on the life detection process, potentially limiting flexibility in interpreting ambiguous or borderline results.

V. POLICY PERFORMANCE RESULTS

We now evaluate the performance of our precomputed autonomous decision-making policies for the LDS POMDP against the performance of the existing Enceladus Orbilander

Algorithm 1 Enceladus Orbilander Concept of Operations for Life Detection

```

1: ORBILANDERLIFEDETECTION
2: Initialize  $b_{life}$                                  $\triangleright$  Belief in life
3: Initialize  $T_{life}$                                  $\triangleright$  Threshold of belief indicating life
4: Initialize  $T_{dead}$                                  $\triangleright$  Threshold of belief indicating dead
5: Initialize sample volume:  $s_V \leftarrow 0$ 
6: Initialize instrument status: all unused
7: while mission not complete do           $\triangleright$  Repeat for entire
   surface mission
8:   while  $T_{dead} < b_{life} < T_{life}$  do
9:     // Phase 1: Accumulate sample
10:    while  $s_V < s_V^{max}$  do
11:      Accumulate sample  $s_V \leftarrow s_V + v_{acc}$ 
12:    // Phase 2: Run instruments
13:    for each instrument  $a_i$  in  $a_1, \dots, a_6$  do
14:      Run instrument  $a_i$  and receive observation  $o_i$ 
15:       $s_V \leftarrow s_V - v_{use}(a_i)$ 
16:    // Phase 3: Update belief
17:    Update  $b_{life} \leftarrow P(\text{life} | o_1)$ 
18:    // Phase 4: Decision
19:    if  $b_{life} \geq T_{life}$  then
20:      Declare Life Detected
21:    else if  $b_{life} \leq b_{dead}$  then
22:      Declare No Life Detected

```

ConOps. All policies were generated using the offline POMDP solver SARSOP. To select an effective policy, we first conducted a parameter sweep over λ , which governs the tradeoff between minimizing uncertainty and maximizing sensitivity to biosignatures. Using the policy corresponding to the optimal λ in our Pareto frontier, we analyze the resulting alpha vector plot and illustrate a decision tree from a single rollout. We then compare the classification performance of the SARSOP-derived policy to that of the current Orbilander concept of operations [21].

A. Selecting the Best Design Tradeoff

We evaluate a range of λ and τ values between $[0, 1]$, comparing the resulting policies by two metrics: uncertainty in final agent declaration and average sensor usage.

The former is the final uncertainty margin—the agent’s belief gap between its chosen declaration (biotic or abiotic) and the alternative hypothesis at the final timestep. This metric ranges from 0 to 1, where a value of 0 indicates the agent was fully confident in its final declaration, and values closer to 1 indicate greater uncertainty. This measure captures how decisively the agent reaches conclusions based on the available evidence, regardless of whether the decision is ultimately correct.

The second metric is the average cumulative sensor usage per rollout. Longer rollouts imply more sensing, allowing the agent to gather data and build stronger beliefs about each sample’s biology. However, excessive accumulation and testing can have diminishing returns. In the context of the LDS system, each new batch introduces additional sample material,

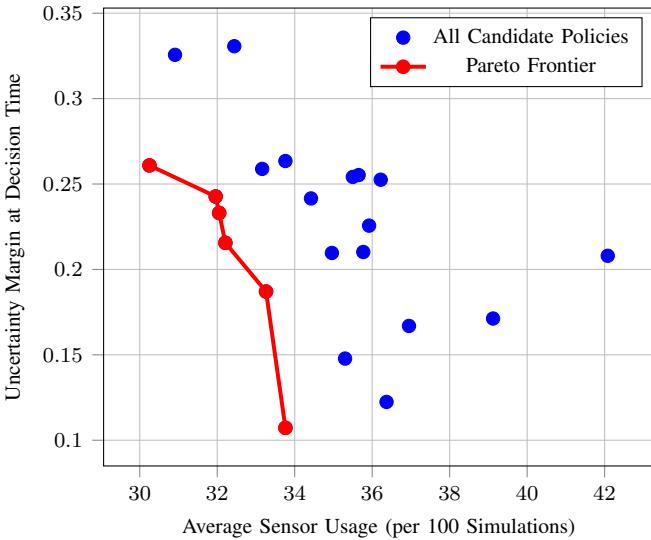


Fig. 5. Pareto frontier indicating optimal tradeoff between sensor usage and sample analysis accuracy.

potentially diluting or contaminating previous biosignatures. As a result, strong biotic signals may become obscured by an increasing concentration of abiotic material, making it harder for the agent to confidently detect and declare life when it is actually present.

For every policy, 100 rollouts were performed to assess uncertainty margins and average sensor usage. Metrics for each policy are plotted in Figure 5 to compare performance, and a Pareto frontier of most optimal solutions is identified in red. Although multitudes of λ values were tested during this process, most values of λ less than 0.8 had unrealistic policies that did not provide useful behavior, with premature halting of testing. Generally, smaller λ values corresponded to very risk-tolerant policies, with a willingness to declare biotic or abiotic signs in a sample with high certainty. A linear increase in λ did not produce a similar linear correlation in either of the two metrics, and only certain regions of $\lambda \in [0, 1]$ created comprehensible policies. For policies depicted in Figure 5, all values of $\lambda \in [0.85, 0.9999]$. Higher values of λ corresponded to very risk-avoidant policies, which became overly conservative with declaring samples as biotic/abiotic once as $\lambda \rightarrow 1$. We find that the λ values leading to the best depicted policies on the Pareto frontier were $[0.85295, 0.996, 0.998, 0.999, 0.9995, 0.9998]$. We identified an optimal balance with $\lambda = 0.85295$, which had the lowest uncertainty margin among all tested policies while maintaining reasonable average sensor usage. All subsequent experiments are performed with this policy.

B. Alpha Vector Policy

To better interpret the $\lambda = 0.85295$ policy generated with SARSOP, we examined the produced set of alpha vectors, which encode the expected utility of each action under each possible state. As each alpha vector is linked with a specific action, the dominating alpha vector returns the most optimal

action to take under a given belief over s_L . Examples of dominating alpha vectors are shown in Figure 6.

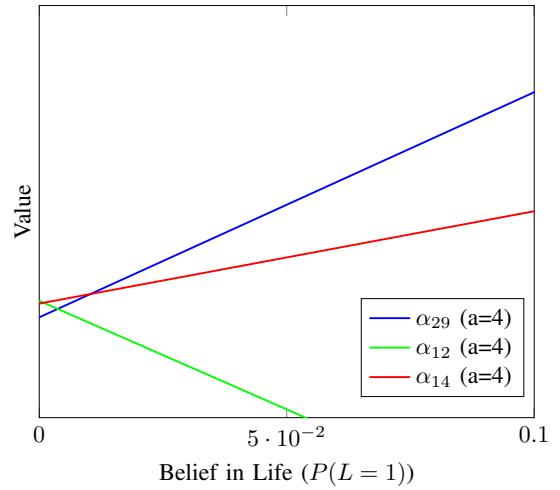


Fig. 6. Example set of alpha vectors over belief in s_L at a sample volume of 46 units.

Due to the multi-dimensional state space of our problem, a set of alpha vectors, similar to Figure 6, exists for every possible sample volume s_V . To visualize this, in Figure 7, we plot the dominant alpha vector per sample volume and current belief over s_L , revealing a complicated set of action sequences depending on how much sample is accumulated by the Orbilander LDS system. This plot can be used not only as a look-up table, but as a validation tool to check the policy's recommendations of possible belief states to be used in any situation. In Figure 7, we found that the policy alternates between using instrument actions a_2, a_3, a_4 and a_5 , and only declares samples abiotic a_8 or biotic a_9 once the agent's belief in life passes a certain range. The range for when the agent declares samples abiotic a_9 , is much larger than when the agent declares samples abiotic a_8 , indicating that the policy attempts to find more opportunities to detect life, rather than rule it out. This behavior aligns with the reward function outlined in Section IV-C. The policy also has a clear preference for using instrument action a_4 , the ESA, for evaluating samples. Other instrument actions, a_2, a_3 , and a_5 are used only when the agent needs to verify that life is abiotic, which brings in the use of additional sensors for verification. Samples are only accumulated using a_7 when the sample volume s_V is lower than 20%.

C. Decision Tree

To interpret spacecraft behavior over time, we depict simulations using a decision tree as shown in Figure 8. Red arrows indicate negative or absent observations of a Boolean variable (e.g. polyelectrolytes detected or not detected), while green arrows show a biosignature was indeed measured. Black arrows denote numerical, real-valued measurements. The belief in the probability of life in the analyzed samples $P(\text{life})$ is tracked at each time step. Following the tree structure, it is clear that certain observations guide the agent toward increasingly confident beliefs about $P(\text{life})$, while others are inconclusive.

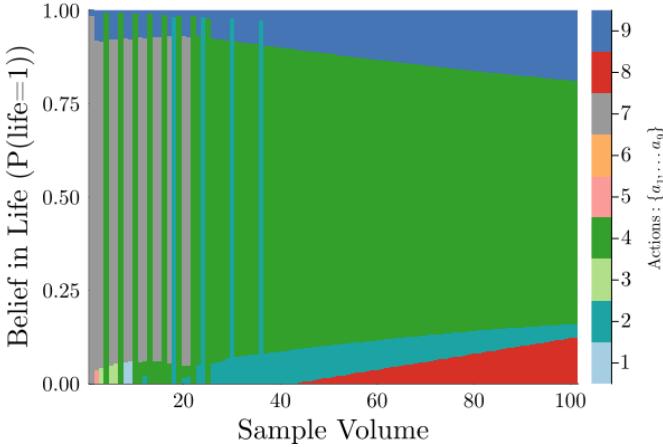


Fig. 7. Dominating alpha vector at each belief state and sample volume. Associated actions are given in the legend for the $\lambda = 0, \tau = 0, \gamma = 0$ policy from the offline solver, SARSOP.

The tree reveals early-stage exploration dominated by sensor actions, later followed by biotic or abiotic declarations as confidence is built.

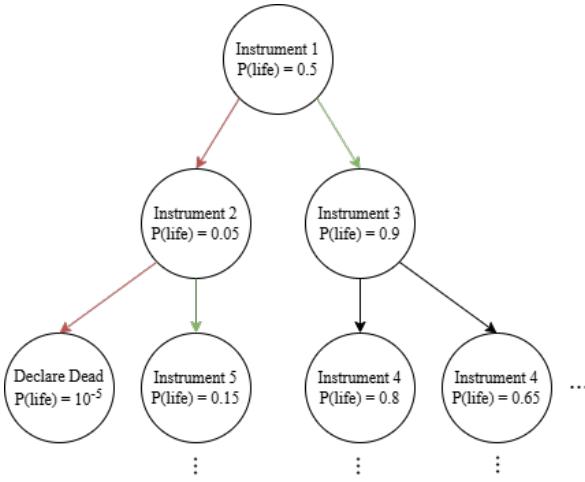


Fig. 8. A decision tree visualizing possible rollouts of the SARSOP optimal policy. Edges correspond to observations, and nodes show both optimal actions and updated beliefs over the prior.

D. Binary Classifier Performance

Though accuracy in predicting the presence of life is only one performance measure (others being total data gathered, total value of information, measurement efficiency, etc.), we compare each policy's false positive, true positive, false negative, and true negative rates in the confusion matrices shown in Table IV. Overall, we show improved performance over the Orbilander ConOps in simulation in all categories, especially reducing false negatives due to the policy balancing both objectives of minimizing declaration uncertainty and quickly reacting to measurements.

TABLE IV
CONFUSION MATRICES OF POLICY ACCURACY

| | | (a) SARSOP Policy (Ours) | |
|--------|---------|--------------------------|---------|
| | | Predicted | |
| | | Biotic | Abiotic |
| Actual | Biotic | 100 | 0 |
| | Abiotic | 9 | 91 |

| | | (b) Orbilander ConOps Policy | |
|--------|---------|------------------------------|---------|
| | | Predicted | |
| | | Biotic | Abiotic |
| Actual | Biotic | 90 | 10 |
| | Abiotic | 7 | 93 |

VI. CONCLUSION

This work presents a solution for performing autonomous adaptive science operations for deep space missions, using Enceladus life detection instrument operations as a case study. We propose a novel integration of a Bayesian network to represent the observation space of biosignatures captured by science instruments, combined with a POMDP formulation to guide autonomous decision-making under uncertainty using verifiable precomputed policies.

Although our modeling is grounded in the Enceladus Orbilander mission, the framework generalizes to a wide range of science operations in resource-constrained, uncertain environments. The modular architecture supports adaptation across missions by simply updating the Bayesian network structure or tuning environmental and operational parameters in the POMDP. In doing so, we lay the foundation for autonomous systems that are not only efficient and responsive but also thoroughly verifiable.

In future work, we seek to show improved performance, science collection, and minimized risk with SARSOP during off-nominal events, such as fluctuating sample volume (Enceladus plume ejecta rates are highly uncertain) as well as degraded or broken life detection instruments. Another future goal is to integrate the station keeping subsystem into the autonomy stack, which directly influences plume volume accumulated and passed to the life detection instruments. This may also involve incorporating more engineering constraints into the decision-making framework, like power and memory budgets. By expanding our methods to handle degraded sensing, related subsystems, and dynamic environments, we move closer to building resilient, autonomous science platforms capable of real-time decision-making at the edge of the solar system.

ACKNOWLEDGMENT

This work was supported by the NASA Ames Center Innovation Fund (CIF) under the grant “SHERPA: Robust Pre-computed Autonomy (RPA) Module,” as well as the Stanford Graduate Fellowship (SGF) program.

APPENDIX

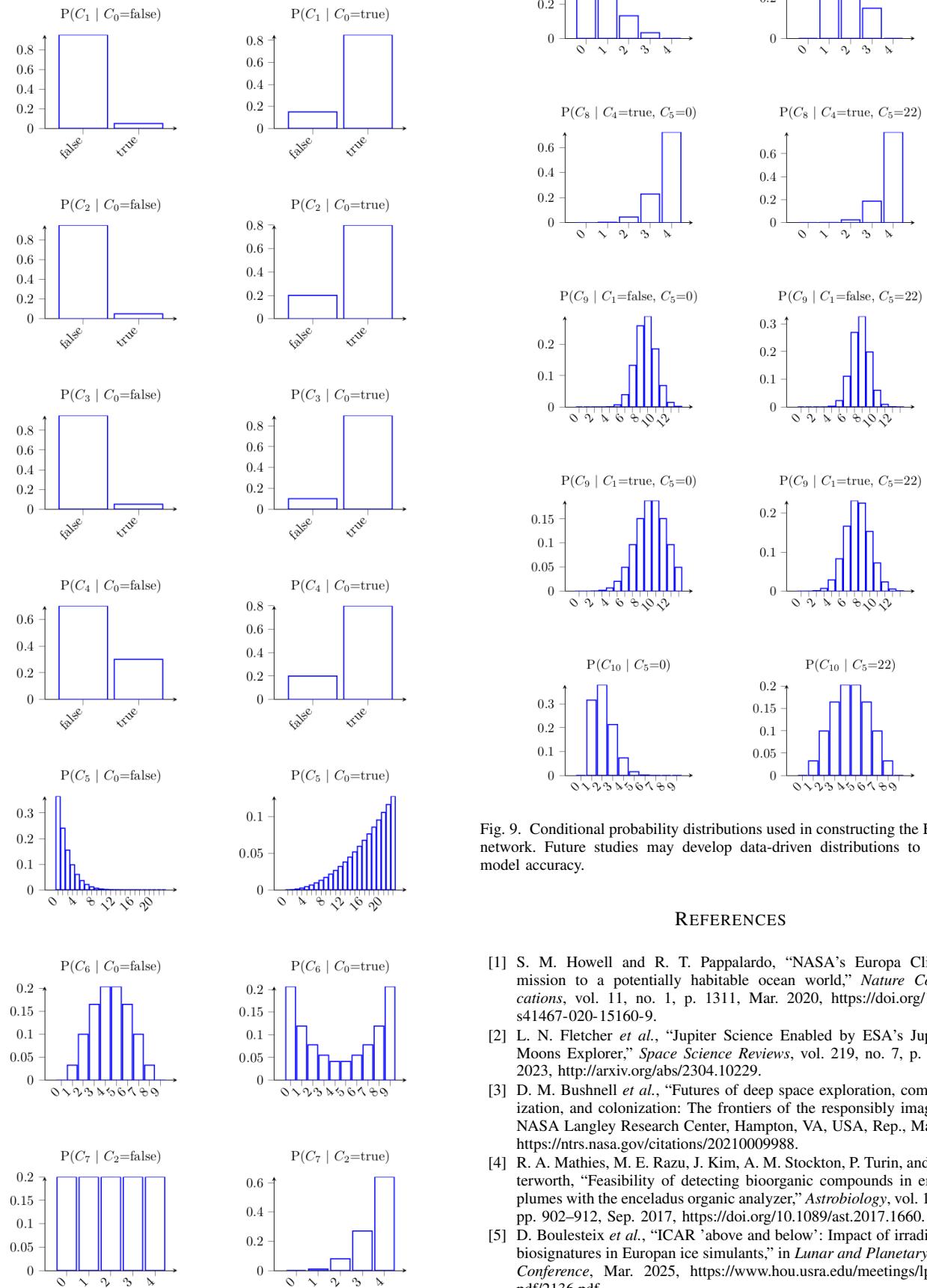


Fig. 9. Conditional probability distributions used in constructing the Bayesian network. Future studies may develop data-driven distributions to improve model accuracy.

REFERENCES

- [1] S. M. Howell and R. T. Pappalardo, "NASA's Europa Clipper—a mission to a potentially habitable ocean world," *Nature Communications*, vol. 11, no. 1, p. 1311, Mar. 2020, <https://doi.org/10.1038/s41467-020-15160-9>.
- [2] L. N. Fletcher *et al.*, "Jupiter Science Enabled by ESA's Jupiter Icy Moons Explorer," *Space Science Reviews*, vol. 219, no. 7, p. 53, Oct. 2023, <https://arxiv.org/abs/2304.10229>.
- [3] D. M. Bushnell *et al.*, "Futures of deep space exploration, commercialization, and colonization: The frontiers of the responsibly imaginable," NASA Langley Research Center, Hampton, VA, USA, Rep., Mar. 2021, <https://ntrs.nasa.gov/citations/20210009988>.
- [4] R. A. Mathies, M. E. Razu, J. Kim, A. M. Stockton, P. Turin, and A. Butterworth, "Feasibility of detecting biogenic organic compounds in Enceladus plumes with the Enceladus Organic Analyzer," *Astrobiology*, vol. 17, no. 9, pp. 902–912, Sep. 2017, <https://doi.org/10.1089/ast.2017.1660>.
- [5] D. Boulesteix *et al.*, "ICAR 'above and below': Impact of irradiation on biosignatures in Europa ice simulants," in *Lunar and Planetary Science Conference*, Mar. 2025, <https://www.hou.usra.edu/meetings/lpsc2025/pdf/2136.pdf>.

- [6] A. Banerjee, M. Mukherjee, S. Satpute, and G. Nikolakopoulos, “Resiliency in space autonomy: A review,” *Current Robotics Reports*, vol. 4, no. 1, pp. 1–12, Mar. 2023, <https://doi.org/10.1007/s43154-023-00097-w>.
- [7] V. Verma *et al.*, “Autonomous robotics is driving Perseverance rover’s progress on mars,” *Science Robotics*, vol. 8, no. 80, p. eadi3099, 2023, <https://www.science.org/doi/abs/10.1126/scirobotics.adi3099>.
- [8] K. Echigo *et al.*, “Autonomy in the real-world: Autonomous trajectory planning for asteroid reconnaissance via stochastic optimization,” arXiv, Dec. 2024, <http://arxiv.org/abs/2412.06816>.
- [9] D. Bernard *et al.*, “Design of the remote agent experiment for spacecraft autonomy,” in *IEEE Aerospace Conference*, vol. 2, Mar. 1998, pp. 259–281, <https://ieeexplore.ieee.org/document/687914>.
- [10] B. P. Theiling *et al.*, “Science autonomy for ocean worlds astrobiology: A perspective,” *Astrobiology*, vol. 22, no. 8, pp. 901–913, Aug. 2022, <https://pubmed.ncbi.nlm.nih.gov/35507950/>.
- [11] J. M. Srinivasan *et al.*, “Europa clipper flight system overview,” *Space Science Reviews*, vol. 221, no. 1, p. 14, Feb. 2025, <https://doi.org/10.1007/s11214-025-01139-9>.
- [12] C. Wagner *et al.*, “Demonstrating autonomy for complex space missions: A europa lander mission autonomy prototype,” *Journal of Aerospace Information Systems*, vol. 21, no. 1, pp. 37–57, Jan. 2024, <https://arc.aiaa.org/doi/10.2514/1.I011294>.
- [13] H. Zerrouki, H. D. Estrada-Lugo, H. Smadi, and E. Patelli, “Applications of bayesian networks in chemical and process industries: A review,” in *European Safety and Reliability Conference*, Sep. 2019, pp. 3122–3129, <http://rpsonline.com.sg/proceedings/9789811127243/html/0914.xml>.
- [14] T. Beuzen *et al.*, “Bayesian networks in coastal engineering: Distinguishing descriptive and predictive applications,” *Coastal Engineering*, vol. 135, pp. 16–30, May 2018, <https://www.sciencedirect.com/science/article/pii/S0378383917303678>.
- [15] A. Harris, T. Valade, T. Teil, and H. Schaub, “Generation of spacecraft operations procedures using deep reinforcement learning,” *Journal of Spacecraft and Rockets*, vol. 59, no. 2, pp. 611–626, Mar. 2022, <https://doi.org/10.2514/1.A35169>.
- [16] W. Kuhl, J. Wang, D. Eddy, and M. J. Kochenderfer, “Markov decision processes for satellite maneuver planning and collision avoidance,” arXiv, Jan. 2025, <https://arxiv.org/abs/2501.02667>.
- [17] A. Herrmann and H. Schaub, “Autonomous small body science operations using reinforcement learning,” *Journal of Aerospace Information Systems*, vol. 21, no. 10, Oct. 2024, <https://doi.org/10.2514/1.I011376>.
- [18] M. J. Kochenderfer, T. A. Wheeler, and K. H. Wray, *Algorithms for Decision Making*. Cambridge, MA, USA: MIT Press, 2022, <https://mitpress.mit.edu/9780262047012/algorithms-for-decision-making/>.
- [19] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*. Cambridge, MA, USA: The MIT Press, 2009, <https://mitpress.mit.edu/9780262013192/probabilistic-graphical-models/>.
- [20] H. Kurniawati, D. Hsu, and W. S. Lee, “(sarsop): Efficient point-based pomdp planning by approximating optimally reachable belief spaces,” in *Robotics: Science and Systems*, Jun. 2008, pp. 65–72, <https://ieeexplore.ieee.org/document/6284837>.
- [21] S. M. MacKenzie *et al.*, “Enceladus Orbilander: A flagship mission concept for astrobiology,” NASA Goddard Space Flight Center, Greenbelt, MD, USA, Rep. 20205008712, Oct. 2020, <https://ntrs.nasa.gov/citations/20205008712>.
- [22] E. D. Dorn, K. H. Nealson, and C. Adami, “Monomer abundance distribution patterns as a universal biosignature: Examples from terrestrial and digital life,” *Journal of Molecular Evolution*, vol. 72, no. 3, pp. 283–295, Mar. 2011, <http://link.springer.com/10.1007/s00239-011-9429-4>.
- [23] S. M. Marshall, A. R. G. Murray, and L. Cronin, “A probabilistic framework for identifying biosignatures using pathway complexity,” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 375, no. 2109, p. 20160342, Nov. 2017, <https://royalsocietypublishing.org/doi/10.1098/rsta.2016.0342>.