2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
Macau, China, November 4-8, 2019

# Non-myopic Planetary Exploration Combining In Situ and Remote Measurements

Suhit Kodgule[1], Alberto Candela[1] and David Wettergreen[1]

***Abstract*— Remote sensing measurements can provide crucial information about the material properties of a planetary surface but their application is limited by their spatial resolution, typically tens of meters per pixel, when constituent materials are mixed at much finer scale. Consequently the orbital observations must be validated with *in situ* measurements from a spectrometer on the ground. In planetary exploration this means that a rover must visit selected locations that jointly improve a model of the environment and satisfy mobility and sampling constraints. Conventional planning methods used in this situation follow sub-optimal greedy strategies that are not scalable to large areas. We show how the problem can be effectively defined in a Markov Decision Process framework and propose a planning algorithm based on Monte Carlo Tree Search, which is efficient but devoid of these drawbacks thereby providing superior performance. We evaluate our approach using hyperspectral imagery of a well-studied geologic site in Cuprite, Nevada.**

## I. INTRODUCTION

An emerging problem in space exploration is how to increase the autonomy of planetary rovers that are utilized for exploration and survey. Figure 1 displays a prototypical rover used for conducting field experiments in science autonomy. These robots must be capable of effectively incorporating scientific data into their planning strategies for satisfying mission objectives. Typical objectives include measuring discrete features of interest [1] or planning paths using experimental design methods to maximize scientific information gain [2].

Remote sensing measurements play an integral role in not only meeting these objectives, but also prescribing new objectives for science-aware exploration. For instance, spectroscopic analysis [3] [4], which is reliant on orbital sensor measurements, is used extensively for understanding mineralogical composition on planetary surfaces. Ensuring that these measurements are accurate and provide an unambiguous interpretation of the surface is critical for the success of the mission. However, orbital imaging spectrometers have limited resolution and as a result, features that are concentrated on a small region can go undetected. *In situ* spectral measurements can aid in detecting these exceptions as they possess a higher resolution and also provide a higher signal to noise ratio.

This work focuses on the *spatio-spectral* exploration (SSE) problem [5]. In this scenario, a field robot gathers a library of *in situ* spectral measurements that facilitates *spectral unmixing* of the remote image. Concretely, it attempts to collect pure spectral signatures that could reconstruct the orbital image. A detailed explanation of SSE is provided in Section II.

Fig. 1. Prototype planetary rover in the Atacama Desert of Chile, exploring terrain to classify and map geology. An on-board spectrometer measures individual rocks to identify mineralogical composition.

In this scenario, a path planning algorithm must collect a set of most informative samples while satisfying a sampling constraint. This task bears similarity to the Informative Path Planning (IPP) problem which requires an efficient trade-off between exploration and exploitation. A number of active learning strategies have been employed for IPP scenarios such as search and rescue [6], sensor placement [7] or autonomous underwater navigation [8]. All of these approaches however, rely on mapping a scalar field representing the environment, which is commonly done using a Gaussian Process (GP). On the other hand, in our scenario the objective function is dependent on highly correlated distributions of minerals which cannot be mapped to a scalar field. Moreover GP models are unsuitable for solving problems with non-stationary environments. We explain how SSE involves navigating a non-stationary environment in Section III. A number of methods have been employed for science-aware navigation. Thompson *et al.* developed a GP model using remote sensing observations to predict the content of basalt on the surface [9]. Smith designed a POMDP based planner for detecting life in the Atacama desert [10]. Woods *et al.* designed an autonomous, opportunistic science system which included a replanning strategy for arranging a series of science activities [11]. Approaches employed for SSE however, continue to utilize naive strategies such as an exhaustive

[1]The Robotics Institute, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA {skodgule, albertoc, dw0s}@andrew.cmu.edu

coverage of the region or the Greedy Spectral Selection (GSS) algorithm [5]: a greedy strategy that attempts to maximize the entropy of the collected samples.

Reinforcement Learning (RL) provides a suitable framework for developing non-myopic plans while trading off exploration against exploitation. These properties make RL well suited for generating plans for SSE. Several Dynamic Programming (DP) approaches such as Value Iteration [12] and Policy Iteration [13] provide provable convergence guarantee. Nevertheless, they require complete knowledge of states and transitions of the environment, which cannot be easily determined in real-world scenarios. In such situations, online Monte-Carlo based approaches prove useful. Morere *et al.* use POMDPs with Gaussian Process based reward functions for monitoring spatial phenomena with a UAV [14]. Clary *et al.* use a Monte Carlo Tree Search (MCTS) with Upper Confidence Bounds for Trees (UCT) for legged locomotion [15]. Ross *et al.* propose a no-regret algorithm for learning non-stationary policies using Imitation Learning [16]. However, acquiring relevant demonstrations for robotic exploration is a non-trivial problem.

In this work, we focus on adapting MCTS for SSE. While adapting MCTS, it is important to note that there is inherent uncertainty present in the environment as we do not know the *in situ* spectral information for each state of the environment *a priori*. In Section III we demonstrate how utilizing orbital sensing information itself can account for this uncertainty. This paper makes the following contributions:

- We propose an MDP framework in non-stationary environments for SSE that uses orbital sensing measurements to account for uncertainty in the environment.
- We present a path planning algorithm based on MCTS to collect the most informative samples for spectral unmixing.
- We provide a detailed comparative analysis of our proposed method against current practices in SSE on a spectroscopic dataset obtained from the canonical mineral site in Cuprite, Nevada.

The paper is organized as follows. In Section II we briefly summarize MDP, SSE and Spectral Mixture Models. In Section III, we explain our approach based on Monte Carlo planning that can be utilized to gain informative spectral samples. In Section IV, we describe the spectroscopic dataset utilized and the planning strategies that we compare against. Finally we provide our results and concluding thoughts in Sections V and VI respectively.

## II. BACKGROUND

We begin by defining an MDP and then provide brief descriptions of spectral mixture models and SSE.

### A. Markov Decision Process

An MDP can be represented by a tuple $M = (S, A, T, R)$.

- $S$ is a finite set of states.
- $A$ is a finite set of actions.
- $T$ is a probabilistic transition function $T(s, a, s') = p(s'|s, a)$ which defines the probability of transitioning into state $s'$ after taking action $a$ from state $s$.
- $R : S \times A \to \mathbb{R}$ is the reward function.

MDPs rely on the Markov assumption which states the distribution on future states is only dependent on the current state. A *policy* $\pi : S \to A$ is a function that maps a state $s \in S$ to an action $a \in A$. An optimal policy $\pi^*$ is defined as the one that maximizes the cumulative discounted rewards. Formally,

$$\pi^* = \underset{\pi}{\operatorname{argmax}}\, E\left[\sum_{t=0}^{\infty} \gamma^t r_t^{\pi}\right] \tag{1}$$

Here, $\gamma$ is the discount factor and $r_t^{\pi}$ is the reward acquired at time $t$ by following the policy $\pi$. A detailed explanation of MDPs can be found in [17].

### B. Spectral Mixture Models

Orbital hyperspectral sensors have poor spatial resolution on account of measuring from high altitudes. As a consequence, these sensors record scenes in which numerous distinct minerals contribute to the same spectrum that is measured from a single pixel of the hyperspectral image. This gives rise to mixed pixels, which is formed by a combination of pure minerals called *endmembers*. The fractions in which these endmembers combine to form the mixed pixel are known as abundances. We consider here a linear mixing model (LMM) which assumes that the mixed pixel's spectrum is formed by a linear combination of the endmembers that constitute the pixel. Formally, if a spectrum has $d$ bands or channels, we can define a pixel of an orbital image by a vector $x \in \mathbb{R}^{\mathrm{d}}$. Further, if there are $K$ endmembers observed in the entire orbital image, then each pixel $x_j$ of an orbital image $X$ can be expressed as:

$$x_j = \sum_{i=1}^{K} a_{ij} y_i \quad \forall x_j \in X \tag{2}$$

where $y_i \in \mathbb{R}^{\mathrm{d}}$ is the $i_{th}$ endmember spectrum and $a_{ij}$ is the abundance scalar for $i_{th}$ endmember and $j_{th}$ pixel. For physical realizability, endmembers must not have negative abundances, that is: $a_{ij} \geq 0,\ \forall i, j$ and $\sum_{i=1}^{K} a_{ij} = 1,\ \forall j$. Equation (2) can be written as $x_j = Y a_j,\ \forall x_j \in X$, where $Y$ is a library of $K$ endmember spectra: $Y = [y_1,\ y_2,\ \ldots, y_K]$.

An unmixing problem attempts to find the abundance vector $a_j = [a_{1j},\ a_{2j},\ \ldots\ a_{Kj}]^{\top}$ for each pixel $x_j$ of $X$ given $Y$. Typically, this is solved using non-negative least squares optimization [18] and involves solving the following objective for each pixel:

$$\begin{aligned} &\underset{a_j}{\operatorname{argmin}} \quad \|Y a_j - x_j\|_2 \\ &s.t.\ a_{ij} \geq 0 \quad \forall a_{ij} \in a_j \end{aligned} \tag{3}$$

### C. Spatio-Spectral Exploration

SSE involves collecting a library of endmembers $Y$ that can solve an unmixing problem. Consider a robot that collects $N$ *in situ* spectral measurements along a path to

form a library of endmembers. Let the set of sampling locations along the path be $B$. Further, let the *in situ* spectral measurements be defined as $y = f(b) + \epsilon,\ \forall b \in B$, where $\epsilon$ is a sensor noise sampled from a normal distribution: $\epsilon \sim \mathcal{N}(0, \sigma)$. The set of all *in situ* measurements on path $B$ is then $Y_B = \{y_i : y_i \in \mathbb{R}^{\mathrm{d}},\ 1 \leq \mathrm{i} \leq \mathrm{N}\}$. Formally, the robot must collect samples that minimize the objective:

$$O(B) = E\left[\sum_{x_j \in X} \min_{a_j} \|Y_B a_j - x_j\|_2\right] \tag{4}$$

$$s.t.\ a_{ij} \geq 0\ \forall a_{ij} \in a_j,\ C(B) \leq \beta$$

where $C(B)$ is the sampling cost for the rover.

In its present form, (4) cannot be solved during the traverse as all the elements of $Y_B$ are not available. However, the solution can be approximated by replacing the elements in $Y_B$ of unvisited locations by the corresponding remote sensing measurements. Let $V_t \subseteq B$ be the set of locations the rover has sampled until time $t$ and $Y_{V_t}$ be the corresponding *in situ* spectra. Further, let $D_t \subseteq B$ be the set of unvisited locations and $X_{D_t}$ be the set of remote sensing measurements at those locations. Then as demonstrated in [5], solving (4) is analogous to solving the following objective function:

$$O(B|V_t) = \sum_{x_j \in X} \min_{a_j} \|[Y_{V_t}\ X_{D_t}]a_j - x\|_2 \tag{5}$$

$$s.t.\ a_{ij} \geq 0\ \forall a_{ij} \in a_j,\ C(V_t) + C(D_t) \leq \beta$$

## III. Methodology

The objective of our planner is to provide sampling locations that satisfy Equation (5). For this purpose, we first describe our planning environment and propose an MDP for navigating in that environment.

### A. MDP Formulation

We first address the inherent uncertainty present in the environment on account of *in situ* sensing. As the future *in situ* spectra will not be available at present time, this makes our environment partially observable. While it is certainly possible to apply POMDP based techniques to this problem, we display how using the inherent structure of SSE allows us to reduce this problem to solving multiple MDPs with complete observability. We now describe this MDP.

*1) State Definition:* Formally, we define the state $s$ as a tuple $(Y, V, X_D, D, X)$ where $V, D$ are sets of visited and unvisited locations respectively, $Y$ is the set of *in situ* spectra collected from the visited locations, $X_D$ is the set of remote observations corresponding to the locations in $D$ and $X$ is the orbital image. This definition is necessary to hold the Markov assumption. However, a consequence of including previous measurements into the state definition is that the size of the state-space increases significantly. In this regard, utilizing MCTS for finding the optimal policy is beneficial as the branching factor for the tree is only dependent on the size of the action-space. We limit our action-space to 8 actions to maintain tractability.

*2) Action Definition:* Let the set of all actions be $A$. As we do not have access to *in situ* measurements, we consider an action as adding a remote observation measurement to the spectral library. Thus, an action in this MDP corresponds to adding a new location to $D$ and the corresponding remote sensing observation to $X_D$. An action is said to be valid if the new location is reachable from the current rover's position. We assume the rover can only traverse on an eight-connected grid and the reachable locations are the immediate neighboring locations on the grid. For verbosity, let the set of valid actions for a state $s$ be $\Omega(s)$. The transition function $T$, is deterministic as every action from a particular state has a unique successor.

Notice that in this definition of MDP, the number of *in situ* measurements remains fixed for a particular MDP. Thus, at each time step $t$, all the states for $MDP(t)$ possess only those *in situ* measurements collected till time $t$, where $MDP(t)$ refers to the MDP generated at time $t$. This ensures full observability. While we cannot utilize Equation (4) for solving this MDP, the simplification described in Equation (5) can be fully defined using the state and actions as described by our MDP. This then allows us to develop planning strategies while maintaining tractability.

*3) Reward Definition:* Computing NNLS error and LS error for a given library of spectra is computationally expensive on account of the matrix inversions required for computing its solution. Consequently, it is infeasible to define our reward function as either of the two metrics. We chose to use the differential entropy as a reward for solving this MDP. Differential entropy has a closed form solution and is independent of the size of the orbital image. Moreover, differential entropy has been previously applied for science-aware exploration and shown to provide competent results [19], [20]. A brief description of it is given below.

Let $X$ be a random variable with a probability density function $f$ whose support is $\chi$. Then the differential entropy is defined as [21]:

$$h(X) = -\int_{\chi} f(x) log\left(f(x)\right) dx. \tag{6}$$

For a multivariate Gaussian random variable with covariance $\Sigma$, (6) has the following closed form solution [22]:

$$h(X) = \frac{1}{2} \ln |2\pi \mathrm{e} \Sigma|. \tag{7}$$

For a state $s$ in our MDP, we define our reward function $R(s)$ as:

$$R(s) = h(S) = \frac{1}{2} \ln |2\pi \mathrm{e} \Sigma_{S,S}| - \tau U(S). \tag{8}$$

where $S = [Y,\ X_D]$, $U(S)$ is a function that penalizes paths with multiple visits to the same locations and $\tau$ is a scalar weight.

### B. Non-myopic Planner for Spatio-Spectral Exploration

Here, we describe our planner named Non-Myopic Planner for Spatio-Spectral Exploration (NMPSE). Our approach comprises of creating and solving multiple trees using MCTS

in a sequential manner for each new MDP generated. The algorithm is shown in Algorithm 1. First, the current state is passed to the MCTS solver. The algorithm constructs a tree with the root as the current state and returns the optimal action for that state. The optimal action for this MDP consists of a sampling location on the map. Next, an *in situ* spectral measurement is taken at the location alluded by the optimal action. A new state is created by appending the new location to the set of visited locations and the sampled *in situ* spectral measurement to the set of *in situ* measurements. This is represented by the $Traverse$ function in the algorithm For this new state, we again construct a tree employing MCTS with the new state as the root node and find the next optimal action. In this way, additional *in situ* samples are collected until the sampling budget is exhausted. We now explain MCTS as applied to SSE.

MCTS is a partial tree search algorithm that efficiently balances the exploration versus exploitation trade-off. Each node in the tree corresponds to a particular state $s$ of the MDP. Applying an action $a$ from $s$ results in a successor state $s'$. This is depicted in the $Step$ function. The function returns $s'$ and a reward $r'$ which is computed according to Equation (8). The algorithm only expands nodes that result from applying an action $b \in \Omega(s)$.

If an action of the current node has not been tried before, that action is given precedence. The algorithm then rolls out from the resulting successor state until a fixed depth and returns the cumulative discounted reward. On the other hand, if all feasible actions of a state have been tried, the next action to take is based on the Upper Confidence Bound metric that efficiently trades-off exploration against exploitation. Line (28) in Algorithm 1 shows this specific metric. This process is repeated for the successor state until a fixed depth is reached.

The first new successor state resulting from either of the two cases is appended to the tree and the reward is backed up to the root of the tree. Finally, the action leading to the successor with the maximum cumulative reward from the root is returned. Note that all the nodes of the tree constructed from a particular MCTS call will have the same set of visited locations as no action in the MDP causes the rover to physically move to another location.

A key distinction between our method and traditional MCTS approaches is we use a particular MCTS tree to find just one optimal action. Once an optimal action has been found, the agent executes the action and constructs a new MCTS tree from the successor state. This is necessary as the previous tree does not possess information about the additional *in situ* measurement.

**Algorithm 1:** NMPSE

1 **function** `NMPSE(`$v_0,\ y_0$`)`
2 $V \leftarrow v_0$, $Y \leftarrow y_0$
3 $X_D, D \leftarrow \emptyset$
4 $s \leftarrow \{Y, V, X_D, D, X\}$
5 **while** $|Y| < \beta$ **do**
6 $T(root) \leftarrow s$
7 $a^* \leftarrow$ MCTS($T$)
8 $v_{new},\ y_{new} \leftarrow$ Traverse($a^*$)
9 $V \leftarrow \{V \cup v_{new}\}$
10 $Y \leftarrow \{Y \cup y_{new}\}$
11 $X_D, D \leftarrow \emptyset$
12 $s \leftarrow \{Y, V, X_D, D, X\}$
13 **end**
14 **function** `MCTS(`$T$`)`
15 **for** $i = 0$ **to** *max_iterations* **do**
16 Simulate($T(root)$, 0)
17 **end**
18 **return** $\underset{a}{\operatorname{argmax}}\, Q(s_0, a)$
19 **function** `Simulate(`$s,\ depth$`)`
20 **if** $\gamma^{depth} < \epsilon$ **then**
21 **return** 0
22 **end**
23 **if** $s$ *has untried actions* **then**
24 Sample $a$ from untried actions
25 $(s', r) \leftarrow$ Step($s$,$a$)
26 **return** $r + \gamma$·Rollout($s'$, $depth + 1$)
27 **end**
28 $a \leftarrow \underset{b \in \Omega(s)}{\operatorname{argmax}}\, Q(s, b) + \kappa\sqrt{\frac{log(N)}{N_a}}$
29 $(s', r') \leftarrow Step(s, a)$
30 $G \leftarrow r' + \gamma$·Simulate($s', depth + 1$)
31 $N(s) \leftarrow N(s) + 1$
32 $N(sa) \leftarrow N(sa) + 1$
33 $Q(s, a) \leftarrow Q(s, a) + \frac{G - Q(s,a)}{N(sa)}$
34 **return** G
35 **function** `Rollout(`$s,\ depth$`)`
36 **if** $\gamma^{depth} < \epsilon$ **then**
37 **return** 0
38 **end**
39 $a \sim \Omega(s)$
40 $(s', r') \leftarrow$ Step($s$, $a$)
41 **return** $r' + \gamma$·Rollout($s'$, $depth + 1$)
42 **function** `Step(`$s,\ a$`)`
43 $\{Y, V, X_D, D, X\} \leftarrow$ Unwrap(s)
44 $d_{new}, x_{new} \leftarrow$ Unwrap(a)
45 $s' \leftarrow \{Y, V, \{X_D \cup x_{new}\}, \{D \cup d_{new}\}, X\}$
46 $r' \leftarrow Reward(s')$
47 **return** $(s',\ r')$

## IV. Experiments

### A. Environment Description

We simulate an exploration scenario using a high resolution data acquisition system. Specifically, we use the data obtained by Airborne Visible Near Infrared Spectrometer - New Generation (AVIRIS-NG) [3], [23] as proxy for *in situ* spectra. AVIRIS-NG observes spectra in the range of

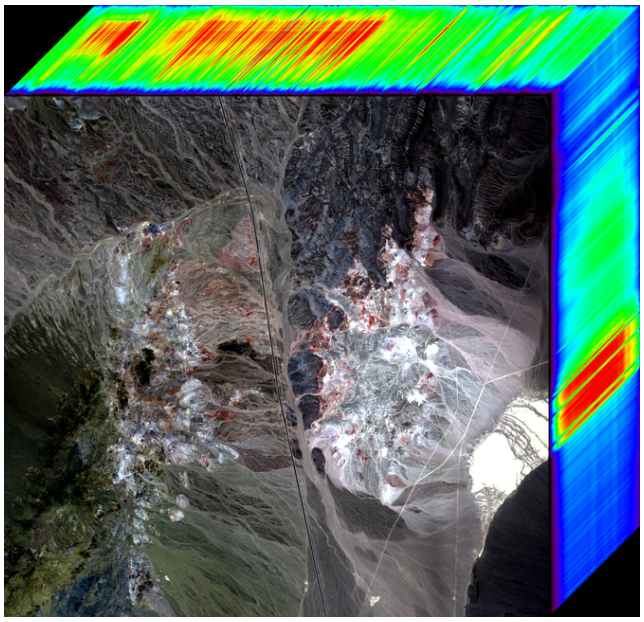

Fig. 2. Spectroscopic map of Cuprite as measured by AVIRIS-NG.

$[0.38\mu m,\ 2.5\mu m]$ at a spectral resolution of $5nm$ and an exceptionally high spatial resolution (3.7 m/pixel). Studies have shown that AVIRIS-NG measurements are a good analog to *in situ* infrared spectra [24], with inferences drawn on AVIRIS-NG transferring well onto data collected by *in situ* sensing instruments.

For remote sensing measurements, we use data obtained by the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) [25]. The instrument consists of multiple cameras possessing three visible short-wave infrared (VSWIR) and six short-wave infrared (SWIR) bands. In fact, other studies have also used AVIRIS-NG and ASTER measurements as proxies for low and high resolution spectral measurements, respectively [5], [26], [27]. We evaluate our approach on measurements taken at from a mining district in Cuprite, Nevada; a well-studied site with high mineralogical diversity [28]. Figure 2 shows an example of the AVIRIS-NG spectroscopic map of Cuprite. We were able to associate the two instruments' observations by aligning them with respect to both their spatial and spectral dimensions. We first registered both images with a planar homography approach. We then used the empirical line method [29] to find the correspondence between the ASTER and AVIRIS-NG reflectance values.

### B. Comparison with Baselines

Next, we empirically evaluate out approaches against a pair of uninformed planners: Fixed Step and Random, and an informed planner: GSS. A brief description of them is provided below.

*1) Fixed Step:* This strategy sequentially samples waypoints uniformly along one direction until the sampling budget is exhausted.

*2) Random:* In this case, the robot randomly samples a location from the reachable locations of the robot.

*3) Greedy Spectrum Selection:* We implemented the static planner named Greedy Spectrum Selection (GSS). The algorithm is displayed in Algorithm 2. For a detailed explanation, refer [5]. The algorithm has been modified to use the MES loss instead of NNLS loss. It greedily selects the next sample from a set of waypoints that maximizes (7) under some traversal constraint. It does not adaptively replan the path after each *in situ* spectral measurement is collected. In order to conduct a fair comparison between the two, we provide the same start locations for both the algorithms. Further, we define the waypoints as all the points on the eight-connected grid which our rover can traverse in the MDP. In addition, GSS algorithm requires a goal location and a path traversal budget. We pass the same goal location and the traversal cost observed during NMPSE evaluation for a particular initial rover configuration.

**Algorithm 2:** Greedy Spectral Selection

**function** GSS ($V_{start}$, $V_{end}$, $X$, $\beta$, $W$)
  $Q \leftarrow \{V_{start}, V_{end}\}$
  **while** $C(Q) < \beta$ **do**
    **forall** $v \in W \setminus Q$ **do**
      $X_V \leftarrow \{X_Q, X_v\}$
      $R(Q \cup v) = \frac{1}{2} \ln |2\pi e \Sigma_{X_V, X_V}|$
      **if** $R(Q \cup v) > R^*$ **then**
        $R^* \leftarrow R(Q \cup v)$
        $v^* \leftarrow v$
      **end**
    **end**
    $Q \leftarrow \{Q, v^*\}$
  **end**

We used mean reconstruction error (MRE) of the collected samples as the performance metric in all of our evaluations. The metric is simply (4) averaged over 50 trial runs. Apart from comparing NMPSE with the planners mentioned above, we also evaluate the effects of varying the hyperparameters of the MCTS in NMPSE on its computational efficiency.

## V. Results

### A. Evaluation of MCTS Parameters

We first evaluate the performance of NMPSE by varying the depth of the MCTS. Increasing the depth of the tree would cause the planner to consider longer horizons in its planning strategy. This in turn, would lead to an increase

TABLE I

Comparison of Mean Reconstruction Error (RE) and average computation time for MCTS with different maximum tree depths.

| Depth | Mean RE NMPSE | Average Action Time (sec) |
|---|---|---|
| 5 | 783.64 | **0.7451** |
| 7 | 776.1989 | 1.0570 |
| 10 | **758.3488** | 1.5304 |

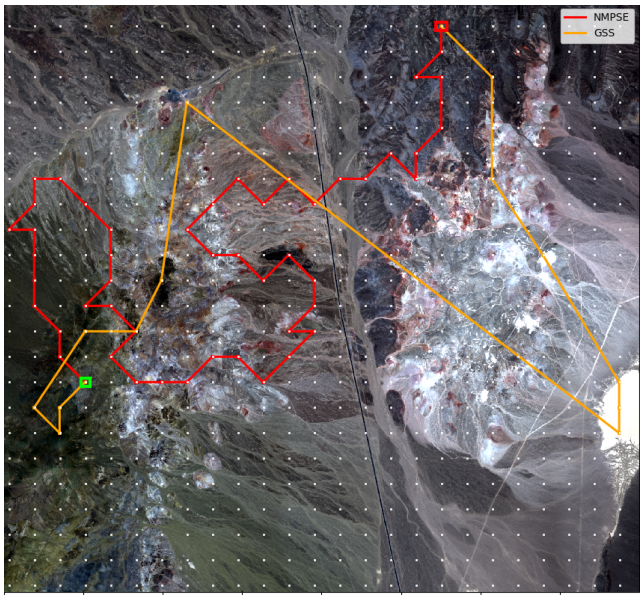


Fig. 3. Example paths from NMPSE (red) and GSS (orange) when constrained by traversal budget overlaid on the map of Cuprite mining site. The white dots on image refer to discretized grid of sampling points. The green and red squares correspond to start and end locations of the 500m traverse.

TABLE II
COMPARISON OF MEAN RECONSTRUCTION ERROR (MRE) AND T-TEST FOR NMPSE AND GSS

| Path Length | MRE NMPSE | MRE GSS | $p < 0.05$ |
|---|---|---|---|
| 250 | **1200.79** | 2533.62 | Y |
| 500 | **783.64** | 2033.13 | Y |
| 750 | **592.19** | 1066.78 | Y |
| 1000 | **469.61** | 563.38 | Y |
| 1250 | 396.32 | **372.53** | **N** |

in the computation cost. The discount factor $\gamma$ was kept at 0.9 and the number of Monte-Carlo simulations during each MCTS call were kept at 500. MRE and the average action time which is the mean time taken by MCTS to provide one optimal action is provided in Table I. The MRE decreased with increasing depth. However, the average action time increased at a faster rate. A depth of 5 seemed ideal for this problem as it allowed us to keep the computation time at less that one second. All the experiments explained below are implemented with 500 Monte-Carlo simulations, a max depth of 5 and a discount factor of 0.9.

### B. Evaluation of NMPSE

We divided the evaluation of NMPSE into two parts. The first evaluation was done by constraining the path length of the rover's traverse and the second evaluation was conducted by constraining the number of samples it can collect. A detailed explanation of each of the evaluations is given below.

*1) Constraining Path Length:* As GSS is constrained by traversal cost instead of sampling cost, it is more equitable to compare the performance of the planners against path length. The path cost is defined as 10 units for traversing to one of the neighbors of a point on the eight-connected grid. In a real-world scenario, this is equivalent to traversing approximately 300m at the Cuprite Site. NMPSE outperforms GSS in all but one case. The advantage of using NMPSE over GSS is that is produces significanly lower mean reconstruction error (MRE) with shorter path lengths. (See Figure 5 and Table II). This is preferable for planetary exploration as rover traverses rarely exceed 1km per command cycle [30]. Surprisingly, GSS is able to achieve comparable performance only with a traversal cost of 1000 and above. It was expected that GSS would more efficiently choose waypoints with a smaller path budget as well but since GSS is myopic in nature, it cannot find additional waypoints that do not exceed the traversal cost with a smaller path budget. Figure 3 illustrates this behavior. GSS greedily samples the most informative location, despite expending a significant portion of the traversal budget. This leaves it with a small portion of the budget in the latter parts of the traverse.

We evaluated statistical significance with a one-tailed t-test: comparing NMPSE with GSS. NMPSE outperforms GSS with a statistically significant difference as observed in Table II. At path length of 1250m, there is slightly better performance by GSS, but it did not achieve statistical significance. Moreover, from a practical point of view, rover traverses rarely exceed more than 1 km in a day [31] which roughly translates to around 750 units of path length in our simulated experiment. For real-world field experiments, which need to be executed with substantial traversal constraints, planners which arrive at good solutions with minimum traverse distance are preferable.

*2) Constraining Sampling Budget:* We evaluated NMPSE against two naive planning approaches; namely, random search and fixed step sampling. Figure 4 displays the performance for each of these planners while varying the sampling budget. NMPSE significantly outperformed both random sampling and fixed sampling planners. With increasing sampling budget, NMPSE reduced the reconstruction error at a faster rate than the other two planners. See Table 4. We evaluated statistical significance with two one-tailed t-tests: comparing NMPSE with random sampling and fixed sampling. As shown, NMPSE achieved a lower mean with statistical significance in all but one case. We believe NMPSE was not able to achieve $p < 0.05$ in that case due to small sample size. Figure 6 displays a path obtained from each of the planning strategies from one of the 50 experiments. Observe that NMPSE spends a significant sampling budget in areas with high mineral diversity.

## VI. CONCLUSION

This work proposed an MDP framework for solving the the SSE problem. By utilizing the inherent structure of the SSE objective, we showed how the problem can be made tractable to apply MDP solvers. We demonstrated the performance of an MCTS based planner against informed and uninformed planning algorithms utilized in this scenario. Our planner reasons about the future states by efficiently evaluating remote sensing measurements with the *in situ*

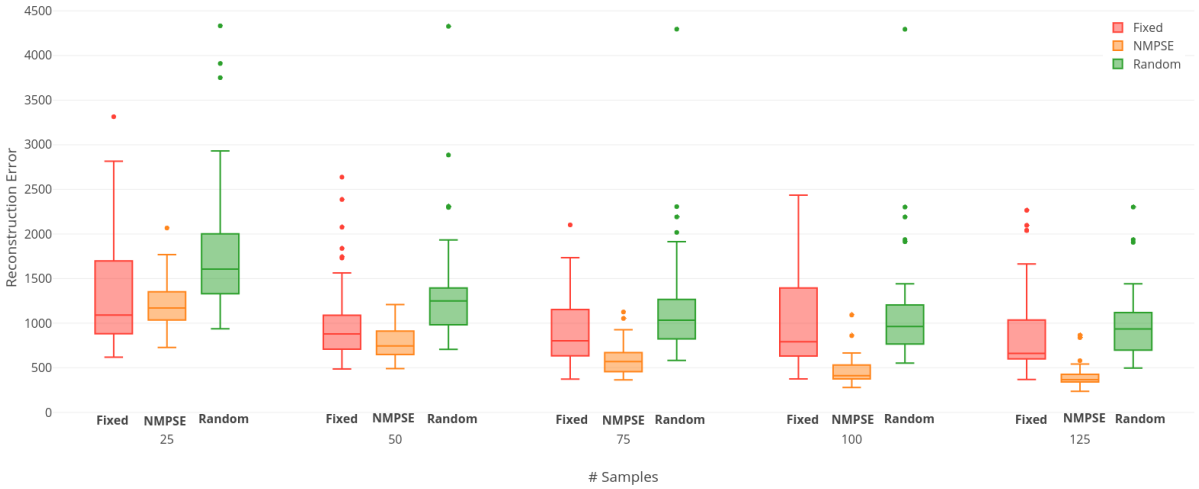


Fig. 4. Comparison of the performance of NMPSE (orange) with Fixed Sampling (red) and Random Sampling (green) with varying sampling budgets. All planners were initialized with the same rover configuration.

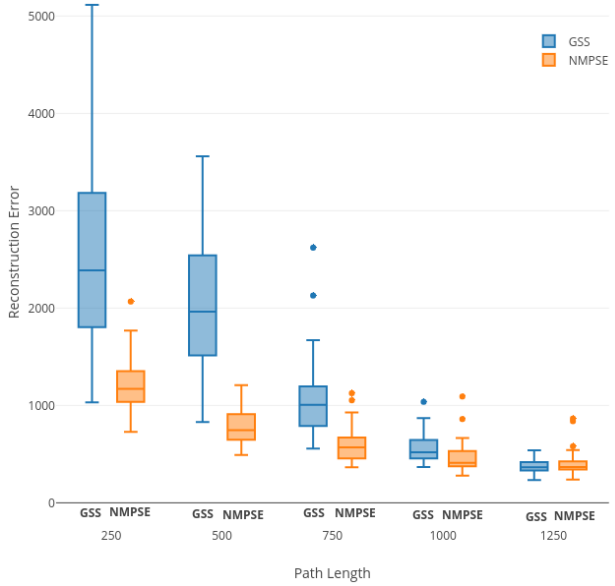


Fig. 5. Comparison of the performance of NMPSE (orange) against GSS (blue) with varying path lengths. Both planners were initialized with the same rover configurations.

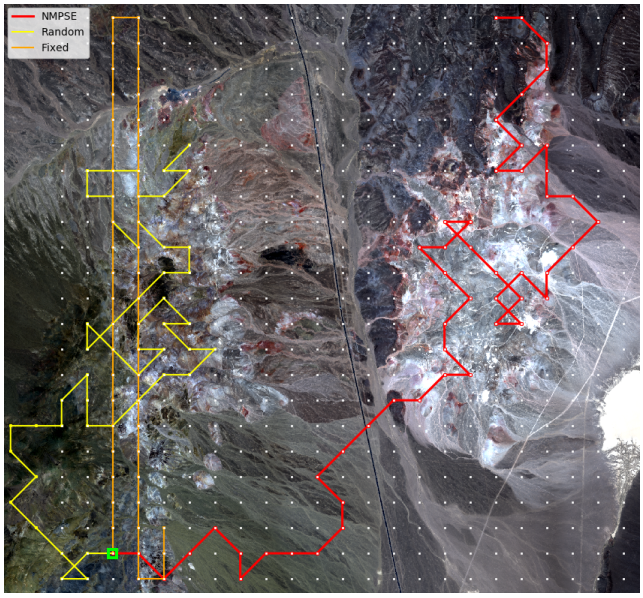


Fig. 6. Example paths from NMPSE (red), Random (yellow) and Fixed Step (orange) when constrained by sampling budget overlaid on the map of Cuprite mining site. The white dots on the image refer to discretized grid of sampling points. The green square corresponds to the start location of the 50-sample traverse.

TABLE III

COMPARISON OF MEAN RECONSTRUCTION ERROR (MRE) AND T-TEST FOR NMPSE, FIXED SAMPLING AND RANDOM SAMPLING PLANNERS WITH DIFFERENT SAMPLING BUDGETS.

| Samples | MRE Fixed | MRE NMPSE | MRE Random | Fixed $p < 0.05$ | Random $p < 0.05$ |
|---|---|---|---|---|---|
| 25 | 1269.95 | **1200.79** | 1779.89 | **N** | Y |
| 50 | 999.96 | **783.64** | 1333.05 | Y | Y |
| 75 | 949.32 | **592.19** | 1163.42 | Y | Y |
| 100 | 976.8 | **469.61** | 1090.02 | Y | Y |
| 125 | 879.55 | **396.32** | 974.09 | Y | Y |

measurements. NMPSE achieves lower mean reconstruction error compared to GSS under the same initial and final conditions. Moreover, NMPSE is *scalable* to large orbital images, unlike GSS which must iterate through each pixel of the orbital image for determining a sampling location.

Having conducted empirical evaluations of the planner using the AVIRIS-NG spectroscopic data, we will evaluate our approach in a field experiment at Cuprite utilizing a rover with a spectrometer on-board. NMPSE will provide waypoints to the robot navigation system which will then generate trajectories between waypoints that satisfy mobility constraints. From a robotics point of view, an interesting

area of future research is incorporating safety in navigation, specifically by considering risk in selecting sampling locations to avoid areas that are unsafe for the robot. Current orbital sensors also provide terrain information in the form of Digital Elevation Models (DEM). These DEMs can be used to compute the slope of a region which can inform the planner about safe regions.

## ACKNOWLEDGEMENT

This research was supported by the National Science Foundation National Robotics Initiative Grant #IIS-1526667. We acknowledge support of the NASA Earth Science Division for the use of AVIRIS-NG data. We gratefully acknowledge the assistance of Dr. David Thompson at Jet Propulsion Laboratory, California Institute of Technology. U.S. government support acknowledged.

## REFERENCES

[1] T. A. Estlin, B. J. Bornstein, D. M. Gaines, R. C. Anderson, D. R. Thompson, M. Burl, R. Castaño, and M. Judd, "Aegis automated science targeting for the mer opportunity rover," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 3, no. 3, p. 50, 2012.

[2] D. R. Thompson, D. S. Wettergreen, and F. J. C. Peralta, "Autonomous science during large-scale robotic survey," *Journal of Field Robotics*, vol. 28, no. 4, pp. 542–564, 2011. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/rob.20391

[3] R. O. Green, M. L. Eastwood, C. M. Sarture, T. G. Chrien, M. Aronsson, B. J. Chippendale, J. A. Faust, B. E. Pavri, C. J. Chovit, M. Solis, M. R. Olah, and O. Williams, "Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (aviris)," *Remote Sensing of Environment*, vol. 65, no. 3, pp. 227 – 248, 1998. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0034425798000649

[4] A. Plaza, J. A. Benediktsson, J. W. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, M. Marconcini, J. C. Tilton, and G. Trianni, "Recent advances in techniques for hyperspectral image processing," *Remote Sensing of Environment*, vol. 113, pp. S110 – S122, 2009, imaging Spectroscopy Special Issue. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0034425709000807

[5] D. R. Thompson, D. Wettergreen, G. Foil, P. M. Furlong, and A. R. Kiran, "Spatio-spectral exploration combining in situ and remote measurements," in *Association for the Advancement of Artificial Intelligence*, January 2015.

[6] A. Singh, A. Krause, and W. J. Kaiser, "Nonmyopic adaptive informative path planning for multiple robots." in *IJCAI*, vol. 3, 2009, p. 2.

[7] A. Krause, A. Singh, and C. Guestrin, "Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies," *Journal of Machine Learning Research*, vol. 9, no. Feb, pp. 235–284, 2008.

[8] J. Binney, A. Krause, and G. S. Sukhatme, "Informative path planning for an autonomous underwater vehicle," in *2010 IEEE International Conference on Robotics and Automation*, May 2010, pp. 4791–4796.

[9] D. R. Thompson, D. S. Wettergreen, and F. J. C. Peralta, "Autonomous science during largescale robotic survey," *Journal of Field Robotics*, vol. 28, no. 4, pp. 542–564, 8 2011. [Online]. Available: https://doi.org/10.1002/rob.20391

[10] T. Smith, "Probabilistic planning for robotic exploration," Ph.D. dissertation, Carnegie Mellon University, 2007.

[11] M. Woods, A. Shaw, D. Barnes, D. Price, D. Long, and D. Pullan, "Autonomous science for an exomars roverlike mission," *Journal of Field Robotics*, vol. 26, no. 4, pp. 358–390, 4 2009. [Online]. Available: https://doi.org/10.1002/rob.20289

[12] R. Bellman, *Dynamic Programming*, 1st ed. Princeton, NJ, USA: Princeton University Press, 1957.

[13] R. A. Howard, *Dynamic Programming and Markov Processes*. Cambridge, MA: MIT Press, 1960.

[14] P. Morere, R. Marchant, and F. Ramos, "Sequential bayesian optimization as a pomdp for environment monitoring with uavs," in *Robotics and Automation (ICRA), 2017 IEEE International Conference on*. IEEE, 2017, pp. 6381–6388.

[15] P. Clary, P. Morais, A. Fern, and J. Hurst, "Monte-carlo planning for agile legged locomotion," in *International Conference on Automated Planning and Scheduling*, 2018.

[16] S. Ross, G. J. Gordon, and J. A. Bagnell, "A reduction of imitation learning and structured prediction to no-regret online learning," in *AISTATS*, 2011.

[17] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[18] C. L. Lawson, . Hanson, Richard J., S. for Industrial, and A. Mathematics, *Solving least squares problems*, [rev. ed.] ed. Philadelphia : SIAM, 1995, "This SIAM edition is an unabridged, revised republication of the work first published by Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1974"–T.p. verso. [Online]. Available: http://www.zentralblattmath.org/zmath/en/search/?an=0860.65029

[19] D. R. Thompson, "Intelligent mapping for autonomous robotic survey," Ph.D. dissertation, Carnegie Mellon University, Pittsburgh, PA, August 2008.

[20] S. Gautam, B. S. Roy, A. Candela, and D. Wettergreen, "Science-aware exploration using entropy-based planning," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Sept 2017, pp. 3819–3825.

[21] T. Cover and J. Thomas, *Elements of Information Theory*, ser. A Wiley-Interscience publication. Wiley, 2006. [Online]. Available: https://books.google.com/books?id=EuhBluW31hsC

[22] N. A. Ahmed and D. V. Gokhale, "Entropy expressions and their estimators for multivariate distributions," *IEEE Transactions on Information Theory*, vol. 35, no. 3, pp. 688–692, May 1989.

[23] L. Hamlin, R. O. Green, P. Mouroulis, M. Eastwood, D. Wilson, M. Dudik, and C. Paine, "Imaging spectrometer science measurements for terrestrial ecology: Aviris and new developments," in *2011 Aerospace Conference*, March 2011, pp. 1–7.

[24] D. R. Thompson, V. Natraj, R. O. Green, M. C. Helmlinger, B.-C. Gao, and M. L. Eastwood, "Optimal estimation for imaging spectrometer atmospheric correction," *Remote Sensing of Environment*, vol. 216, no. May, pp. 355–373, 2018. [Online]. Available: https://doi.org/10.1016/j.rse.2018.07.003

[25] H. Fujisada, "Design and performance of aster instrument," in *Proc. SPIE*, vol. 2583, 1995, pp. 2583 – 2583 – 10. [Online]. Available: https://doi.org/10.1117/12.228565

[26] F. A. Kruse and S. L. Perry, "Regional mineral mapping by extending hyperspectral signatures using multispectral data," in *IEEE Aerospace Conference Proceedings*, 2007, pp. 1–14.

[27] D. R. Thompson, A. Candela, D. S. Wettergreen, E. N. Dobrea, G. A. Swayze, R. N. Clark, and R. Greenberger, "Spatial Spectroscopic Models for Remote Exploration," *Astrobiology*, vol. 18, no. 7, pp. 934–954, 2018.

[28] G. A. Swayze, R. N. Clark, A. F. Goetz, K. E. Livo, G. N. Breit, F. A. Kruse, S. J. Sutley, L. W. Snee, H. A. Lowers, J. L. Post, R. E. Stoffregen, and R. P. Ashley, "Mapping advanced argillic alteration at cuprite, nevada, using imaging spectroscopy," *Economic Geology*, vol. 109, no. 5, p. 1179, 2014. [Online]. Available: http://dx.doi.org/10.2113/econgeo.109.5.1179

[29] G. M. Smith and E. J. Milton, "The use of the empirical line method to calibrate remotely sensed data to reflectance," *International Journal of Remote Sensing*, vol. 20, no. 13, pp. 2653–2662, 1999. [Online]. Available: https://doi.org/10.1080/014311699211994

[30] D. Wettergreen, N. Cabrol, V. Baskaran, F. Calderón, S. Heys, D. Jonak, R. A. Luders, D. Pane, T. Smith, J. Teza *et al.*, "Second experiments in the robotic investigation of life in the atacama desert of chile," in *Proc. 8th International Symposium on Artificial Intelligence, Robotics and Automation in Space*, 2005.

[31] M. Maimone, Y. Cheng, and L. Matthies, "Two years of visual odometry on the mars exploration rovers," *Journal of Field Robotics*, vol. 24, no. 3, pp. 169–186, 2007.