 http://dx.doi.org/10.1109/LRA.2021.3068935

Optimizing Non-Markovian Information Gain under Physics-based Communication Constraints*

Matthew A. Schack†, John G. Rogers‡, Qi Han†, and Neil T. Dantam†

***Abstract*—In many exploration scenarios, it is important for robots to efficiently explore new areas and constantly communicate results. Mobile robots inherently couple motion and network topology due to the effects of position on wireless propagation—e.g. distance or obstacles between network nodes. Information gain is a useful measure of exploration. However, finding paths that maximize information gain while preserving communication is challenging due to the non-Markovian nature of information gain, discontinuities in network topology, and zero-reward local optima. We address these challenges through an optimization and sampling-based algorithm. Our algorithm scales to 50% more robots and obtains 2-5 times more information relative to path cost compared to baseline planning approaches.**

I. Introduction

Many robot tasks combine the need for motion and communication. Inspection, exploration, and search and rescue tasks all require robots to move, observe, and maintain constant communication both between robots and with a human operator. For instance, robots may need to stream sensor data to human operators in search and rescue scenarios or constantly receive instructions from humans to inspect a potential hazard [1]. Furthermore, many scenarios—especially search and rescue or exploration—present challenges for communication due to distance and occlusion between robots. To accomplish tasks with communication requirements, the robots must arrange themselves to both maintain communication and reach necessary positions.

In exploration scenarios, we often want to explore a large area in a limited time. We quantify exploration using *information gain* [2], [3], [4], [5], a metric for how much unknown space is revealed. However, the information gain at a particular configuration in a path inherently depends on the history of configurations. That is, a robot observing an initially unknown feature would gain more information from the first observation than from subsequent observations. This dependence on the history of configurations means information gain is *non-Markovian* when the state is only the current robot configuration. To use information gain as a metric, we must either encode the history of robot configurations in the state or consider an entire path when finding an optimal answer. While adding extra information to each configuration would make the problem Markovian, it would also drastically increase the dimensionality of the problem. Furthermore, a

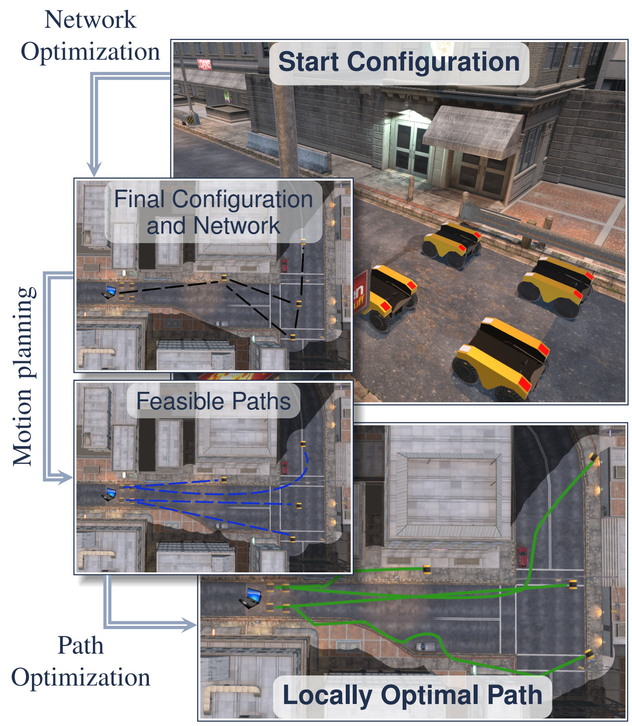


Fig. 1: A robot team exploring an urban area. Unknown regions are shown in black. Our algorithm optimizes information gain while maintaining communication between the team and base station.

configuration in which robots observe fully known space has a local information gain of zero and a local gradient of information gain of zero; such configurations are local optima with zero-reward, confounding the direct use of gradient-based methods. Finally, communication requirements limit the valid configuration space, requiring a mapping from robot configurations to underlying network topologies.

We present an optimization and sampling-based approach for a team of robots to find locally optimal paths with a non-Markovian objective and maintain network connectivity under a physics-based radio communication model (Fig. 1). First, we identify a heuristic final configuration with both positive information gain and all robots in communication via optimization, addressing the zero-reward local optima (Sec. IV-A). Second, we find feasible paths that maintain communication and reach the final configuration via sampling-based planning (Sec. IV-B). Third, we optimize the paths to maximize information gained relative to a cost, addressing the non-Markovian objective (Sec. IV-C). We demonstrate the approach using a physics-based communication model

* This work was supported by ARL DCIST W911NF-17-2-0181 P00007 and NSF CNS-1823245.

† Department of Computer Science, Colorado School of Mines, USA. {mschack,qhan,ndantam}@mines.edu.

‡ United States Army Research Laboratory, Adelphi, MD, USA. john.g.rogers59.civ@mail.mil

(Sec. V). Finally, our evaluation shows improved scalability to many robots and improved efficiency compared to baseline sampling-based methods (Sec. VI).

II. Related Work

Our work combines communication constrained information gain, multi-robot exploration, and motion planning.

Information gain has been used as an accurate metric for exploration problems with a single robot. Prior approaches to maximize information gain either quickly find and move towards a single configuration with high information gain [6], [7] or use optimization [2] to find a locally optimal path. Our work expands the single robot case to scale to large, communicating teams while maintaining locally optimal paths.

Active Information Acquisition problems require a team of robots to learn some feature of the environment—e.g., location of unknown markers or a model for gas concentration. Leading approaches [8], [9] focus on solving this problem by creating a state composed of a robot configuration and extra information to make the problem Markovian, allowing them to consider just the previous state in the algorithm. When the additional information encoded in the state is small, reducing the problem to a Markovian system allows for these algorithms to sample or iterate over many states efficiently. However, for exploration problems, even simple abstractions of the space—e.g., Booleans for whether or not an unknown point is visited—causes *exponential* growth in the state space and the number of possible states that can be associated with a single robot configuration, since every additional unknown point doubles the possible permutations of visited unknown points. Our work instead uses a state comprised of solely robot configurations to avoid this growth in state size, though we must instead address a non-Markovian problem.

Multi-robot exploration must balance scalability to many robots, efficient management of the robots, and potentially communication requirements. Leading approaches that neglect communication have used decentralized control [10], [11], [12] and information gain reward [3], [4], [5] to achieve scalable and efficient algorithms at the cost of decreased connectivity. Algorithms that define static network requirements (i.e., direct communication links between robots are defined initially and cannot be broken) [13] allow for scalability while maintaining communication, but may not achieve every objective to explore or observe, or it may only do so at a higher cost (e.g., distance traveled). Algorithms that permit a dynamic network must instead maintain total graph connectivity [14], [15], [16] and may accomplish more tasks at a lower cost than the static case. All of these approaches typically address one or two of the main problems associated with multi-robot exploration. Our work uses of optimization to scale to large numbers of robots and ensure communication, while using information gain to efficiently prioritize areas to explore, allowing us to address all three requirements of scalability, efficiency, and communication.

High-dimensional motion planning must balance optimality and completeness with scalability. Heuristic search methods—e.g., A* [17], AD* [18], or ARA* [19]—provide optimal results but are not as widely used for high-dimensional problems as alternatives. Sampling-based methods scale well to high-dimensional problems, but only offer probabilistic completeness [20] or asymptotic optimality [21]. Conversely, optimization-based approaches such as CHOMP [22], STOMP [23], and TrajOpt [24] are often highly-efficient but typically guarantee only local optimality rather than the convergence, completeness, or global optimality guarantees of other methods. We integrate a sampling-based approach [25] and an optimization similar to [24] to maximize information gain and maintain communication constraints.

Communication-constrained motion planning combines motion planning with a requirement to maintain network connectivity. Prior work has applied discrete search [26], [27], decentralized control [28], [29], and sampling-based approaches [30], [31]. Such works typically construct paths to a specific goal position; it is possible to find positions with a high amount of information gain [6], [7] and then plan paths there. However, we show that our formulation—optimizing information gain along the entire path—results in more information gained relative to the amount of time to compute the path. Furthermore, prior work on communication-constrained motion planning typically used range-based communication models, whereas we plan using a more accurate physics-based model of radio propagation.

III. Problem Definition

We address a *multi-robot communication and exploration* problem, where a robot team must maintain communication among the team and with a base station, avoid physical collisions, and maximize information gain relative to a cost. We define the problem in terms of deterministic functions for network connectivity, collision distance, information gain, and cost as follows.

Definition 1. *A multi-robot communication and exploration problem is a* $\Sigma = \left(\mathcal{I}, \mathcal{X}, x^{[0]}, b, \mathcal{F}_f, \mathcal{F}_c, \mathcal{F}_g, \mathcal{C}\right)$ *where,*

- $\mathcal{I}$ *is a finite set of robots*
- $\mathcal{X} = \mathfrak{m} \times \mathfrak{m} \cdots \times \mathfrak{m}$ *is the multi-robot configuration space, where each* $\mathfrak{m}$ *is the position of a single robot—e.g.* $\mathcal{SE}(2)$, $\mathcal{SE}(3)$, *or* $\mathbb{R}^n$.
- $x^{[0]} \in \mathcal{X}$ *is the initial configuration of all the robots.*
- $b \in \mathfrak{m}$ *is the position of the base station.*
- $\mathcal{F}_f : \mathfrak{m} \mapsto \mathbb{R}$ *is a signed distance function indicating separation from the closest obstacle. Negative distances indicate penetration (collision) with an obstacle.*
- $\mathcal{F}_c : \mathfrak{m} \times \mathfrak{m} \mapsto \mathbb{R}$ *is a signed communication function. Positive* $\mathcal{F}_c$ *indicates that communication is possible between the two positions.*
- $\mathcal{F}_g : ([0,1] \mapsto \mathcal{X}) \mapsto \mathbb{R}^+$ *is a function that maps from a path in* $\mathcal{X}$ *to the information gain over that path.*
- $\mathcal{C} : ([0,1] \mapsto \mathcal{X}) \mapsto \mathbb{R}^+$ *is a function that maps from a path in* $\mathcal{X}$ *to a cumulative cost—e.g., distance, time, or energy.*

The configuration space $\mathcal{X}$ is the union of disjoint valid space $\mathcal{X}_{\text{valid}}$ and invalid space $\mathcal{X}_{\text{invalid}}$. The valid space $\mathcal{X}_{\text{valid}}$ is the region where there are no collisions ($\mathcal{F}_f$ is positive for every robot) and where the base station can communicate with every robot through one or more network hops. A network hop is possible when the communication function $\mathcal{F}_c$ for the two nodes is positive.

The solution to Definition 1 is a path $\boldsymbol{\sigma}$ that is feasible and maximizes information gain relative to the cost.

Definition 2. *An information-optimal path σ solves,*

$$\max_{\boldsymbol{\sigma}} \quad \frac{\mathcal{F}_g(\boldsymbol{\sigma})}{\mathcal{C}(\boldsymbol{\sigma})}$$
$$\textit{s.t.} \quad \boldsymbol{\sigma}[0,1] \in \mathcal{X}_{\text{valid}} \wedge \boldsymbol{\sigma}(0) = x^{[0]} \ .$$

Additionally, we assume that most of the information gain will come from the final points on the path, which is appropriate when the robots are exploring previously unexplored areas. Other works have achieved promising results by considering information gain only at the final point [6], [7], indicating that this assumption is reasonable.

IV. Algorithm

We solve the multi-robot communication and exploration problem using a hybrid of sampling and optimization to overcome the non-Markovian objective and zero-reward optima problems while scaling to large teams. The algorithm (see Alg. 1) follows three steps. First, we address the zero-reward local optima by finding a heuristic final configuration with positive information gain (see Sec. IV-A). Second, we use the heuristic configuration as a goal to find a feasible path via bidirectional, sampling-based planning (see Sec. IV-B). Third, we overcome the non-Markovian objective by optimizing the feasible path according to Definition 2 (see Sec. IV-C).

Algorithm 1: Information Gain Optimization

Input: $b, \mathcal{F}_f, \mathcal{F}_c, \mathcal{F}_g, \mathcal{C}, x^{[0]}$; `// Definition 1`
Output: $\boldsymbol{\sigma}_{\text{opt}}$; `// Definition 2`
1 **function** `isValid`(x) **is**
2   **foreach** $x_i \in x$ **do** `// Physical Collision`
3   **if** $\mathcal{F}_f(x_i) \leq 0$ **then return** False;
  `// Sec. IV-A2: Comms. constraint`
4   **if** `calculateComCost` $(b, \mathcal{F}_c, x) \neq 0$ **then**
5   **return** False;
6   **return** True;
7 $\mathcal{X}_{\text{valid}} \leftarrow \{x \in \mathcal{X} | \texttt{isValid}(x)\}$;
`// Sec. IV-A: Heuristic config`
8 **repeat** `// find non-zero info-gain`
9   $x, r \leftarrow$ `goalConfiguration` $(\mathcal{F}_g, \mathcal{X}_{\text{valid}})$;
10 **until** $r \neq 0$;
`// Sec. IV-B: Feasible path`
11 $\boldsymbol{\sigma}_{\text{feas}} \leftarrow$ `sampleBasedPlanning` $(x^{[0]}, x, \mathcal{X}_{\text{valid}})$;
`// Sec. IV-C: Optimize info-gain`
12 $\boldsymbol{\sigma}_{\text{opt}} \leftarrow$ `trajectoryOpt` $(\boldsymbol{\sigma}_{\text{feas}}, \mathcal{F}_g, \mathcal{C}, \mathcal{X}_{\text{valid}})$;
13 **return** $\boldsymbol{\sigma}_{\text{opt}}$

A. Heuristic Final Configuration

We first evaluate a heuristic to maximize information gain at a single configuration (line 8). The heuristic configuration

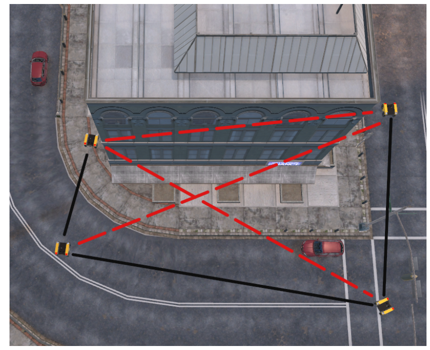


Fig. 2: Robots team and corresponding, valid, communication graph. Communication is possible along zero-weight edges (solid lines). Communication is not possible along nonzero-weight edges (dashed lines).

is the solution to the following,

$$\begin{aligned} \max_{x \in \mathcal{X}} \quad & \mathcal{F}_g(x) \\ \text{s.t.} \quad & x \in \mathcal{X}_{\text{valid}} \ . \end{aligned} \tag{1}$$

Previously explored areas may create zero-reward local optima for the heuristic configuration, so we ensure that information gain is non-zero. Since a single configuration is lower-dimensional than an entire path, it is faster to identify the zero-reward local optima and restart the optimization here compared to optimizing the entire path.

We define two constraints to ensure that the configuration is in $\mathcal{X}_{\text{valid}}$. First, the configuration must be free from physical collisions. Second, all robots and the base station must be able to communicate at the configuration.

1) Collision-Free Constraint: The robots must maintain positive clearance from every obstacle.

$$\mathcal{F}_f(x_i) \geq 0 \ \forall i \in \mathcal{I} \ , \tag{2}$$

where x_i is the robot i's position. For a position defined in a Cartesian configuration space, the derivative of $\mathcal{F}_f$ for robot i is the derivative of Euclidean distance between i and the closest object.

2) Communication Constraints: The base station must be able to communicate with every robot via one or more network hops. The communication function, $\mathcal{F}_c$, implicitly represents the configurations with valid communication; we model typical radio communication in Sec. V. We define the constraints in terms of $\mathcal{F}_c$ by constructing a strongly-connected weighted graph that represents the possible communication network topology (see Fig. 2). The graph nodes are the robots and base station. Each edge weight w_{ij} represents the communication strength between two nodes, according to communication function $\mathcal{F}_c$. If communication is impossible, w_{ij} is negative value denoting the degree to which communication is impossible. Otherwise, w_{ij} is 0.

$$w_{ij} = \begin{cases} \mathcal{F}_c(x_i, x_j), & \text{if } \mathcal{F}_c(x_i, x_j) \leq 0 \\ 0, & \text{if } \mathcal{F}_c(x_i, x_j) > 0 \end{cases} \tag{3}$$

where x_i is the position of robot i.

The optimization constraint is that the maximum cost network route from the base station to every node must equal zero, indicating that communication is possible from the base station to any robot. We find this route using a graph search to calculate the cost to go from the base station to any node. For a constant network route, the gradient is the sum of the gradients for each edge weight according to (3). Changes in the route create discontinuities in the gradient; we empirically show that the optimization formulation is robust to this nonlinearity (see Sec. VI).

B. Initial Feasible Path

Next, we find a feasible path from the current configuration $x^{[0]}$ to the heuristic final configuration found in Sec. IV-A. The valid space for this path consists of configurations where robots avoid collisions and can communicate; these requirements are the same those of Sec. IV-A defined by collision (2) and communication (3) constraints. The final configuration from Sec. IV-A allows for bidirectional search.

In this step, we find a feasible (rather than optimal) path because our information gain objective is non-Markovian and contains many local optima with a zero reward. Optimizing sampling-based methods (e.g., RRT* [21]) do not directly address non-Markovian objectives such as information gain over the path. Path optimization does not robustly address the zero-reward local optima. In such scenarios, we found that path optimization beginning from an invalid seed will typically move the final configuration to a valid—but zero-reward—configuration, resulting in a locally optimal—but zero-reward—path. Instead, we use the feasible path from motion planning in this step as an initially valid seed for the path optimization in Sec. IV-C, ensuring that path optimization will be at least as good as the feasible path.

C. Path Optimization

Finally, we optimize the feasible path from Sec. IV-B to maximize information gain. We perform path optimization over a set of waypoints and we ensure that the segment between each waypoint is in the valid space. The optimal path from Definition 2, restated over discrete waypoints, is as follows,

$$\begin{aligned} \max_{x^{[0]},\dots x^{[K]}} \quad & \frac{\mathcal{F}_g\left(\left[x^{[0]},x^{[1]}\right],\left[x^{[1]},x^{[2]}\right],\dots\left[x^{[K-1]},x^{[K]}\right]\right)}{\mathcal{C}\left(\left[x^{[0]},x^{[1]}\right],\left[x^{[1]},x^{[2]}\right],\dots\left[x^{[K-1]},x^{[K]}\right]\right)} \\ \text{s.t.} \quad & \left[x^{[k]},x^{[k+1]}\right] \in \mathcal{X}_{\text{valid}} \quad \forall k \le K-1 \ , \end{aligned} \tag{4}$$

where $\left[x^{[k]},x^{[k+1]}\right]$ is a lineally interpolated segment between step k and $k+1$. We ensure any segment is in valid space by checking collisions using the same continuous collision checking constraint as [24] and check communication by discretizing the segment and ensuring that the communication constraint (3) is met for each discrete configuration.

Once we find an optimal path, the robots explore according to the path. Then, using the map updates from exploration, the algorithm restarts to find the next optimal path.

V. Communication model

While our problem formulation (Definition 1) and algorithm (Sec. IV) are independent of a particular communication function $\mathcal{F}_c$, we consider specifically radio signal power loss for the experiments in Sec. VI. Many other communication models have been developed [32], [33], [34], [35], [36], [37] and also would apply to our framework. Power loss in particular is an important metric based on the underlying physics of radio propagation [32], [34]. Moreover, power loss presents important properties—such as discontinuities—that pose challenges for optimization.

Power loss between a source and a receiving antenna may occur for three reasons: distance (pathloss), occluding objects (shadowing), and reflection in the signal (multi-path fading) [32], [34]. We combine these three effects to find received power, P_r, in dB.

$$P_r = P_{r0} - 10n\log_{10}\left(||x_i - x_j||\right) - v_s - \epsilon \tag{5}$$

P_{r0} is the received power at a standard distance (i.e., 1m). The term $-10n\log_{10}\left(||x_i - x_j||\right)$ represents pathloss, where n is the pathloss exponent and $||x_i - x_j||$ is the distance between the robots. The effects of shadowing, v_s, are discontinuous and change when more or fewer objects occlude the straight line path between two antennas. Term ϵ represents the effects of multi-path fading. Prior communication modeling work estimates parameters P_{r0}, n, v_s, and ϵ online [32], [34] or uses standardized static values [38].

We convert (5) to the form of communication function $\mathcal{F}_c$ for static parameters, P_{r0}, n, and ϵ and for shadowing v_s as a function of the two nodes' positions. Our experiments use standardized values for P_{r0}, n, and ϵ from [38]. To successfully communicate, received power P_r must exceed some minimum strength s [37], which varies based on the receiver's properties (e.g., the antenna). The resulting $\mathcal{F}_c$ is,

$$\begin{aligned} &\mathcal{F}_c\left(x_i,x_j\right) = \\ &\quad \left(P_{r0} - 10n\log_{10}\left(||x_i - x_j||\right) - v_s\left(x_i,x_j\right) - \epsilon\right) - s\ . \end{aligned} \tag{6}$$

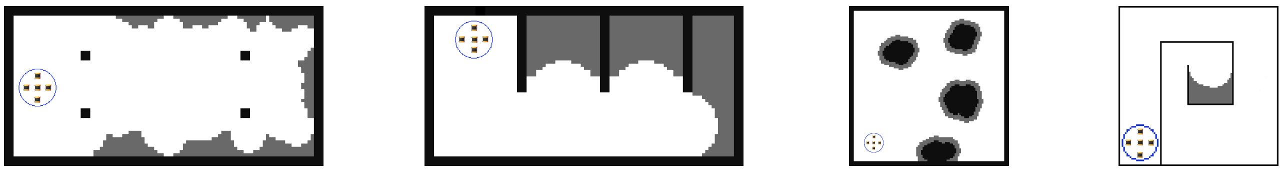
We use a conservative model of shadowing v_s that limits communication to Line of Sight (LoS). If LoS is not present between positions x_i and x_j, then $v_s\left(x_i,x_j\right)$ has a large value, M, thus preventing communication.

$$v_s\left(x_i,x_j\right) = \begin{cases} 0, & \text{If } \texttt{LoS}\left(x_i,x_j\right) \\ M, & \text{If } \neg\texttt{LoS}\left(x_i,x_j\right) \end{cases} \tag{7}$$

We calculate the gradient of (7) analytically disregarding the discontinuity in (7) as we move into and out of LoS. We show empirically that we are robust to this discontinuity.

VI. Evaluation

We evaluate the performance of our algorithm for the environments in Fig. 3. Typical exploration scenarios will update plans as the map is updated; we evaluate specifically the planning process in terms of scalability, optimality, and efficiency. We compare planning efficiency in terms of the reward for a path divided by the time to compute the path—i.e., the benefit achieved per second of computation time or *Reward per Computational Second* (R/CS). The experimental environments vary in distribution of unknown space and difficulty of communication. Unknown space lies in discrete areas in Fig. 3c and Fig. 3d and is widely distributed in Fig. 3a and Fig. 3b. Communication is easy to maintain in Fig. 3a,



(a) A simple room with pillars (b) A hallway with many corridors (c) Hotspot monitoring (d) Spiral hallway

Fig. 3: Maps of the simulated environments used in our experiments. All robots start within the blue circle with the base station in the center of the circle. Gray points are unknown space.

	Our work	RRT	RRT*	Our work	RRT	RRT*
Simple room (Fig. 3a)		5 robots			10 robots	
Total time (stdev)	10.0 (6.1)	**1.72 (3.2)**	60.4 (0.2)	52.2 (19.1)	**7.44 (13.0)**	60.6 (0.4)
Reward (stdev)	**1.12 (0.74)**	0.12 (0.07)	0.38 (0.04)	**0.76 (0.48)**	0.06 (0.03)	0.21 (0.02)
R/CS (stdev)	**0.159 (0.148)**	0.12 (0.08)	0.006 (0.001)	**0.018 (0.015)**	0.012 (0.008)	0.004 (0.000)
Hallway (Fig. 3b)		5 robots			10 robots	
Total time (stdev)	25.7 (16.7)	19.9 (17.2)	**19.8 (7.7)**	**58.0 (9.1)**	60.1 (0.04)	60.1 (0.04)
Reward (stdev)	**0.64 (0.54)**	0.10 (0.06)	0.40 (0.10)	**0.24 (0.92)**	0 (0.00)	0 (0.00)
R/CS (stdev)	**0.036 (0.044)**	0.007 (0.004)	0.023 (0.012)	**0.004 (0.015)**	0 (0.000)	0 (0.000)
Hotspot (Fig. 3c)		5 robots			10 robots	
Total time (stdev)	5.6 (5.0)	**2.3 (2.8)**	10.5 (2.65)	38.6 (20.3)	**13.7 (20.17)**	39.1 (13.54)
Reward (stdev)	**0.76 (0.57)**	0.06 (0.03)	0.41 (0.08)	**0.80 (0.51)**	0.05 (0.03)	0.23 (0.03)
R/CS (stdev)	**0.186 (0.193)**	0.051 (0.44)	0.042 (0.015)	**0.027 (0.030)**	0.009 (0.006)	0.006 (0.002)

TABLE I: Experimental results. Scores are the average of 30 trials, with standard deviation in parenthesis. We evaluate efficiency as Reward per Computational Second (R/CS). All methods did not find a non-zero reward on Fig. 3d before the timeout, and so it is not included in the table.

while it is difficult in Fig. 3b and Fig. 3d due to shadowing from walls.

A. Baseline Methods

We compare our approach to baselines which apply communication-aware, sampling-based motion planning from [30], [39] using frontier positions containing unknown space [6] to define a goal state. [30], [39], use unidirectional, sampling-based motion planning to find a communication-aware path for a group of robots to a user-specified goal; we define a state as a valid goal for our baselines using convex hulls of unknown space from [6]. The baselines use RRT [25] and RRT* [21] to plan paths that preserve communication [30], [39]; we found that direct use of optimization-based methods to reach frontiers did not robustly find valid paths with non-zero information gain. Importantly, the frontiers are *positions* for individual robots rather than valid *configurations* of the team that maintain communication. Since the baselines do not provide a goal configuration, we cannot directly apply bidirectional search. As in [30], [39], the baseline RRT and RRT* are unidirectional, but we bias search towards configurations with a robot at the frontier. We assume the final configuration provides most of the information, so we expect shortening to improve reward (information gain over distance), which holds in our experiments.

The non-Markovian nature of information gain poses a challenge for RRT*. Classically, RRT* optimizes the accumulated, per-configuration objective over a path; however, information gain depends on the path as a whole. RRT* can also optimize an objective for a specific configuration—e.g., maximizing the minimum path clearance (an example in the OMPL tutorial). Thus, our objective for RRT* is to maximize information gain at the final configuration in the path.

We apply the baseline methods to maximize efficiency in terms of reward over computation time, R/CS. Neither RRT or RRT* fully address the non-Markovian information gain objective. Instead, we greedily terminate the search after achieving some level of information gain. The resulting plans are not optimal, but they are fast to find, which improves efficiency (R/CS). We terminate the RRT once we find a configuration with positive information gain. We terminate RRT* after reaching a configuration that exceeds a ratio of the sum of information gain from all frontiers; if RRT* times-out, we return the best plan it found. The specific termination ratio affects the trade-off between reward and computation. A higher ratio will take longer, but will find a better plan (or timeout). A lower ratio will approach the performance of RRT. In the experiments, we use a ratio of 0.5.

We use the implementations of RRT and RRT* in OMPL [40], set a timeout of 60 seconds, an extension distance of one meter, and use the frontier positions from [6] as a heuristic. If either planner returns before the timeout, we spend up to the remaining time shortening the path.

B. Experiments

In the experiments, the implementation of our algorithm uses Sequential Least-Squares Quadratic Programming (SLSQP) [41], [42], [43] to solve for the heuristic final configuration, RRT-connect [25] in OMPL [40] to find feasible paths, and gradient descent to optimize the path. Our algorithm was given a timeout of 60 seconds. RRT-connect used an extension distance of one meter and simplified the path if it returned before the timeout.

Both our algorithm and the baseline methods use same communication model (see Sec. V) and objective, though the baselines cannot fully address the non-Markovian objectives. We use static communication parameters from the ITU model [38]: power received at standard distance $P_{r0} = 39.6$db, pathloss exponent $n = 3$, and the effects of multipath fading are constant and included in P_{r0}. We conservatively model shadowing to limit communication to LoS, with the penalty for no LoS $M = 100$. We assume a minimum communication strength $s = 10$db. We model collisions and check LoS with FCL [44]. We evaluate information gain, $\mathcal{F}_g$, based on observations from independent beams [2] with a Euclidean distance path cost $\mathcal{C}$. We performed the experiments on an Intel Xeon CPU at 3.40GHz.

Table I shows the average information gain, runtime, and efficiency over 30 trials for each environment for both five and ten robots. While our method is typically slower than the baselines, it gains more reward. In terms of R/CS, our method is more efficient than the baselines.

All methods have a high variance due to the random sampling during planning, and, in our method, random seeds to find a final configuration. In exploration scenarios, we would typically plan many times as we update the map. While there would be variance in individual planning times, we expect the average efficiency over the entire exploration scenario to approach the mean as in our results.

Both RRT and RRT* were unable to find a path with 10 robots in the hallway due to the high-dimensional configuration space and the limitation to unidirectional search. A key part of our algorithm is finding locally optimal final configurations (see Sec. IV-A), enabling bidirectional search.

Our work uses both optimization and sampling and is thus subject to the general limitations of both methods—i.e., non-convexity in optimization and narrow passages in sampling, though our approach does address the small amounts of non-convexity and narrow passages present in the scenario in Fig. 3b better than the baselines. Non-convexity during optimization can arise from communication shadowing effects, as even convex obstacles in Cartesian space may cause non-convexity in the configuration space. By restarting the optimization process when it fails to find a nonzero reward configuration, we can empirically still find a goal configuration in environments with a small amount of non-convexity—e.g., Fig. 3b. However, when there is a large amount of non-convex areas—e.g. Fig. 3d—it takes many restarts to find a nonzero reward configuration, slowing down the approach. Additionally, narrow areas by themselves pose a challenge for the sampling-based methods as they may result in narrow passageways in the configuration space. For the small amount of non-convexity and narrow passages in Fig. 3b, our method was able to find greater reward through optimization and bidirectional sampling compared to the unidirectionally sampling baselines. However, when non-convexity is greater and passages are narrower—e.g., Fig. 3d—our method, along with the baselines, fails to find a path to a non-zero reward configuration.

VII. Conclusion

We have presented an optimization and sampling-based approach to find paths that maximize information gain relative to a cost. Information gain over a path is both non-Markovian, due to dependence on previous states, and has many zero-reward local optima. Our approach finds locally optimal paths by leveraging efficient methods to find heuristic configurations and feasible paths to use as a seed for path optimization. This formulation is general to a variety of communication, collision, cost, and information gain models. Our experimental results show, for the tested domains, that our approach has a higher reward per second spent calculating than baseline sampling-based methods.

In future work, we evaluate the approach with physical robots and radios, and we will extend the work to allow robots to disconnect from the network to reach previously unreachable areas.

References

[1] K. Nagatani, S. Kiribayashi, Y. Okada, K. Otake, K. Yoshida, S. Tadokoro, T. Nishimura, T. Yoshida, E. Koyanagi, M. Fukushima *et al.*, "Emergency response to the nuclear accident at the Fukushima Daiichi nuclear power plants using mobile rescue robots," *J. of Field Robotics*, vol. 30, no. 1, pp. 44–63, 2013.

[2] B. Charrow, G. Kahn, S. Patil, S. Liu, K. Goldberg, P. Abbeel, N. Michael, and V. Kumar, "Information-theoretic planning with trajectory optimization for dense 3D mapping."

[3] R. Zlot, A. Stentz, M. B. Dias, and S. Thayer, "Multi-robot exploration controlled by a market economy," in *ICRA*, vol. 3. IEEE, 2002.

[4] M. Corah, C. O'Meadhra, K. Goel, and N. Michael, "Communication-efficient planning and mapping for multi-robot exploration in large environments," *RA-L*, vol. 4, no. 2, 2019.

[5] R. G. Colares and L. Chaimowicz, "The next frontier: combining information gain and distance cost for decentralized multi-robot exploration," in *ACM Symposium on Applied Computing*, 2016.

[6] N. C. Fung, C. Nieto-Granda, J. M. Gregory, and J. G. Rogers, "Autonomous exploration using an information gain metric," *US Army Research Laboratory Adelphi United States*, 2016.

[7] L. Schmid, M. Pantic, R. Khanna, L. Ott, R. Siegwart, and J. Nieto, "An efficient sampling-based method for online informative path planning in unknown environments," *RA-L*, vol. 5, no. 2, pp. 1500–1507, 2020.

[8] Y. Kantaros, B. Schlotfeldt, N. Atanasov, and G. J. Pappas, "Asymptotically optimal planning for non-myopic multi-robot information gathering." in *Robotics: Science and Systems*, 2019.

[9] N. Atanasov, J. Le Ny, K. Daniilidis, and G. J. Pappas, "Decentralized active information acquisition: Theory and application to multi-robot slam," in *ICRA*. IEEE, 2015.

[10] R. T. Vaughan, K. Stoy, G. S. Sukhatme, and M. J. Mataric, "Lost: Localization-space trails for robot teams," *T-RO*, vol. 18, no. 5, 2002.

[11] N. R. Hoff, A. Sagoff, R. J. Wood, and R. Nagpal, "Two foraging algorithms for robot swarms using only local communication," in *Intl. Conf. on Robotics and Biomimetics*. IEEE, 2010.

[12] D. A. Shell and M. J. Mataric, "On foraging strategies for large-scale multi-robot systems," in *IROS*. IEEE, 2006, pp. 2717–2723.

[13] G. Notarstefano, K. Savla, F. Bullo, and A. Jadbabaie, "Maintaining limited-range connectivity among second-order agents," in *ACC*. IEEE, 2006.

[14] P. Yang, R. A. Freeman, G. J. Gordon, K. M. Lynch, S. S. Srinivasa, and R. Sukthankar, "Decentralized estimation and control of graph connectivity for mobile sensor networks," *Automatica*, 2010.

[15] L. Sabattini, N. Chopra, and C. Secchi, "On decentralized connectivity maintenance for mobile robotic systems," in *CDC / ECC*. IEEE, 2011.

[16] M. M. Zavlanos, M. B. Egerstedt, and G. J. Pappas, "Graph-theoretic connectivity control of mobile robot networks," *Proc. of the IEEE*, vol. 99, no. 9, pp. 1525–1540, 2011.

[17] P. E. Hart, N. J. Nilsson, and B. Raphael, "A formal basis for the heuristic determination of minimum cost paths," *Trans. on Systems Science and Cybernetics*, vol. 4, no. 2, pp. 100–107, 1968.

[18] M. Likhachev, G. J. Gordon, and S. Thrun, "ARA*: Anytime A* with provable bounds on sub-optimality," in *Advances in neural information processing systems*, 2004, pp. 767–774.

[19] S. Aine and M. Likhachev, "Anytime truncated D*: Anytime replanning with truncation," in *Symposium on Combinatorial Search*, 2013.

[20] S. M. LaValle, "Rapidly-exploring random trees: A new tool for path planning," Computer Science Deptartment, Iowa State University, Tech. Rep. TR-98-11, October 1998.

[21] S. Karaman and E. Frazzoli, "Sampling-based algorithms for optimal motion planning," *IJRR*, 2011.

[22] N. Ratliff, M. Zucker, J. A. Bagnell, and S. Srinivasa, "CHOMP: Gradient optimization techniques for efficient motion planning," in *ICRA*. IEEE, 2009.

[23] M. Kalakrishnan, S. Chitta, E. Theodorou, P. Pastor, and S. Schaal, "STOMP: Stochastic trajectory optimization for motion planning," in *ICRA*. IEEE, 2011.

[24] J. Schulman, Y. Duan, J. Ho, A. Lee, I. Awwal, H. Bradlow, J. Pan, S. Patil, K. Goldberg, and P. Abbeel, "Motion planning with sequential convex optimization and convex collision checking," *IJRR*, vol. 33, no. 9, pp. 1251–1270, 2014.

[25] J. J. Kuffner and S. M. LaValle, "RRT-connect: An efficient approach to single-query path planning," in *ICRA*, vol. 2. IEEE, 2000.

[26] Y. Chen, X. C. Ding, and C. Belta, "Synthesis of distributed control and communication schemes from global ltl specifications," in *CDC / ECC*. IEEE, 2011.

[27] E. Stump, N. Michael, V. Kumar, and V. Isler, "Visibility-based deployment of robot formations for communication maintenance," in *ICRA*. IEEE, 2011.

[28] N. Majcherczyk, A. Jayabalan, G. Beltrame, and C. Pinciroli, "Decentralized connectivity-preserving deployment of large-scale robot swarms," in *IROS*. IEEE, 2018, pp. 4295–4302.

[29] D. Tardioli, A. R. Mosteo, L. Riazuelo, J. L. Villarroel, and L. Montano, "Enforcing network connectivity in robot team missions," *IJRR*, 2010.

[30] J. Fink, "Communication for teams of networked robots," 2011.

[31] Y. Solana, M. Furci, J. Cortés, and A. Franchi, "Multi-robot path planning with maintenance of generalized connectivity," in *Intl. Symposium on Multi-Robot and Multi-Agent Systems*. IEEE, 2017.

[32] M. Malmirchegini and Y. Mostofi, "On the spatial predictability of communication channels," *Trans. on Wireless Communications*, 2012.

[33] R. Parasuraman, P. Oegren, and B.-C. Min, "Kalman filter based spatial prediction of wireless connectivity for autonomous robots and connected vehicles," in *Vehicular Technology Conf.* IEEE, 2018.

[34] Y. Mostofi, M. Malmirchegini, and A. Ghaffarkhah, "Estimation of communication signal strength in robotic networks," in *ICRA*. IEEE, 2010, pp. 1946–1951.

[35] Y. Mostofi and P. Sen, "Compressed mapping of communication signal strength," in *Military Communications Conf.* IEEE, 2008.

[36] J. Fink and V. Kumar, "Online methods for radio signal mapping with mobile robots," in *ICRA*. IEEE, 2010, pp. 1940–1945.

[37] J. Fink, J. Twigg, P. Yu, and B. M. Sadler, "A parsimonious model for wireless connectivity in robotic networks," in *Global Conf. on Signal and Information Processing*. IEEE, 2013.

[38] P. Series, "Propagation data and prediction methods for the planning of indoor radio communication systems and radio local area networks in the frequency range 300 MHz to 100 GHz," *Recommendation ITU-R*, pp. 1238–8, 2015.

[39] J. Stephan, J. Fink, V. Kumar, and A. Ribeiro, "Concurrent control of mobility and communication in multirobot systems," *IEEE Transactions on Robotics*, vol. 33, no. 5, pp. 1248–1254, 2017.

[40] I. A. Şucan, M. Moll, and L. E. Kavraki, "The open motion planning library," *RAM*, vol. 19, no. 4, 2012.

[41] S. G. Johnson, "The NLopt nonlinear-optimization package," http://github.com/stevengj/nlopt.

[42] D. Kraft, "A software package for sequential quadratic programming," Institut für Dynamik der Flugsysteme, Oberpfaffenhofen, Tech. Rep. DFVLR-FB 88-28, July 1988.

[43] ——, "Algorithm 733: TOMP–fortran modules for optimal control calculations," *Trans. on Mathematical Software*, 1994.

[44] J. Pan, S. Chitta, and D. Manocha, "FCL: A general purpose library for collision and proximity queries," in *ICRA*. IEEE, 2012, pp. 3859–3866.