Robotic Exploration using Generalized Behavioral Entropy

Aamodh Suresh, Carlos Nieto-Granda, Sonia Martínez

Abstract—This work presents and evaluates a novel strategy for robotic exploration that leverages human models of uncertainty perception. To do this, we introduce a measure of uncertainty that we term "Behavioral entropy", which builds on Prelec's probability weighting from Behavioral Economics. We show that the new operator is an admissible generalized entropy, analyze its theoretical properties and compare it with other common formulations such as Shannon's and Renyi's. In particular, we discuss how the new formulation is more expressive in the sense of measures of sensitivity and perceptiveness to uncertainty introduced here. Then we use Behavioral entropy to define a new type of utility function that can guide a frontier-based environment exploration process. The approach's benefits are illustrated and compared in a Proof-of-Concept and ROS-Unity simulation environment with a Clearpath Warthog robot. We show that the robot equipped with Behavioral entropy explores faster than Shannon and Renyi entropies.

Index Terms—Robot Exploration, Human-Centered Robotics, Planning under Uncertainty, Information Theory

I. INTRODUCTION

AUTONOMOUS robotic exploration is critical as it alleviates the inaccessibility and dangers of remote, unsafe, and risky environments. Depending on the environment and task at hand, the robot might be expected to exhibit a variety of exploration behaviors from coarse and fast to fine and detailed. These tasks are invariably evaluated or supervised by humans and the exploration objective could change dynamically. Thus having an intuitive, diverse and theoretically grounded exploration framework is a necessity. In this work, we infuse human perception characteristics into robotic exploration AI, and investigate both theoretically and practically, the effects of this infusion.

The fundamental approach to exploration consists of reducing uncertainty by observing an area with noisy sensors. Uncertainty is typically quantified by entropy [1] and majority of exploration policies aim at reducing it (implicitly or explicitly). However, unlike the typical robotic AI designs, it is well known that humans perceive uncertainty in a fundamentally non-rational manner [2], [3], [4], especially in sensory perception and evaluating outcomes [5], causing a variety of behaviors and decision making. Motivated by this, we first propose and characterize a novel measure of uncertainty called Behavioral entropy that incorporates human

models of uncertainty perception [2] that is widely used in the Behavioral Economics literature. Then, we investigate the design of new behavioral robotic exploration policies using our proposed uncertainty measure.

Related Work: Robotic exploration broadly involves the cyclic interplay between *perception* (knowledge of environment and self via SLAM [6]), *decision* (evaluating areas of interest (AOIs) from current knowledge) and *action* (navigation policies to reach the selected AOI). Several SLAM techniques have been proposed to produce environment maps utilizing visual odometry [7] and LiDAR [8] that are key components for exploration. Similarly, there exists navigation pipelines utilizing a combination of global planning [9], [10], local planning [11], optimal control [12] and reactive control [13] to reach an AOI safely, reliably and efficiently. We focus on the *decision* aspect of identifying and evaluating potential AOIs to navigate from a given map. In particular, we use the popular frontier-based exploration [14], that has been successfully used to explore challenging environments [15], [16]. These strategies are driven by optimizing utility functions that measure notions of information gain or uncertainty reduction by visiting these AOIs. Thus, the metrics employed in these utility functions are a key factor in designing exploration policies. The most commonly used metric is based on the Shannon information gain [14] that uses the Shannon entropy [1] as a measure of uncertainty. More recently, information metrics derived from Renyi's entropy have been proposed and shown to perform better than Shannon's information gain [17]. Renyi's entropy generalizes Shannon's to a family of generalized entropies [18]. However, these entropies and the corresponding metrics do not incorporate human behavioral models. Moreover, their tuning is non-intuitive and Renyi's entropy is not continuously differentiable w.r.t. its parameter (at 1) and thus making tuning also difficult. Some works try to address modelling human choice and uncertainty perception through data intensive approaches like Boltzmann machines [19] and Bayesian learning [20]. However data-based approaches require large datasets and computational effort, provide limited theoretical insights, can suffer from poor generalization and are non-intuitive to tune and implement. Thus, we formulate a new entropy that is based on human models of uncertainty perception and which provides the most broad set of uncertainty perception tools to tackle the exploration problem, while also being smooth, theoretically grounded and intuitive to tune.

Contributions: In this work, we introduce a novel operator "Behavioral entropy" to quantify subjectively perceived uncertainty of a given probability distribution. Our entropy design is intuitive and empirically grounded from Behavioral Economics, as well as "admissible" as a generalized entropy

and theoretically grounded from Information theory perspective. We provide novel measures of "Sensitivity" and "Perceptiveness" to characterize behavioral uncertainty perception of entropic measures; which provides a basis to characterize and compare any admissible entropic measure. We show that Behavioral entropy is more general in the sense that it can capture the widest range of uncertainty "perceptions" w.r.t. Shannon's and Renyi's (which is made more precise later). We also show that in this regard, Renyi's entropy is more general than Shannon's. Then, we construct a novel exploration utility function based on Behavioral entropy and combine it into a frontier-based exploration approach. We further analyze and show that our exploration formulation is more expressive and can capture a wider range of perceived utility than others that use standard Bolzmann-Gibbs-Shannon (BGS) and Renyi entropies. We perform proof of concept simulations with ideal SLAM and robot navigation in a custom occupancy map to evaluate our proposed method in isolation, and show that our method is faster than methods with other entropies. Finally, we evaluate our framework in a complex and realistic Unity-ROS simulation environment used in the DARPA Subterreanean challenge [21], with a Clearpath Warthog robotic system with a LiDAR and IMU based SLAM system. We show that the robot equipped with our behavioral entropy based method, explores the environment quicker than other methods.

II. GENERALIZED ENTROPY AND PROBLEM FORMULATION

We start introducing basic notations[1], a concise description of generalized entropy, followed by our problem formulation.

Generalized entropy: Entropy is a measure of uncertainty associated with the probability distribution of a random variable. More precisely, consider the set of probability distributions $\mathscr{P}_M$ of discrete random variables of M outcomes; that is, $(p_1,\ldots,p_M) \in \mathscr{P}_M$, if $0 \le p_i \le 1$, for $i = 1,\ldots,M$, and $\sum_{i=1}^{M} p_i = 1$. Let $H : \mathscr{P} \triangleq \cup_{M=1}^{\infty} \mathscr{P}_M \to \mathbb{R}_{\ge 0}$ be a permutation-invariant operator. We say that H belongs to the class of entropy functions if it satisfies the following Shannon-Kinchin axioms [22], [1]:

A1 *Continuity:* H is continuous over each $\mathscr{P}_M$.
A2 *Maximality:* $H(p) \le H(\frac{1}{M},\ldots,\frac{1}{M})$, and the equality holds if and only if $p_i = \frac{1}{M}$, for all $i \in \{1,\ldots,M\}$. This axiom implies that the uniform distribution attains the highest entropic measure.
A3 *Expansibility:* A sure outcome added to a probability distribution p does not change its entropy. i.e.:
$H(p_1,\ldots,p_i,0,p_{i+1},\ldots,p_M) = H(p_1,\ldots,p_M)$.
A4 *Separability or Strong Additivity:* For any two $p,q \in \mathscr{P}_M$, let $H(p,q)$ denote the entropy of a joint distribution of p,q, then $H(p,q) = H(p) + H(q|p)$, where $q|p$ is the conditional distribution of q given p.

Shannon proved [1] that an entropic function satisfying all the four axioms is necessarily a scaled BGS entropy (1a). We focus on generalized entropies defined as follows.

Definition 1 (Generalized entropy)**.** *An entropic function H is an admissible generalized entropy, if it satisfies the first three entropy axioms A1-A3.*

BGS (Shannon) entropy H^{S} and Renyi entropy H^{R} (generalises many other entropies like Tsallis entropy [18]) are well known admissible generalized entropic functions, defined as follows:

$$H^{\text{S}}(p) = -k \sum_{i=1}^{M} p_i \log(p_i), \ k \in \mathbb{R}_{>0}, \tag{1a}$$

$$H^{\text{R}}(p) = \frac{1}{1-\gamma} \log\left(\sum_{i=1}^{M} p_i^{\gamma}\right), \ \gamma \in \mathbb{R}_{>0}, \ \gamma \ne 1, \tag{1b}$$

Robotic Exploration Setting: Our application of interest is robotic exploration. To this end, we represent the environment where the robot is deployed as a compact set $\mathscr{X} \subset \mathbb{R}^2$. The discretization of $\mathscr{X}$ using square grid cells results into the set of grid elements $\mathscr{D}$ and an associated occupancy map $\mathscr{M}$ representing the likely location of obstacles in the environment. More precisely, consider an occupancy function $f_{\text{occ}} : \mathscr{D} \to [0,1]$, which assigns probability of occupancy of each cell in $\mathscr{D}$. The occupancy map $\mathscr{M}$ is the discrete field of f_{occ} over $\mathscr{D}$. We partition $\mathscr{D}$ and, subsequently, $\mathscr{M}$ into the unknown space $\mathscr{D}^{\text{UK}} \triangleq \{x \in \mathscr{D} : f_{\text{occ}}(x) \triangleq 0.5\}$ (resp. $\mathscr{M}^{\text{UK}}$), the uncertain space $\mathscr{D}^{\text{U}} \triangleq \{x \in \mathscr{D} : f_{\text{occ}}(x) \in (0,0.5) \cup (0.5,1)\}$ (resp. $\mathscr{M}^{\text{U}}$), and the known space $\mathscr{D}^{\text{K}} \triangleq \{x \in \mathscr{D} : f_{\text{occ}}(x) \in \{0,1\}\}$ (resp. $\mathscr{M}^{\text{K}}$). The process of "exploration" aims to reduce the volume of unknown space $\mathscr{D}^{\text{U}}$ and increasing the volume of known space $\mathscr{D}^{\text{K}}$. In an ideal setting, exploration is complete when $\mathscr{D}^{\text{K}} = \mathscr{D}$. However, factors like sensing noise and obstacles might hinder uncertainty reduction and could make unknown areas unobservable and remain uncertain.

The robot uses LiDAR to perceive its immediate environment, IMU and odometry to approximate its internal state. We use the popular SLAM [6] method called "Omnimapper" [8] to simultaneously estimate the robot's pose $x \in SE(2)$ and estimated occupancy map $\mathscr{M}$. The LiDAR provides sensor measurements of an area $\mathscr{A} \subseteq \mathscr{D}$ around the robot and the SLAM algorithm updates $\mathscr{A}$ in $\mathscr{M}$ using physical models of sensor.

The frontiers $F \triangleq \{f_1,\ldots,f_N\}$ where $f_i \in SE(2)$ are identified from the current occupancy map $\mathscr{M}$, and contain locations to explore next. We employ a utility function $u : SE(2) \times SE(2) \times \mathbf{P}(\mathscr{M}) \to \mathbb{R}_{\ge 0}$ to quantify the reward of the robot being at a pose x and observing an area $\mathscr{A} \subset \mathscr{M}$ around a frontier f_i; $u(x,f_i,A) \in \mathbb{R}_{\ge 0}$. In this work, the rewards of exploring an area A are given in terms of an information gain. After identifying the best frontier to go next, the navigation manager in the robot plans global paths and local control actions to reach the selected frontier location.

The framework is visualized in Fig. 1 and will be explained in detail in Sec. IV.

III. BEHAVIORAL ENTROPY AND ITS ANALYSIS

We derive a new measure of uncertainty, $H^{\text{B}} : \mathscr{P} \to \mathbb{R}_{\ge 0}$, which we term Behavioral entropy. This function aims to quantify the subjectively perceived randomness present in a probability distribution.
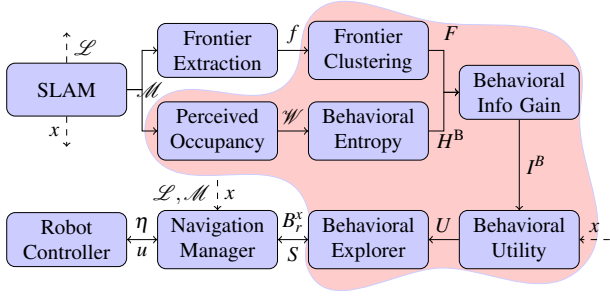
Fig. 1: Proposed exploration framework with our contribution in the red cloud. The Omnimapper SLAM system provides a local map $\mathscr{L}$, a global map $\mathscr{M}$ and robot localization x. Then, frontiers f and perceived occupancy $\mathscr{W}$ are extracted from $\mathscr{M}$. The frontiers f are clustered (F) and Behavioral entropy H^{B} is calculated from $\mathscr{W}$. For each cluster in F, a behavioral information gain I^B is obtained from H^{B} that corresponds to the measurement model. Then, utilities U are calculated from the information I^B and pose estimate x. The explorer then picks a suitable goal region B_r^x and sends it to the navigation manager, which in turn sends global and local plans η to the controller.
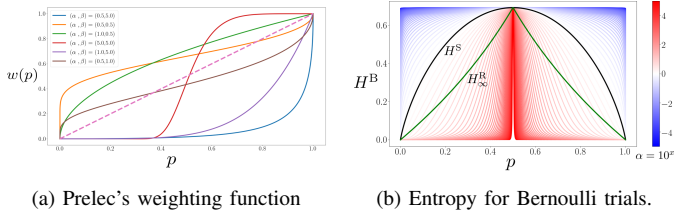
(a) Prelec's weighting function   (b) Entropy for Bernoulli trials.

Fig. 2: (Left) Prelec's probability weighting function with a few parameter choices. Y-axis $w(p)$ indicates the perceived uncertainty associated with probability p (X-axis). Dotted line indicates the unity curve $w(p) = p$. (Right) Behavioral entropy variation with α in log scale for a Bernoulli trial. Black curve indicates Shannon's entropy.

Consider a probability vector $p = (p_1, \ldots, p_M) \in \mathscr{P}_M$, and let $w : [0,1] \to [0,1]$ be *a probability weighting function*, which maps $p_i \mapsto w(p_i) \equiv w_i$ and transforms the ith outcome p_i into a perceived probability w_i. Among the several probability weighting functions from the literature [3], [23], we employ the popular Prelec's model [2]:

$$w(p) = e^{-\beta(-\log p)^\alpha}, \quad \alpha > 0, \ \beta > 0, \ w(0) = 0, \quad (2)$$

see some examples in Fig. 2a. The parameter α controls the convexity or concavity of w: taking $\alpha \approx 0$ results in a highly concave w that over-weights probabilities (i.e $w(p) \gg p$), while taking $\alpha \to \infty$ makes of w a highly convex function that under-weights probabilities (i.e $w(p) \ll p$). The parameter β controls the unique fixed point p^* of w, i.e. the intersection with $w(p) = p$, the dotted line in Fig. 2a. A $\beta \approx 0$ indicates fixed point $p^* \approx 0$, while a $\beta \to \infty$ brings $p^* \approx 1$. By varying α, β, we can model different behaviors. With low α and β values results into an "uncertainty averse" behavior (green and orange curves in Fig. 2a), with $w(p) > p$ which implies more certainty in unlikely outcomes. With high β values we get "uncertainty insensitive" behavior when $w(p) < p$ (blue and brown curves in Fig. 2a), implying that the decision maker only considers more certain outcomes. For more properties of these weighting functions, see [23].

A. Behavioral entropy and its analysis

Our proposed "Behavioral Entropy" H^{B} combines the notion of BGS entropy with Prelec's weights w from (2).

$$H^{\mathrm{B}}(p_1, \ldots, p_M) = -\sum_{i=1}^{M} w(p_i) \log(w(p_i)). \quad (3)$$

Thus, incorporating the behavioral modelling capacity of Prelec's weighting function into an entropic function. To analyze the new operator, we first list some properties of Prelec's probability weighting functions w.

Proposition 1 (Perceived probability [23], Prop. 2.11). *The class of weighting functions w in* (2) *satisfies:*

1) *w is continuous, strictly increasing, and $w(1) = 1$.*
2) *In addition to 0 and 1, w only has another fixed point at $p^* = e^{-(\frac{1}{\beta})^{\frac{1}{\alpha-1}}} \in (0,1)$.* $\square$

We use Proposition 1 to control fixed points. Specifically, to obtain $p^* = 1/M$ for any integer $M \geq 2$, we substitute $p^* = \frac{1}{M}$ in 2) above and simplify terms to get conditions on α, β that guarantee $p^* = 1/M$:

$$\beta = e^{(1-\alpha)\log(\log(M))}. \quad (4)$$

Now we recall the following well-known result.

Lemma 1 (Maximality of H^{S}). *The BGS entropy function H^{S} attains its unique global maximum at $p \triangleq (\frac{1}{M}, \ldots, \frac{1}{M})$ (uniform distribution) over $\mathscr{P}_M$.* $\square$

Next, we show that H^{B} is admissible as a generalized entropy according to Definition 1 under (4).

Theorem 1. *Fix $M \in \mathbb{N}$. If α, β satisfy* (4)*, then the function H^{B} is admissible as a generalized entropy.*

Proof. We need to show axioms A1, A2 and A3 are satisfied by H^{B}. Axioms A1 and A3 are trivially satisfied from the properties of w in Proposition 1 and the fact that H^{S} satisfies both A1 and A3. Regarding, A2, we observe that $H^{\mathrm{B}}(\mathbf{p}) = H^{\mathrm{S}}(\mathbf{w})$, where $\mathbf{p} = (p_1, \ldots, p_M) \in \mathscr{P}_M$ and $\mathbf{w} = (w_1, \ldots, w_M)$ is the corresponding Prelec weights (2). From Lemma 1, H^{S} is maximized only with $\mathbf{w}$ with $w_i = \frac{1}{M}$, for all i. Then for any α, β satisfying (4), the non-trivial fixed point of w is necessarily $1/M$ and is unique. Hence $w_i = \frac{1}{M}$, for all i only when $p_i = \frac{1}{M}$, for all i. Thus, H^{B} is maximized only when $\mathbf{p}$ is the uniform distribution. $\square$

Remark 1. *The Prelec's function* (2) *with conditioning* (4) *allows us to control the fixed point of the function which is critical for admissibility guarantees, as well as the shape of the perceived probability curve, ensuring different behavioral perceptions. Whereas, other popular probability weighting functions [23] like Karmakar's (fixed point is fixed at 0.5) or Tversky and Kahneman's (shape and fixed point can't be controlled independently) cannot be used to generate meaningful and admissable generalized entropies.* •

Next, we will compare the properties and performance of Behavioral entropy with other entropic functions.

B. Entropy comparison

First, we introduce the metrics that we use to make comparisons of different entropies.

Definition 2 (Sensitivity). *Fix any generalized entropy function H and assume that H is normalized so that it is bounded above by $\log(M)$ (the maximizer of H^S). Then, $S(H) \in \mathbb{R}_{\geq 0}$, the entropic sensitivity of H, is defined as*

$$S(H) = \int_{\mathscr{P}_M} H(\mathbf{p}) d\text{Vol}, \tag{5}$$

where Vol is the standard Lebesgue measure over $\mathscr{P}_M$. The value $S(H)$ provides a measure of the average sensitivity of H to uncertainty over all distributions $\mathbf{p} \in \mathscr{P}$. If $S(H_1) > S(H_2)$, then the overall randomness perceived by H_1 is greater than that of H_2, and H_1 is more sensitive towards uncertainty.

Definition 3 (Perceptiveness). *Let H_θ be any generalized entropy function as in Definition 2 parametrized by $\theta \in \Theta \subseteq \mathbb{R}^A$. Let H_θ be continuous in θ, then $P(H) \in \mathbb{R}_{\geq 0}$, the perceptiveness of H, is defined as:*

$$P(H) = \max_{\theta \in \Theta} S(H_\theta) - \min_{\theta \in \Theta} S(H_\theta). \tag{6}$$

The above notion of perceptiveness captures the range of sensitivity towards uncertainty for a class of entropies H_θ, by considering all possible $\theta \in \Theta$.

Now, we evaluate the perceptiveness of the family of H^B. We begin by identifying a useful bound.

Lemma 2. *Consider a generalized entropy H_θ, which is bounded above by some $b \in \mathbb{R}_{>0}$, for all $\theta \in \Theta$. Then, $S(H_\theta)$ lies in the interval $(0, \frac{b}{M-1!})$, where M is the number of outcomes. Subsequently, $0 \leq P(H_\theta) \leq \frac{b}{M-1!}$.*

Proof. The volume of the probability simplex $\mathscr{P}_M \subseteq \mathbb{R}^M$ consisting of M outcomes is $\frac{1}{(M-1)!}$ [24]. The limits of $S(H_\theta)$ can be computed by considering that $0 \leq \inf_{\theta \in \Theta} \int_{\mathscr{P}_M} H_\theta(\mathbf{p}) d\text{Vol} \leq S(H_\theta) \leq \sup_{\theta \in \Theta} \int_{\mathscr{P}} H_\theta(\mathbf{p}) d\text{Vol}$. $\qquad\square$

With this let us look at the perceptiveness of H^B.

Theorem 2. *Consider $H^B_{\alpha,\beta} \equiv H^B_\alpha$ conditioned according to (4). The sensitivity $S(H^B_\alpha)$ lies in the interval $(0, \frac{\log(M)}{M-1!})$ its perceptiveness $P(H^B)$ reaches the maximum possible $\frac{\log(M)}{M-1!}$.*

Proof. We will first evaluate the limits of H^B in α and then apply it to sensitivity measurement s. We note that H^B is continuous in α and has a unique maximum at the uniform distribution $\mathbf{p}_{\text{uni}}$ when conditioned according to (4). The limits of $S(H^B)$ can be calculated as follows:

$$s_0 = \lim_{\alpha \to 0} S(H^B_\alpha) = \lim_{\alpha \to 0} \int_{\mathscr{P}_M} H^B_\alpha(\mathbf{p}) d\text{Vol}. \tag{7}$$

Since H^B is continuous and bounded above by $\log(M)$ for every $P \in \mathscr{P}$, and the volume of $\mathscr{P}$ region is finite, we can take the limit inside the integral and obtain:

$$
\begin{aligned}
s_0 &= \int_{\mathscr{P}_M} \lim_{\alpha \to 0} H^B_\alpha(P) dP \\
&= \int_{\mathscr{P}_M} \sum_{i=1}^{M} -\lim_{\alpha \to 0} (e^{-\beta_\alpha(-\log(p_i))^\alpha} \beta_\alpha(-\log(p_i))^\alpha)) \\
&= \int_{\mathscr{P}_M} M \times \frac{1}{M} \times \log(M) = \frac{\log(M)}{M-1!}.
\end{aligned}
$$

Next for the lower bound,

$$
\begin{aligned}
s_\infty &= \int_{\mathscr{P}_M} \lim_{\alpha \to \infty} H^B_\alpha(P) dP \\
&= \int_{\mathscr{P}_M} \sum_{i=1}^{M} -\lim_{\alpha \to \infty} (e^{-\beta_\alpha(-\log(p_i))^\alpha} \beta_\alpha(-\log(p_i))^\alpha)) \\
&= \int_{\mathscr{P}_M} M \times 0 \times \log(0) = 0,
\end{aligned}
$$

and the result follows. $\qquad\square$

Now we compare H^B with Renyi and Shannon entropies on the basis of perceptiveness.

Theorem 3. *Consider the family of $H_B \equiv (H^B_\alpha)$ conditioned according to (4). Then, $P(H^B) > P(H^R) > P(H^S)$.*

Proof. From (1), the value $P(H^R)$ can be obtained from appropriate limits of $S(H^B)$. The upper limit of Renyi sensitivity is reached when $\gamma \to 0$ and it is easy to see that coincides with that of Behavioral entropy:

$$s_0 = \lim_{\gamma \to 0} \int_{\mathscr{P}_M} H^R_\gamma(\mathbf{p}) d\text{Vol} = \frac{\log(M)}{M-1!}. \tag{8}$$

For the lower bound, since Renyi entropy is decreasing in $\gamma \in (0,1) \cup (1,\infty)$, the lowest sensitivity is reached when

$$s_\infty = \int_{\mathscr{P}_M} \lim_{\gamma \to \infty} H^R_\gamma(\mathbf{p}) d\text{Vol} = \int_{\mathscr{P}_M} -\log(\max\{p_1, \ldots, p_M\}) d\text{Vol}.$$

Since $0 < \max_{\mathbf{p} \in \mathscr{P}_M}\{p_1, \ldots, p_M\} \leq 1$, as $\mathbf{p}$ is a probability distribution, then $-\log(\max_{\mathbf{p} \in \mathscr{P}_M}\{p_1, \ldots, p_M\}) \geq 0$ for all $\mathbf{p}$. Hence, the lower limit $s_\infty(H^R) \geq 0$. Thus, we get $P(H^B) > P(H^R)$ from Theorem 2. Next it is trivial to show that $P(H^R) > P(H^S)$ as the perceptiveness of H^S is 0. $\qquad\square$

The concept of perceptiveness and its comparison with different entropies is visualized in Fig. 2b using Bernoulli trials. Behavioral entropy H^B is able to capture the entire behavior spectrum from "Uncertainty Averse" (blue region) to "Uncertainty ignorant" (red region). Shannon entropy H^S is a single curve (black). Renyi entropy H^R can capture "Uncertainty Averse" (blue region) behavior however, it can't be more "uncertainty ignorant" than H^R_∞, thus being significantly less perceptive than H^B. Next, we connect the Behavioral entropy notion to a robotic exploration setting.

IV. BEHAVIORAL EXPLORATION

Here, we provide details on the frontier selection and clustering process for robotic exploration, and explain the construction of the utility function used to choose the best frontier, which defines our framework in Fig. 1.

Frontier Extraction and Clustering: The known areas $\mathscr{M}^K$, unknown areas $\mathscr{M}^{UK}$, and uncertain areas $\mathscr{M}^U$ of the occupancy map $\mathscr{M}$ are extracted from the occupancy values as described in Sec. II. Then, the free space $\mathscr{M}^F$, resp. the occupied space $\mathscr{M}^O$, is obtained by collecting cells whose occupancy probability is less than τ^{FS}, resp. greater than τ^{OB}. Gradient vectors are computed over $\mathscr{M}$ and the non-zero cells are stored in $\mathscr{M}'$ defining all possible frontier candidates. Obstacle neighbor cells $\mathscr{M}^{ON}$ are given by the convolution of $\mathscr{M}^O$ with a suitable kernel K^N. Frontier cells $\mathscr{F}$ are the free cells that are part of the gradient but not neighbors of obstacles. As this can create a large number of frontier cells,

these are clustered $\mathscr{F}^C$ according to the cluster size $\tau^{CL} \in \mathbb{Z}_{\geq 0}$, which indicates the size of square grid that clusters frontiers into a single frontier. Greater τ^{CL} provides lesser number of clusters while $\tau^{CL} \triangleq 1$ results in no clustering. A representative frontier is picked at random in each cluster to give the final list of chosen frontiers $\mathscr{F}^L$, whose size is equal to the number of clusters. This procedure is summarized in Algorithm 1.

Algorithm 1: Frontier extraction and clustering

---
1 Input: $\mathcal{M}, \tau^{OB}, \tau^{FS}, \tau^{CL}, K^N$;
2 Output : $\mathscr{F}^L$;
3 $\mathcal{M}^K, \mathcal{M}^{UK}, \mathcal{M}^U \leftarrow \mathcal{M}$;   // Partition map
4 $\mathcal{M}^F \leftarrow (\mathcal{M}^K \cup \mathcal{M}^U) < \tau^{FS}$;   // Get free space
5 $\mathcal{M}^O \leftarrow (\mathcal{M}^K \cup \mathcal{M}^U) > \tau^{OB}$; // Get occupied space
6 $\mathcal{M}' \leftarrow \mathcal{M}$;   // map areas with non-zero gradients
7 $\mathcal{M}^{ON} \leftarrow \mathcal{M}^O \circ K^N$;   // obstacle neighbor cells
8 $\mathscr{F} \leftarrow (\mathcal{M}^F \cap \mathcal{M}') - \mathcal{M}^{ON}$;   // free gradient cells
9 $\mathscr{F}^C \leftarrow$ Cluster$(\mathscr{F}, \tau^{CL})$;   // cluster frontiers
10 $\mathscr{F}^L \leftarrow$ Random Pick$(\mathscr{F}^C)$

Information Gain and Utility function: A sensor footprint $\mathscr{A} \subset \mathcal{M}$ is obtained around each $f \in \mathscr{F}$, which is a collection of occupancy cells according to the sensor model. We use the LiDAR beam model [6] with range and number of beams to determine the cells that get observed. We then update the occupancy for each cell in $\mathscr{A}$ based on the number of beams that hit a cell. Next, a perceived occupancy is calculated using (9). Here, the behavioral information gain I^B is calculated as the reduction in entropy from observing the area $\mathscr{A}$ around f. We then use the following utility function to determine the value $u \in \mathbb{R}_{\geq 0}$ of going to f' from the current pose x and map $\mathcal{M}$:

$$u^B(f', x; \mathcal{M}) = \frac{I^B(f'; \mathcal{M})}{|\eta(x, f')|}, \ I^B(f'; \mathcal{M}) = \sum_{x' \in \mathscr{A}} H^B(p((x'))) \quad (9)$$

Here, $I^B(f'; \mathcal{M})$ is the information gained or entropy reduced by observing area $\mathscr{A}$ around frontier f' based on the current occupancy map $\mathcal{M}$. $p((x'))$ indicates the Bernoulli occupancy distribution at x'. $|\eta(x, f')|$ is the path length from position x to frontier f'. The frontier with the highest utility is chosen to visit; see Algorithm 2.

$$f^* = \underset{f' \in \mathscr{F}^L}{\text{argmax}} \, u^B(f', x; \mathcal{M}). \quad (10)$$

Lemma 3. *Consider the entropies H^S, H^R and H^B conditioned according to (4), the corresponding utilities u^S, u^R, and u^B. For any given map $\mathcal{M}$, pose x and frontier location f', the range $\mathscr{R}(u^x) \in \mathbb{R}_{>0}$ of possible utilities u^x follows the order $\mathscr{R}(u^B) > \mathscr{R}(u^R) > \mathscr{R}(u^S)$*

Proof. The proof trivially follows from Theorem 3   □

Remark 2. *As utilities calculated from Behavioral entropy H^B have the highest range, chosen frontier f^* from (10) with different parameter choices are also the most diverse. For instance, frontiers with high quality information will be valued higher with high α irrespective of path length, forcing aggressive exploration farther away. On the other end, a low*

α *will treat any uncertainty almost equally, thus focusing exploration in nearby areas.*

The proof of concept (POC) simulations in Sec. V provides more insight. We describe the exploration pipeline (Fig. 1).

Algorithm 2: Behavioral frontier selection

---
1 Input: $\mathcal{M}, \mathscr{F}^L, \alpha, x$;
2 Output : f^*;
3 **for** $f \in \mathscr{F}^L$ **do**
4    $\mathscr{A} \leftarrow f, \mathcal{M}$;   // predicted sensor footprint
5    $\mathscr{W} \leftarrow \mathscr{A}, \alpha$;   // perceived probabilities
6    $H^B \leftarrow \mathscr{W}$;   // Behavioral entropy
7    $I \leftarrow H^B$;   // information gain
8    $|\eta| \leftarrow$ planner(f, x) ;   // path length
9    $u \leftarrow$ utility$(I, |\eta|)$;   // frontier utility
10 **end**
11 $f^* \leftarrow$ argmax$U$;   // most informative location

---

Exploration pipeline: Algorithm 3 describes the behavioral exploration framework in Fig 1. The robot starts with an initial pose x. The observations at x, Y, are used to construct the occupancy map $\mathcal{M}$ using SLAM. Then, the frontier list $\mathscr{F}^L$ is extracted by Algorithm 1 and the best frontier f^* is selected according to Algorithm 2. Then, a circular goal region $\mathbf{B}_r^{x^*}$ is computed, where x^* is the transformed pose of f^* in world frame and r is radius of the region. The navigation manager then provides appropriate global and local plans η to the goal region $\mathbf{B}_r^{x^*}$, to generate a control policy Π that ensures this is reached. SLAM, navigation manager, controller and frontier clustering are continually running in the background to ensure accurate and updated information. The frontier selection process updates the new frontier f^* once the navigation manager either confirms that the goal is reached or aborts when no feasible plans are found. The algorithm runs until there are no more frontiers to explore or the perceived information gain is negligible (≈ 0) from remaining frontiers.

Algorithm 3: Behavioral explorer

---
1 Input: $\mathcal{M}, x, K$;
2 Output : Π;
3 Initialise : $\mathcal{M}, x, r, \mathscr{F}^L$;
4 Initialise : x, r;
5 **while** $\mathscr{F}^L \neq \phi$ & $I^B \neq 0$ **do**
6    $\mathcal{M}, x \leftarrow$ SLAM(Y, Π)
     $\mathscr{F}^L \leftarrow$ Frontier Cluster$(\mathcal{M}, K)$
     $f^* \leftarrow$ Frontier Selection$(\mathcal{M}, \mathscr{F}^L, x, K)$
     $x^* \leftarrow$ Pose transform$(f^*, \mathcal{M})$
     $r \leftarrow$ Goal Radius$(\mathbf{B}_r^p)$
     $\mathbf{B}_r^{x^*} \leftarrow$ Goal region$(r, x^*, \mathcal{M})$
     $\eta \leftarrow$ Navigation Manager$(\mathbf{B}_r^{x^*}, \mathcal{M})$
     $\Pi \leftarrow$ Robot Controller$(\eta, x, \mathcal{M})$
7 **end**

V. PROOF OF CONCEPT SIMULATIONS

To evaluate our proposed approach, we consider first a proof of concept (POC) environment and then focus on real

(a) Environment used for POC simulation.

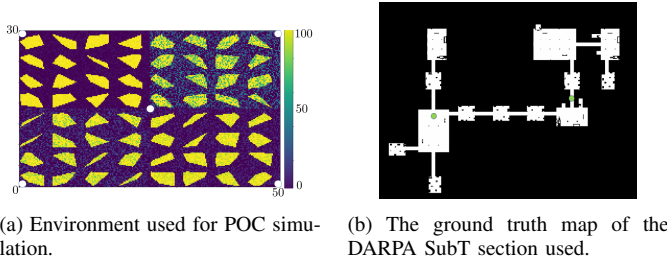(b) The ground truth map of the DARPA SubT section used.

Fig. 3: Environments used for simulations

world scenarios in Sec. VI. Here, we assume all other aspects of exploration: SLAM, navigation and controls are working optimally to focus exclusively on exploration.

Environment Setup: Consider a 2D rectangular environment $\mathscr{X} = [0,30] \times [0,50]$ (Fig 3a) discretized by 0.1 unit square grid cells ($\mathscr{D} = [0,300] \times [0,500]$) containing randomly shaped polygonal obstacles (yellow/bright). The ground truth map $\mathscr{M}^*$ consists of free-space (0 occupancy value) and obstacles (100 occupancy value). The initial occupancy map $\mathscr{M}$ is constructed by adding noise to each cell value by 1) sampling from uniform distribution over each quadrant, 2) adding those to ground truth values of free-space cells (0) and subtracted from obstacle cells (100), see Fig 3a where this is shown. Blue/dark areas (0) represent free space, bright/yellow areas (100) represent obstacles and other values indicate uncertain areas. The environment is divided into four quadrants which contain different levels of noise, sampled from a uniform distribution with different intervals from $[0,5]$ (top-left quadrant), to $[0,50]$ (top-right), and others with $[0,15]$ (bottom-right) and $[0,25]$ (bottom-left). The white circles in the map are the five initial conditions for the robot, $x_i^0, i \in \{1,2,3,4,5\}$. From these, given initial map in Fig 3a, the robot explores the entire region via Algorithm 3 to recover $\mathscr{M}^*$.

POC exploration pipeline: First, the agent updates a circular area $\mathscr{A} = \mathbf{B}_r^x$ in $\mathscr{M}$, with a radius r and centered around its current pose x. Four different mapping radii $\{2,3,4,5\}$ to capture different sensor ranges and three levels of mapping noise $\sigma^m \in \{0,1,2\}$ are employed with the following meaning: *a)* In perfect mapping ($\sigma^m \triangleq 0$) each cell in $\mathscr{A}$ is updated to the ground truth. *b)* In imperfect mapping ($\sigma^m \triangleq 1$ and $\sigma^m \triangleq 2$) the occupancy values in $\mathscr{A}$ are reduced/increased randomly[2], with higher σ^m implying lesser reduction in uncertainty after observation.

Second, frontiers $\mathscr{F}$ are obtained as those cells with occupancy values < 2 and a non-negative gradient w.r.t. to the cells around them. We do not cluster the frontiers here to make sure all possible frontiers are considered.

Then, for each $f_i \in \mathscr{F}$, we calculate the information gain by considering $\mathscr{A} = \mathbf{B}_r^{f_i}$. That is, for each cell in $\mathscr{A}$, the information gain per cell is given by the Behavioral entropy in these cells and the total information gain I is the sum of all of these values. Then, for each f_i, the utility is calculated as the information gain per Euclidean path length required to reach f_i from x. The best frontier f^* is that with the maximum information gain. Finally, the robot moves from x to f^* in a

---

[2]with a random number generated by a uniform distribution with range $[0,35]$ for $\sigma^m \triangleq 1$ and $[0,15]$ for $\sigma^m \triangleq 2$ for each cell in $\mathscr{A}$

(a) Entropy reduced all trials

(b) Area explored all trials

(c) Entropy reduced ($r=2, \sigma^m = 0$)

(d) Entropy reduced ($r=2, \sigma^m = 2$)

(e) Entropy reduced ($r=4, \sigma^m = 0$)
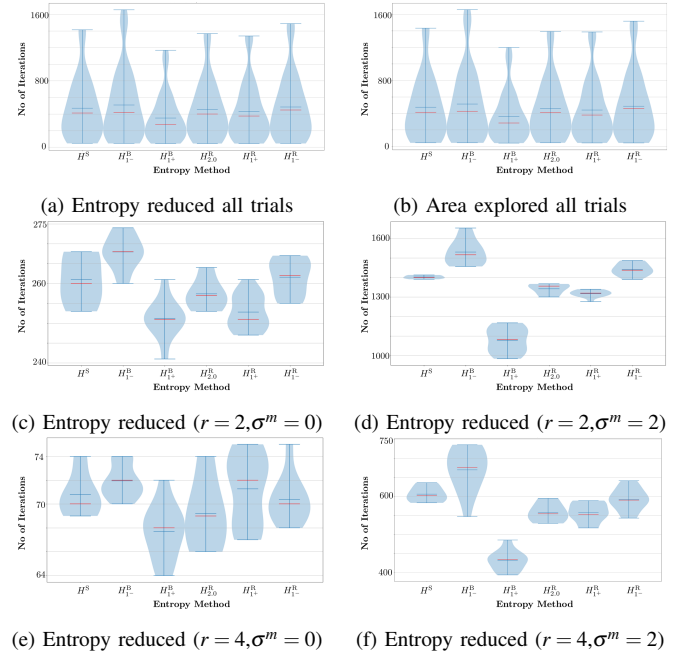
(f) Entropy reduced ($r=4, \sigma^m = 2$)

Fig. 4: Violin plots indicating the number of iterations used to complete 99% of exploration under different behavioral exploration strategies, sensing radius and noise conditions

straight line while mapping (first step) all possible circular regions within its sensor range along the path.

Trials and Metrics: In summary, we have the following control variables: a) sensor radius $r \in \{2,3,4,5\}$, b) mapping noise $\sigma^m \in \{0,1,2\}$, c) starting point $x_i^0, i \in \{1,2,3,4,5\}$, d) entropy method and parameter choice: $\alpha \in \{0.2,0.5,0.8,2.0,3.0,5.0\}$ for Behavioral entropy, $\gamma \in \{0.2,0.5,0.8,2.0,10.0,100.0,1000.0\}$ for Renyi's entropy, and Shannon's entropy. Resulting in 60 trials for each entropy method. For the ease of comparison, we group the behavioral and Renyi parameters into "Uncertainty averse" H_{1-}^B, H_{1-}^R corresponding to $\alpha, \gamma < 1$ and "Uncertainty ignorant" H_{1+}^B, H_{1+}^R corresponding to $\alpha, \gamma > 1$ respectively. We also consider Renyi quadratic entropy (RQE) with $\gamma = 2$, as this is widely used in the literature [17], [25]. We use the ground truth to measure the total uncertainty remaining in $\mathscr{M}$, which is given by Shannon's entropy for a fair comparison. After every iteration of Algorithm 3, we record the entropy remaining and the percentage of completion. We keep the same random number generator for sampling the mapping noise per trial to ensure accountability.

Results and Discussion: The results in Fig 4 show the distribution of the number of iterations taken to explore 99% of the map. The Y-axis indicates the distribution of number of iterations used with a particular group to complete 99% of area/entropy. The largest and smallest number of iterations are indicated by blue horizontal lines limiting the violin shapes, while mean and median are given by horizontal blue and red lines inside the violin, respectively. The wider the violin, the more trials resulted in similar number of iterations, while the narrower the violin, the fewer the trials.

First, we report the performance for each entropy method group across all trials, see Fig 4. We found that the results of area explored and entropy reduced have similar trends as seen

in Fig 4a and Fig 4b. Henceforth, we will report and discuss both metrics as the same. Fig 4a shows that the minimum time for full exploration is similar for all groups, resulting from trials with very good sensor capability ($r = 4, 5$) under perfect mapping $\sigma^m = 0$. This is illustrated in Fig 4e for $r = 4$, we see that across all behaviors, the number of iterations are very similar ranging from $64 - 75$. We also see a similar trend with $r = 4$ and $\sigma^m = 0$, see Fig. 4c. Thus with perfect mapping $\sigma^m = 0$, all behavioral methods seem to provide similar results with marginally better performance from "Uncertainty ignorant" H^B_{1+} behavior. In general, from Fig. 4a and Fig. 4b mean and median iterations are significantly less for "Uncertainty ignorant" Behavioral entropy compared to the other groups. This is because "Uncertainty ignorant" entropies focus on exploring larger areas of uncertainty first as the information gain from almost/partially known area is significantly less than that from highly uncertain or unknown areas. Thereby, reducing uncertainty (Shannon's entropy) much faster than other groups. This is clearly evident under very noisy mapping ($\sigma^m = 2$): it can be seen from Fig. 4d and Fig. 4f, that the "Uncertainty ignorant" group of Behavioral entropy clearly outperforms all other groups in all statistics, resulting into a significantly lower range, mean and median iterations. This is because larger areas of uncertainty are explored first, not waiting for an area to get an almost/complete information gain before moving to the next. However, we observed that, for $\alpha \geq 10$, as the behavioral information gain of the explorer becomes ≈ 0 for almost/partially known areas for the remaining frontiers, exploration terminates. We observed a completion (Shannon's entropy reduced) of around $75 - 90\%$ for $\alpha = 10$. This behavior is well justified looking at the entropy for Bernoulli trials (Fig 2b). As alpha increases, the tail ignorance grows and more likely outcomes also start getting ignored. Thus care needs to be taken to appropriately choose α.

Although Renyi's entropy offers some tail ignorance with $\gamma > 1$, it is not enough to speed up exploration comparing to Behavioral entropy. This is evident from Fig. 2b, where the variation for H^R from $1 \leftarrow \gamma$ (black H^S curve) to $\gamma \to \infty$ (green) is minimal as compared with Behavioral entropy H^B (red).

VI. UNITY-ROS EXPERIMENTS AND RESULTS

Here, we describe results in a ROS-Unity environment from the DARPA subterranean robotic challenge [21].

Environment and Robot setup: The simulated hardware experiments use a Clearpath Warthog robot[3] fitted with Ouster LiDAR and IMU shown in Fig 5 in ROS. Real hardware sensor and actuator models are used including noise sampled from real data. We test our exploration algorithm in the "Urban Challenge" Unity environment, part of the DARPA subterranean robotic challenge, see supplementary videos for scenarios that contains challenges like uneven surfaces, rubble and cluttered environment to make it as close to reality as possible. As we have an UGV that is incapable of climbing stairs, we will use a part of the environment whose ground truth is shown in Fig 3b. The white areas are traversable whereas the black areas are obstacles or walls. We consider

[3]https://clearpathrobotics.com/warthog-unmanned-ground-robot/
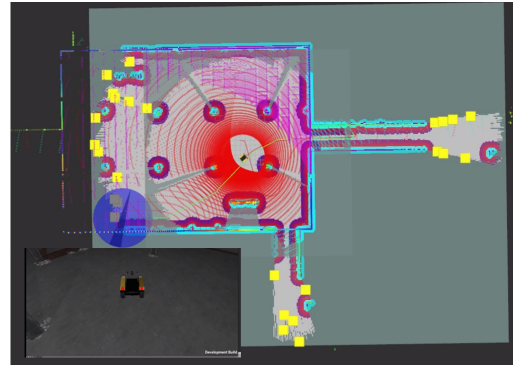

Fig. 5: The unity-ROS environment used in DARPA subterranean challenge. Snapshot of execution shown with map and costmap, frontier clusters (yellow box), laser scans and robot going to region (blue circle), trying to follow a planned path (green line).

two different starting positions (green dots in Fig 3b) for the robot facing towards left of the map and use Omnimapper [8] to map and localize the robot.

Experiment setup: A snapshot of the execution is shown in Fig 5. The Omnimapper's updated occupancy map $\mathcal{M}$ and robot localization x are shown in RViZ in Fig 5 (left). Frontiers f and perceived occupancy $\mathcal{W}$ are extracted from $\mathcal{M}$. The frontiers f are clustered (F) (Yellow boxes in Fig 5) and Behavioral entropy H^B is calculated from $\mathcal{W}$. For each cluster in F, A Behavioral information gain I^B is calculated as described earlier that corresponds to a beam based measurement model as in [6]. Utilities U are calculated from I^B and estimated x as in (9). The explorer then picks a suitable goal region B_r^x (shown as blue disk in Fig 5(right)) and sends it to the navigation manager, which in turn sends global and local plans η to the controller. We use a Generalized Lazy Search (GLS) [26] path planner for global planning, and MPPI [11] for local planning. The framework is visualised in Fig 1.

The explorer algorithm is implemented with a time limit of 30 minutes, and terminates if there are no more frontiers or no perceived information to be gained. We consider 7 different entropy parameter choices: Shannon's entropy H^S, Renyi's entropy H^R with $\gamma \in \{0.5, 2.0\}$, and Behavioral entropy H^B with $\alpha \in \{0.5, 0.8, 2.0, 5.0\}$. We perform 10 trials at each starting position for each entropy or parameter choice and we measure the area explored and entropy reduced every 1 second, and compute percentage completion from the ground truth. As the results for area explored and entropy reduced is similar, here we consider percentage entropy reduction as shown in Fig 6. We again use violin plots to depict the distribution of time taken in seconds for $x\%$, $x \in \{50, 75, 90, 95\}$ completion. Due to inherent uncertainty in the mapping framework, it is very difficult to completely recover the ground truth for every cell of $\mathcal{M}$. Thus we go up to 95% completion in this case. As before, the Y-axis indicates the distribution of time in seconds for a particular group to complete a percentage of area/entropy. The interpretation of these plots is similar to those of Fig. 4.

Results and Discussion: We note that, unlike the POC experiments, here the inherent uncertainty and complexity of mapping, localization, planning, navigation and controls are present in the system. Thus they add significant aleatoric uncertainty and make the trends in the results less apparent. From the trials, the "Uncertainty ignorant" group ($\alpha, \gamma > 1$)

(a) 50 % completion (b) 75 % completion

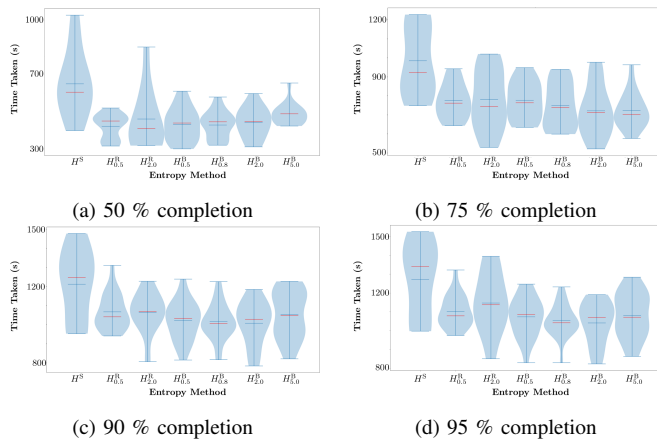(c) 90 % completion (d) 95 % completion

Fig. 6: Experimental results for the DARPA SubT environment.

focuses on exploring larger areas of uncertainty first as the information gain from a partially known area is significantly less. This results in a "Breadth-First-Search" type of exploration. Whereas, the "Uncertainty averse" group ($\alpha, \gamma < 1$) focuses on clearing an area before moving to the next one. Thus, the information gain from a partially known area is still high and the utilities of frontiers in new far away areas are similar to the partially-known nearby frontiers. This group produced a "Depth-First-Search" type of exploration.

The global and local planners performed better (time taken and path quality) for the goal regions that are closer to the current pose. This results in larger wait times, poorer navigation and re-planning for further away goals, thus adversely affecting the "Uncertainty ignorant" group ($\alpha, \gamma > 1$), which tend to produce further away goals than the other group. Interestingly, Shannon's entropy is performing the worst among all the other choices as it suffers from implementation disadvantages of both the "Uncertainty ignorant" and "Uncertainty averse" cases: neither was an area explored completely in depth or breadth causing much longer execution times. Since Renyi's entropy has limited perceptiveness as compared to Behavioral entropy, we see that it is not performing as good as our proposed entropy, following a similar trend from POC experiments. These behaviors are visualized in the accompanying video. Thus, our proposed entropy outperforms the existing entropies used in the literature in both the POC environment and the ROS-Unity environment. Based on the results, we recommend H^{B} parameter α to be in the range of $0.5 - 0.8$ for good exploration conditions or about $2.0 - 3.0$ when there is excessive sensing or environmental noise.

VII. CONCLUSIONS AND FUTURE WORK

This works presents a novel robotic exploration strategy, which relies on human models of perception uncertainty. To do this, we first define a novel human-inspired measure of uncertainty called "Behavioral entropy", which is theoretically analyzed and compared with other common entropy measures like Shannon's and Renyi's. We then used this family of entropies to define utility functions to guide exploration in a frontier-based approach. After this, we illustrate and show that our method is superior in proof of concept simulations as well as a DARPA Subterranean Challenge ROS-Unity

simulation environment with a Clearpath Husky robot. Future work will study the benefits of the spatial/temporal variation of α and conduct real world experiments. We would also like to investigate other applications of information-adaptive path planning algorithms using this entropy class.

REFERENCES

[1] C. E. Shannon, "A mathematical theory of communication," *The Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.
[2] D.Prelec, "The probability weighing function," *Econometrica*, vol. 66, no. 3, pp. 497–527, 1998.
[3] A. Tversky and D. Kahneman, "Advances in Prospect theory: Cumulative representation of uncertainty," *Journal of Risk and Uncertainty*, vol. 5, no. 4, pp. 297–323, 1992.
[4] A. Suresh, A. Taylor, L. Riek, and S. Martínez, "Robot Navigation in Risky, Crowded Environments: Understanding Human Preferences," *IEEE Robotics and Automation Letters*, vol. 8, no. 9, pp. 5632–5639, 2023.
[5] D. R. Bach and R. J. Dolan, "Knowing how much you don't know: a neural organization of uncertainty estimates," *Nature reviews neuroscience*, vol. 13, no. 8, pp. 572–586, 2012.
[6] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*, ser. Intelligent Robotics and Autonomous Agents. The MIT Press, 2005.
[7] Y. Tian, Y. Chang, F. H. Arias, C. Nieto-Granda, J. P. How, and L. Carlone, "Kimera-multi: Robust, distributed, dense metric-semantic slam for multi-robot systems," *IEEE Transactions on Robotics*, vol. 38, no. 4, 2022.
[8] C. Nieto-Granda, I. John G. Rogers, and H. I. Christensen, "Coordination strategies for multi-robot exploration and mapping," *The International Journal of Robotics Research*, vol. 33, no. 4, pp. 519–533, 2014.
[9] A. Suresh and S. Martínez, "Planning under non-rational perception of uncertain spatial costs," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 4133–4140, 2021.
[10] A. Suresh, C. Nieto, and S. Martínez, "Purr-rrt*: Behavioral path planning in uncertain, risky and rewarding environments," in *2023 20th International Conference on Ubiquitous Robots (UR)*. IEEE, 2023, pp. 211–216.
[11] G. Williams, N. Wagener, B. Goldfain, P. Drews, J. M. Rehg, B. Boots, and E. A. Theodorou, "Information theoretic mpc for model-based reinforcement learning," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 1714–1721.
[12] S. Koga, A. Asgharivaskasi, and N. Atanasov, "Active exploration and mapping via iterative covariance regulation over continuous se (3) trajectories," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2021, pp. 2735–2741.
[13] A. Suresh and S. Martínez, "Risk-Perception-Aware Control Design under Dynamic Spatial Risks," *IEEE Control Systems Letters*, vol. 6, pp. 1802 – 1807, 2022.
[14] B. Yamauchi, "Frontier-based exploration using multiple robots," in *Proceedings of the second international conference on Autonomous agents*, 1998, pp. 47–53.
[15] Z. Sun, B. Wu, C.-Z. Xu, S. E. Sarma, J. Yang, and H. Kong, "Frontier detection and reachability analysis for efficient 2d graph-slam based active exploration," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 2051–2058.
[16] K. Otsu, S. Tepsuporn, R. Thakker, T. S. Vaquero, J. A. Edlund, W. Walsh, G. Miles, T. Heywood, M. T. Wolf, and A.-A. Agha-Mohammadi, "Supervised autonomy for communication-degraded subterranean exploration by a robot team," in *2020 IEEE Aerospace Conference*. IEEE, 2020, pp. 1–9.
[17] H. Carrillo, P. Dames, V. Kumar, and J. A. Castellanos, "Autonomous robotic exploration using a utility function based on rényi's general theory of entropy," *Autonomous Robots*, vol. 42, no. 2, pp. 235–256, 2018.
[18] M. Ribeiro, T. Henriques, L. Castro, A. Souto, L. Antunes, C. Costa-Santos, and A. Teixeira, "The entropy universe," *Entropy*, vol. 23, no. 2, p. 222, 2021.
[19] T. Osogami and M. Otsuka, "Restricted boltzmann machines modeling human choice," *Advances in Neural Information Processing Systems*, vol. 27, 2014.
[20] Ádám Koblinger, J. Fiser, and M. Lengyel, "Representations of uncertainty: where art thou?" *Current Opinion in Behavioral Sciences*, vol. 38, pp. 150–162, 2021.

[21] J. G. Rogers, A. Schang, C. Nieto-Granda, J. Ware, J. Carter, J. Fink, and E. Stump, "The DARPA SubT Urban Circuit Mapping Dataset and Evaluation Metric," in *International Symposium on Experimental Robotics (ISER) 2020*. IFRR, 2021, pp. 391–401.

[22] J. M. Amigó, S. G. Balogh, and S. Hernández, "A brief review of generalized entropies," *Entropy*, vol. 20, no. 11, 2018.

[23] S. Dhami, *The Foundations of Behavioral Economic Analysis*. Oxford University press, 2016.

[24] P. Stein, "A note on the volume of a simplex," *The American Mathematical Monthly*, vol. 73, no. 3, pp. 299–301, 1966.

[25] B. Charrow, S. Liu, V. Kumar, and N. Michael, "Information-theoretic mapping using cauchy-schwarz quadratic mutual information," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2015, pp. 4791–4798.

[26] A. Mandalika, S. Choudhury, O. Salzman, and S. Srinivasa, "Generalized lazy search for robot motion planning: Interleaving search and edge evaluation via event-based toggles," in *Proceedings of the International Conference on Automated Planning and Scheduling*, vol. 29, 2019, pp. 745–753.