EE 290S Project Writeup: Information-Optimal Autonomous Robot Exploration

David Fridovich-Keil

***Abstract*— This report considers the general problem of autonomous exploration, in which a mobile robot seeks to build a map of its environment as quickly as possible. Often, this is formulated as an information maximization problem and solved using Monte Carlo methods. In this report, we introduce a novel reformulation in which the optimization variable is not a single trajectory, but rather the distribution over all trajectories. Finally, a more sample-efficient version of this idea is presented. The main contribution of this work is a theoretical guarantee that all three methods are equivalent in their simplest forms, in the sense that all will yield the same information-optimal trajectory.**

I. Introduction

Autonomous exploration is a well-studied problem in the robotics literature, and in practice it is often a central component in many industrial applications. There are a number of common approaches to formulating the exploration problem, and even more approaches to solving the problem efficiently. Despite this variety of techniques, a significant stumbling block persists: computational intractability and the curse of dimensionality.

This work approaches the problem from the perspective of information theory, as a modified instance of the channel coding problem. A sequence of algebraic manipulations allows us to prove one of our main results: that the noiseless modified channel coding problem reduces to the current state of the art. We begin to address the issue of computational intractability by proposing a robust version of the channel coding problem which allows us to reduce sample complexity. Our second main result is that the simplest incarnation of this robust problem reduces to the initial problem.

Section II provides a brief overview of the current state of the art in this area. We formulate the general problem as a variant of the classical channel coding problem in Sec. III. Section III-C presents a recurring example – the *radiation detection problem* – which we will treat as a toy problem to illustrate our approach. Section IV presents our first attempt to solve the modified channel coding problem, and proves our first main result. Section V introduces a robust version of the previous result, with the intent of improving sampling complexity. Section VI describes a software implementation of these ideas. Section VII presents an empirical validation of our earlier results, and Sec. VIII offers some final remarks on the utility of these methods and suggests steps for future work.

Berkeley Artificial Intelligence Research Laboratory (BAIR), UC Berkeley, Berkeley CA, USA. dfk@eecs.berkeley.edu

II. Related Work

There are at least three major approaches to solving the exploration problem: Markov decision processes (MDPs), frontier exploration, and information maximization.

MDPs actually predate much of the modern work on exploration, and are in fact more general. As most commonly formulated, e.g. in [7], MDPs model the robot's *state* as a Markov process, where the state is taken to be a unique representation of all relevant information for a particular task. At each time step, the robot may transition from its current state i to another state j (both in state space Σ) and thereby collect some *reward* r_{ij}. The solution to the MDP is then taken to be a *policy* $\pi : \Sigma \longrightarrow \Sigma$ that maps the current state to the optimal next state such that the time-discounted expected reward is maximized for all time. There exist dynamic programming approaches to solving this problem: unfortunately they are known to be computationally intractable for large enough (though still rather small) problems.

In practice, greedy methods are often used in place of an MDP, among them frontier exploration and information maximization. Frontier exploration arose early in the robotics literature as an algorithm to guide robots toward unexplored territory, as in [8]. The idea is to keep track of which areas of space are explored and unexplored, and to move toward the nearest unexplored area. Once the unexplored area is within the robot's field of view (for whatever sensor it is using), it acquires measurements of the space. Over time, the explored area grows and the unexplored area shrinks, yielding a trivial proof of convergence.

Unfortunately, frontier exploration is often a poor heuristic in three dimensions due to the prevalence of small, partially-occluded regions. However, information maximization approaches tend to be more robust to such distractions, as they are specifically designed to balance the information gained from exploring occluded areas with that from exploring larger regions that may be further away, or even from re-exploring territory that may have been hastily passed over. For example, in [2] robot actions are chosen at each time step to maximize an information metric defined over its environment representation and its internal state estimate, using a finite, receding horizon. In recent years, this approach has become dominant in the robotics literature: for a more modern implementation, see [3].

Unfortunately, information maximization over a receding time horizon suffers two major difficulties. First, as with

most greedy algorithms, it is not optimal in general (although convergence was shown in [5]). Second, as the time horizon over which planning occurs increases, the number of feasible trajectories increases exponentially which can make the problem computationally intractable. In practice, the first issue is completely ignored, and the second is often mitigated by choosing trajectories from a small subset of feasible trajectories (again, leading to sub-optimality).

In this work, we examine the information maximization approach from the perspective of channel coding, and leverage the structure of that problem to begin to design a more efficient exploration algorithm.

III. Problem Formulation

Figure 1 provides a block-level overview of the autonomous exploration problem. At every time step, the robot is required to choose some trajectory X, based on its current understanding of the "map" m, which encodes the robot's knowledge of the true map M. This trajectory results in a (possibly noisy) "measurement" Z of the environment. These measurements are then accumulated by the robot and incorporated into an improved map estimate. In the information maximization formulation of exploration, the objective of the robot in choosing each successive trajectory is to gather information as quickly as possible, so that, ideally, $m \longrightarrow M$ in as few steps as possible.

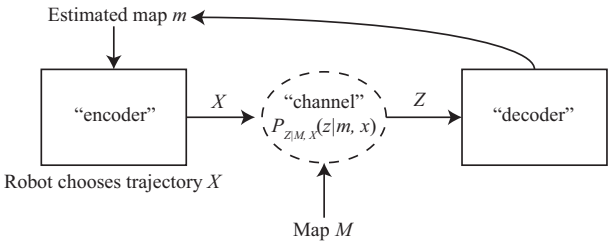

Fig. 1. Block diagram of the robot exploration problem. A robot is required to choose its next trajectory based on its current map estimate.

In the information maximization approach described in Sec. II, it is common to formulate the problem of choosing an optimal trajectory x^* as follows:

$$x^* = \arg\max_x I(Z; M|X = x) \tag{1}$$

i.e. that the optimal trajectory is the one which, on average, results in a measurement Z which reveals the most information about the map M.

A. Through the Lens of Channel Coding

The "encoder," "channel," and "decoder" blocks in Fig. 1 are intended to evoke the classical channel coding problem, in which the goal is to encode some data source in such a way that the data can be recovered despite being corrupted by a noisy transmission process. When the noise statistics – in our case, the conditional measurement distribution $p_{Z|X}$ for a fixed known map M – are known, the so-called channel capacity C is given by Shannon's channel coding theorem:

$$C = \max_{p_X} I(Z; X) \tag{2}$$

Moreover, there exists an efficient method, the Blahut-Arimoto algorithm, for solving Eq. 2 – which is a convex optimization problem – efficiently.

Notice that to achieve the channel capacity, we are required to find a probability distribution over trajectories X, rather than a single, optimal trajectory. Intuitively, it seems plausible that solving this convex relaxation may be more efficient than solving the integer program in Eq. 1.

The natural way to apply the channel coding idea to Eq. 1 is as follows:

$$p_X^* = \arg\max_{p_X} I(Z; M|X) \tag{3}$$

where we choose the input distribution p_X such that the average information revealed by the resulting sequence of measurements Z is maximized.

B. Preliminaries

Before proceeding, we state several major assumptions about the nature of the trajectories, measurements, and maps in this problem.

Assumption 1: (Discreteness) Throughout this work, it is assumed that all quantities (trajectories, measurements, and maps) are discrete. Formally, we denote this by $X \in \mathcal{X}, Z \in \mathcal{Z}, M \in \mathcal{M}$ where the sets $\mathcal{X}, \mathcal{Z}, \mathcal{M}$ are countable.

Assumption 2: (Finiteness) For computational tractability, we also assume that the sizes of these sets are finite and "small" – by which we mean that enumerating them is not a significant computational burden. Note that this requirement is not necessary for any of the theoretical results; it is only required for an efficient implementation in practice.

Assumption 3: (Determinism) We assume that measurements Z are deterministic, i.e. given a fixed trajectory x and map m, the corresponding measurement Z is no longer random. That is, $Z = z(x, m)$ for some well defined function $z : \mathcal{X} \times \mathcal{M} \longrightarrow \mathcal{Z}$.

C. Toy Problem: Radiation Mapping

The preceding formulation is intentionally abstract. To be more explicit, consider the following instance of the exploration problem, shown in Fig. 2.

A robot is tasked with finding k point sources of radiation $\{s_i\}_{i=1}^k$ known to lie somewhere on a $N \times P$ grid, i.e. $s_i \in \mathbb{F}_N \times \mathbb{F}_P$. It has a sensor which counts the number of sources in its θ-degree field of view. At each time step, it can move from its current position on the grid to any diagonal or adjacent grid cell, and rotate by $\pm\theta_0$, where θ_0 is a small angular step size. The goal is to choose a distribution over short sequences of such movements, i.e. over trajectories X, such that the resulting sequence of measurements Z reveals the most information about the underlying map M as specified in Eq. 3.

IV. Finding the Best Distribution

Our first main result is that the channel coding approach always yields the same optimal trajectory as the classical information maximization approach:

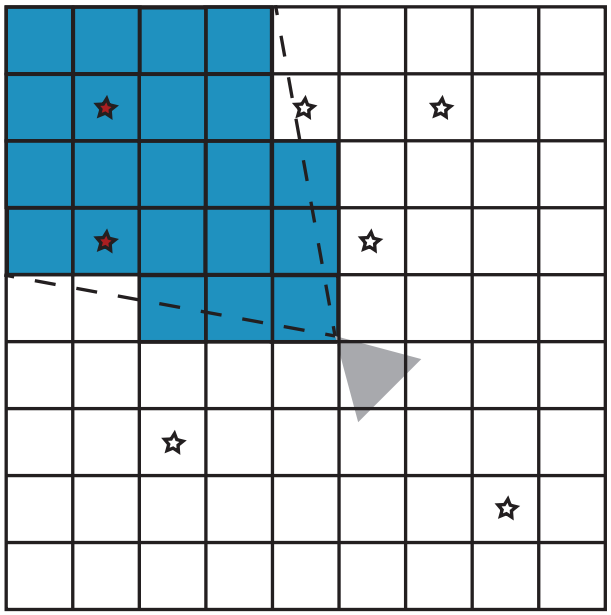

Fig. 2. Diagram of radiation detection scenario with $k = 7$ sources, two of which are in the blue-shaded field of view.

Theorem 1: The solution p_X^* to Eq. 3 is a delta function, and all mass is concentrated on the solution x^* to Eq. 1.

We devote the remainder of this section to proving Theorem 1.

A. Proof of Theorem 1

Begin by expanding the objective function in Eq. 3, as follows:

$$\begin{aligned} I(Z;M|X) &= H(M|X) - H(M|Z,X) && (4)\\ &= H(Z|X) - H(Z|M,X) && (5)\\ &= H(Z|X) - 0 && (6)\\ &= -\sum_x p_X(x) \sum_z p_{Z|X}(z|x) \log p_{Z|X}(z|x) && (7) \end{aligned}$$

Note that $H(Z|M,X) = 0$ in Eq. 6 because we have assumed in Assumption 3 that a sequence of measurements is completely determined by a trajectory and a map.

Observe that $I(Z;M|X)$ is linear in p_X, i.e. if we define

$$c(x) \doteq -\sum_z p_{Z|X}(z|x) \log p_{Z|X}(z|x) \tag{8}$$

then we can rewrite Eq. 3 as follows:

$$p_X^* = \arg\max_{p_X} c^T p_X \tag{9}$$

Further, the only constraint on the optimization problem in Eq. 9 is that p_X be a probabilty distribution on $\mathcal{X}$, i.e. that

$$p_X \geq 0 \text{ and } \mathbf{1}^T p_X = 1 \tag{10}$$

Thus, Eq. 9 is a linear program (LP), which implies that its solution is a vertex of the feasible set defined in Eq. 10, i.e. a vertex of the probability simplex over $\mathcal{X}$. Explicitly, $p_X^*(x) = 0$ for all x except the optimal one, at which $p_X^* = 1$. Had the feasible set been only the set of vertices, the solution would be the same. Optimizing over the set of vertices is equivalent to choosing the single best trajectory in Eq. 1, which proves Theorem 1.

V. Reducing Sample Complexity

In practice, the vector c must be computed at each time step in order to set up the LP of Eq. 9. This is done by first estimating the distribution $p_{Z|X}$ with the empirical distribution generated by Monte Carlo sampling, and then plugging into the formula in Eq. 8. Unfortunately, estimating $p_{Z|X}$ accurately requires *lots* of samples: the sample complexity scales at least as fast as $\text{card}(\mathcal{X}) \times \text{card}(\mathcal{M}) \times \text{card}(\mathcal{Z})$ (the total support size of all variables). Even in the toy radiation detection problem, this can become prohibitive.

Suppose, however, that we draw only n samples from the true joint distribution $p_{Z,X,M}$, where n is smaller than the coarse scaling law above. The resulting estimate $\hat{p}_{Z|X}$ will converge to the true distribution $p_{Z|X}$ as $n \longrightarrow \infty$. For any finite n, the error $p_{Z|X} - \hat{p}_{Z|X}$ results in an error η in computing c, i.e. $\eta = c - \hat{c}$. If we had an upper bound on some norm of η, we might formulate the following "robust" version of Eq. 9:

$$\begin{aligned} p_X^* = \arg\max_{p_X} \min_{\eta} (\hat{c} + \eta)^T p_X \\ \text{subject to } \|\eta\| \leq \epsilon \end{aligned} \tag{11}$$

where the robot is now trying to choose the best distribution over trajectories such that, for the worst case sampling-induced error η the resulting mutual information is maximized.

If we take the norm on η to be the ℓ_∞ norm – or equivalently, if we have a bound on $|c(x) - \hat{c}(x)|$ for all x – then our second main result states that the solution to Eq. 11 again coincides with the solution to Eq. 1.

Theorem 2: The solution to Eq. 11 where the bound on η is taken to be in ℓ_∞ coincides with the solution to the non-robust version in Eq. 9 and by Theorem 1 to that of Eq. 1.

The proof of Theorem 2 is given in Sec. V-B. First, however, in Sec. V-A we show how to obtain an upper bound ϵ on $\|\eta\|_\infty$ as a function of the number of samples n. In Sec. V-C we consider a small adjustment to Eq. 11 in which Theorem 2 does not apply.

A. Toward a Bound on Sampling Error

We state the following lemma without proof (the proof may be found in [4]):

Lemma 3: If $p \doteq p_{Z|X=x}$ is estimated from n IID samples of the true distribution, with the estimate denoted $\hat{p}$, then

$$\mathbb{P}\{D(\hat{p}\|p) \leq \delta\} \geq 1 - 2^{-n\delta + \text{card}(\mathcal{Z}) \log_2(n+1)} \tag{12}$$

Lemma 3 provides an upper bound on the relative entropy between the estimated distribution $\hat{p}$ and the true distribution p, that holds with high probability for sufficiently large n.

Further application of Pinsker's inequality yields a bound on the ℓ_1 deviation:

$$\frac{\log_2(e)}{2}\|p-\hat{p}\|_1^2 \leq D(\hat{p}\|p) \leq \delta \tag{13}$$

Continuing, the Fannes-Audenaert inequality [6] yields:

$$\eta(x) = |H(p) - H(\hat{p})| \tag{14}$$

$$\leq \frac{\|\hat{p}-p\|_1}{2}\log_2(k) + H_2\left(\frac{\|\hat{p}-p\|_1}{2}\right) \tag{15}$$

$$\leq \log_2(k)\delta' + H_2(\delta'), \delta' \doteq \sqrt{\frac{\delta}{2\log_2(e)}} \tag{16}$$

where H_2 is the binary entropy function and $\delta \ll 1$. This yields the following ℓ_∞ bound on η:

$$\|\eta\|_\infty \leq \epsilon, \epsilon = \log_2(k)\delta' + H_2(\delta') \tag{17}$$

Thus, we have derived a reasonably tight upper bound on the ℓ_∞ norm of estimation error η.

B. Proof of Theorem 2

We prove Theorem 2 with a short sequence of manipulations that reduce Eq. 11 to Eq. 9:

$$p_X^* = \arg\max_{p_X} \min_{\eta:\|\eta\|_\infty \leq \epsilon} (\hat{c}+\eta)^T p_X \tag{18}$$

$$= \arg\max_{p_X} \hat{c}^T p_X + \min_{\eta:\|\eta\|_\infty \leq \epsilon} \eta^T p_X \tag{19}$$

$$= \arg\max_{p_X} \hat{c}^T p_X - \epsilon \cdot \text{sgn}\,(p_X)^T\, p_X \tag{20}$$

$$= \arg\max_{p_X} \hat{c}^T p_X - \epsilon\|p_X\|_1 \tag{21}$$

$$= \arg\max_{p_X} \hat{c}^T p_X - \epsilon \tag{22}$$

$$= \arg\max_{p_X} \hat{c}^T p_X \tag{23}$$

where the reduction in Eq. 22 holds because p_X is a probability vector, which implies that its elements sum to unity.

We have shown that the solution to Eq. 11 coincides with that to Eq. 9 when estimation error η is bounded in the ℓ_∞ sense. This suffices to prove the theorem.

C. Loosening the Bound on Sampling Error

One may interpret Theorem 2 optimistically or pessimistically. On the one hand, we have shown that, even accounting for an adversarial sampling error, the information-optimal distribution over trajectories remains the same delta function as in the literature. This should give us some measure of comfort, that we can do no better than choosing the best trajectory (indeed, it may be why the state of the art is so good!). On the other hand, we may expect that choosing a non-degenerate distribution over trajectories – in a sense, probabilistically interpolating between trajectories – might be advantageous given that our estimates of information gain are imprecise. Theorem 2 tells us that, under a box uncertainty model, this is not so.

In this section, we consider another natural uncertainty model which arises when we loosen the bound in Eq. 17. Since $\|\eta\|_\infty \leq \|\eta\|_2$, we know that $\{\eta : \|\eta\|_2 \leq \epsilon\} \subset \{\eta : \|\eta\|_\infty \leq \epsilon\}$, which implies the following inequality:

$$\max_{p_X} \min_{\eta:\|\eta\|_2 \leq \epsilon} (\hat{c}+\eta)^T p_X \geq \max_{p_X} \min_{\eta:\|\eta\|_\infty \leq \epsilon} (\hat{c}+\eta)^T p_X \tag{24}$$

i.e. the loose ℓ_2 bound results is a comparatively optimistic predictor of information gain.

Regardless, we can compute the value of p_X at optimum, following a similar line of reasoning as in Sec. V-B:

$$p_X^* = \arg\max_{p_X} \min_{\eta:\|\eta\|_2 \leq \epsilon} (\hat{c}+\eta)^T p_X \tag{25}$$

$$= \arg\max_{p_X} \hat{c}^T p_X + \min_{\eta:\|\eta\|_2 \leq \epsilon} \eta^T p_X \tag{26}$$

$$= \arg\max_{p_X} \hat{c}^T p_X - \epsilon \cdot \frac{p_X^T}{\|p_X\|_2} p_X \tag{27}$$

$$= \arg\max_{p_X} \hat{c}^T p_X - \epsilon\|p_X\|_2 \tag{28}$$

Eq. 28 is a second order cone program (SOCP), and may be solved efficiently with an off-the-shelf commercial solver. Effectively, the ℓ_2 penalization causes the solution p_X^* to move away from the vertices of the feasible set, balancing the contributions from different trajectories. This is illustrated in 2D in Fig. 3, where the solution to the original LP (the red dot) is shifted a small amount along the probability simplex toward the other vertex.

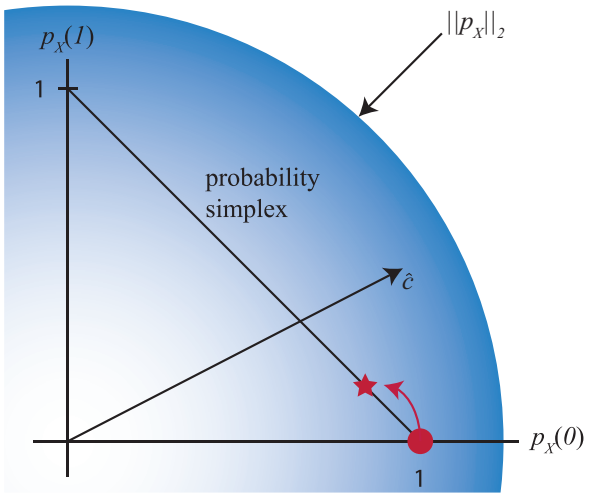

Fig. 3. Illustration of the ℓ_2-constrained robust problem from Eq. 28. For the vector $\hat{c}$ shown, the solution to the original LP is given by the red dot. Introduction of the ℓ_2-penalization, visualized by the blue gradient field, causes the optimum to move upward along the probability simplex to the location marked by the red star.

As mentioned above, such probabilistic interpolation – choosing a trajectory at random according to the distribution p_X^* – is an intuitively desirable property; it may be viewed as "hedging" across different possible values of estimation error η.

VI. Implementation

We provide open-source C++ and Python implementations of the LP-based formulation in Eq. 9 and the SOCP-based formulation in Eq. 28 to solve the toy radiation

detection problem: `https://dfridovi.github.io/exploration/`.

Although the statement of the radiation detection problem given in Sec. III-C is fairly simple and straightforward, there is one non-trivial step which has yet to be discussed. That is, how to fuse measurements Z into a posterior distribution on map M at each time step; effectively, how to close the loop in Fig. 1. For a detailed treatment of this topic, see the appendix.

VII. Empirical Results

Figure 4 shows the results of an empirical comparison between the linear information maximization of Eq. 9, its ℓ_2-robust counterpart Eq. 28, and a simple random walk explorer, where the grid size is 10×10 and there are $k = 2$ sources. The LP- and SOCP-based explorers are optimizing over trajectories of length one, using $n = 1000$ Monte Carlo samples to generate $\hat{c}$, with $\epsilon = 0.1$. Results are based on 1000 random initializations.

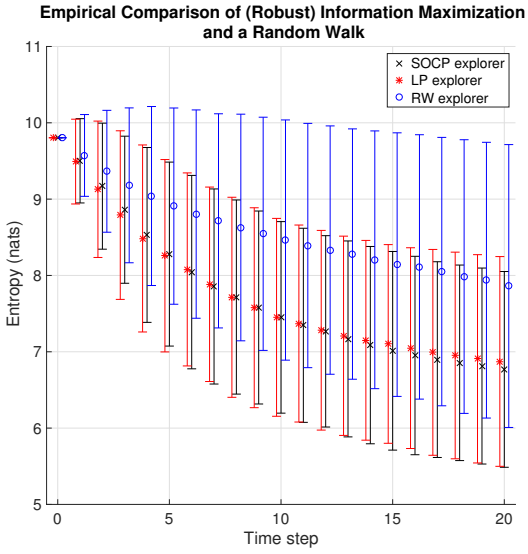

Fig. 4. Empirical exploration rates for the LP-based non-robust explorer of Eq. 9, the SOCP-based robust explorer of Eq. 28, and a random walk (RW) on a 10×10 grid with $k = 2$ sources. Error bars are at ± 1 standard deviation. Note that the information maximization algorithms are more efficient on average than a random walk (which was expected), while there is little difference between robust and non-robust explorers. In larger problems, we expect the difference to be more striking.

As shown in Fig. 4, even in this extremely small version of the problem, the information maximization algorithms outperform the baseline. This pattern should become even more pronounced in more complicated problems. Moreover, although we do not observe a significant difference between robust and non-robust information maximizers here, we speculate that in larger problems the robust version may outperform the non-robust version.

VIII. Remarks and Future Directions

Three remarks are in order, each of which concludes with a suggestion for future work.

Remark 1: (Noiseless measurements) We have assumed throughout this work that measurements were noiseless (see Assumption 3). Although this is almost never true in practical problems, it was necessary for the original convex information maximization problem in Eq. 3 to reduce to a linear objective, which was crucial to the proof of both of our main results. If we allowed noise in the measurements, then the simplification in Eq. 6 ($H(Z|M, X) = 0$) would no longer hold. However, for any particular noise model, it should be possible to compute (or estimate from samples) that conditional entropy term, and include it in the computation. The resulting problem could be solved using the Blahut-Arimoto algorithm, for example. Regardless, the noisy case would not reduce to the state of the art information maximization of Eq. 1. It would be useful to examine this case in detail in future work.

Remark 2: (Conservativeness of bounds) The bounds in Sec. V-A, though in general tight, may not be always be useful or meaningful. For example, the bound in Lemma 3 suggests that, for $k = 2$ sources and $\delta = 0.01$, at least $n \approx 10^4$ samples are required *for each valid trajectory!* Considering three-step trajectories of the sort described in Sec. III-C, this works out to $\sim 2 \times 10^8$ samples total – which is clearly impractical for real-time decision making. As suggested in Sec. V-C, one way to sidestep this issue is to artificially loosen the bound. Still, there may be other, more parsimonious solutions which we have not considered.

Remark 3: (Direct computation of mutual information) A particularly promising approach to reducing the computational complexity of these methods is to replace the Monte Carlo estimation process of Sec. V with a direct computation based on *a priori* knowledge of the sensor and noise models. We focused on Monte Carlo sampling for two reasons: (1) it is common in practice, and (2) it is very simple to extend to new and more complicated sensor and noise models. An interesting direction for future work will be to combine these methods efficiently in more complicated planning problems.

IX. Conclusion

In this report, we have presented a channel coding-inspired reformulation of the information maximization approach to robot trajectory planning. Under the assumption of deterministic measurements – or equivalently, that all channel noise is due to an unknown environment – we showed in Theorem 1 that this new formulation coincides with the state of the art. Additionally, we considered how this approach might be made robust to errors induced by decreasing the sample complexity. In Theorem 2, we showed that a simple ℓ_∞ bounding scheme again coincided with the original approach. Finally, we proposed an artificially loose ℓ_2 bound which exhibits an interesting "interpolation" property, presented empirical results, and discussed several shortcomings and directions for future work.

References

[1] Sameer Agarwal, Keir Mierle, et al. *Ceres Solver*. `http://ceres-solver.org`.

[2] F. Bourgault et al. "Information based adaptive robotic exploration". In: *Proc. IEEE/RSJ Intl. Conf. on Intelli. Robots and Systs.* Vol. 1. Lausanne, Switzerland, 2002, pp. 540–545.
[3] B. Charrow et al. "Information-Theoretic Mapping Using Cauchy-Schwarz Quadratic Mutual Information". In: *Proc. IEEE Intl. Conf. on Robotics and Autom.* Seattle, USA, 2015.
[4] Thomas M Cover and Joy A Thomas. *Elements of information theory*. John Wiley & Sons, 2012.
[5] Brian J Julian, Sertac Karaman, and Daniela Rus. "On mutual information-based control of range sensing robots for mapping applications". In: *Proc. IEEE/RSJ Intl. Conf. on Intelli. Robots and Systs.* Tokyo, Japan, 2013, pp. 5156–5163.
[6] Igor Sason. "Entropy Bounds for Discrete Random Variables via Maximal Coupling". In: *IEEE Trans. Info. Th.* (July 2013).
[7] Sebastian Thrun, Wolfram Burgard, and Dieter Fox. *Probabilistic robotics*. MIT press, 2005.
[8] B Yamauchi. "A frontier-based approach for autonomous exploration". In: *Proc. IEEE Sym. on Comput. Intelli. in Robot. and Autom.* Monterey, USA, 1997, pp. 146–151.

APPENDIX

Here, we describe the approach taken to fusing successive simulated radiation measurements together efficiently into a posterior distribution over maps: $p_{M|X,Z}$. Naive application of Bayes' rule reveals:

$$p_{M|X,Z} = \frac{p_{M,X,Z}}{p_{X,Z}} \tag{29}$$

$$= \frac{p_{M,X}}{p_X} \tag{30}$$

under the noiseless measurement assumption. Unfortunately, this is little help: the joint distribution $p_{M,X}$ is just as hard to write down as the conditional $p_{M|X,Z}$.

As a simple alternative, we model the distribution of M as a function f supported on the grid where, at each grid cell, the function value is the probability that a source lies in that cell, independent of all other cells. More precisely, $f : \mathbb{F}_N \times \mathbb{F}_P \longrightarrow [0, 1]$ and

$$p_M(m = \{(i_1, j_1), ..., (i_k, j_k)\}) = \prod_{(i,j)\in m} f(i,j) \prod_{(i,j)\notin m} 1 - f(i,j) \tag{31}$$

Under this assumption, it is relatively straightforward to enforce consistency across the observed data. At each time step t, we record the set $\mathcal{V}_t$ of grid cells that were in the field of view, and also the corresponding measurement z_t. We then solve for the function f_t which is consistent with the measurements, and with the constraint that the total number of sources on the grid is k:

$$f_t = \arg\min_f \sum_{\tau\le t} \left(z_\tau - \sum_{(i,j)\in\mathcal{V}_\tau} f(i,j) \right)^2 + \lambda t \left(k - \sum_{(i,j)} f(i,j) \right)^2 \tag{32}$$

Note that we have introduced a regularization constant $\lambda > 0$ and scaled it by t. In practice, we set $\lambda = 1$ so that the terms are equally weighted. This is a linear least squares problem with implicit bounds constraints ($f(i,j) \in [0, 1]$), and we solve it using the Google Ceres solver [1]. Although this model is not fully general (for instance, it cannot model multiple sources at the same location), it is sufficient for our toy problem and is not so different from the data structures used in more complicated problems.