

Information Gain Is Not All You Need

Ludvig Ericson^{*,1}, José Pedro^{*,1,2}, Patric Jensfelt¹

Abstract—Autonomous exploration in mobile robotics is driven by two competing objectives: coverage, to exhaustively observe the environment; and path length, to do so with the shortest path possible. Though it is difficult to evaluate the best course of action without knowing the unknown, the unknown can often be understood through models, maps, or common sense. However, previous work has shown that improving estimates of information gain through such prior knowledge leads to greedy behavior and ultimately causes backtracking, which degrades coverage performance. In fact, any information gain maximization will exhibit this behavior, even without prior knowledge. Information gained at task completion is constant, and cannot be maximized for. It is therefore an unsuitable choice as an optimization objective. Instead, information gain is a decision criterion for determining which candidate states should still be considered for exploration. The task therefore becomes to reach completion with the shortest total path. Since determining the shortest path is typically intractable, it is necessary to rely on a heuristic or estimate to identify candidate states that minimize the total path length. To address this, we propose a heuristic that reduces backtracking by preferring candidate states that are close to the robot, but far away from other candidate states. We evaluate the performance of the proposed heuristic in simulation against an information gain-based approach and frontier exploration, and show that our method significantly decreases total path length, both with and without prior knowledge of the environment.

I. INTRODUCTION

Autonomous exploration is a fundamental problem in mobile robotics, found in a wide variety of applications ranging from search and rescue [3, 4] to industrial inspection [18]. Its definition varies between applications, and we identify two main reoccurring variations: *budget-constrained exploration*, and *quality-constrained exploration*. In budget-constrained exploration, it is assumed that the robot’s budget is limited, e.g., by its battery, a deadline, or a camera roll. The budget is insufficient to fully cover the environment and the optimal set of views must be selected. In quality-constrained exploration, the constraint is flipped and it is assumed that the robot can cover its environment, and instead must collect views such that the whole map is of some minimum level of quality, i.e., the remaining *gain* falls below a given threshold.

^{*} Equal contribution.

¹ Authors are with the Division of Robotics, Perception and Learning at KTH Royal Institute of Technology, Stockholm, SE-10044, Sweden.

² Author is with Ericsson Research, Stockholm, Sweden.

This work was supported by the Swedish Research Council, and the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation. For e-mail correspondence, contact ludv@kth.se. Experiments were run on the Berzelius cluster provided by NSC at Linköping University, Sweden.

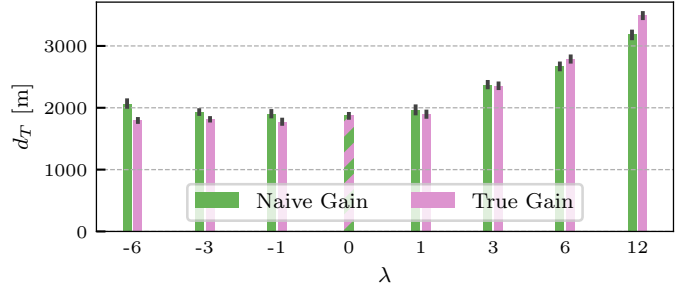


Fig. 1: Distance at completion d_T for a selection of gain affinities λ , where a higher λ means stronger preference for gain and a lesser concern with the length of the path to acquire it. Naive gain refers to the assumption that unknown space is occlusion-free, i.e., yields maximal gain; in true gain, the real would-be sensor scan is used for gain computation. Tellingly, negative affinities, i.e., *minimizing* gain, results in a lower d_T than maximization, and no choice is substantially better than nearest frontier, i.e., $\lambda = 0$.

Crucially, no part of the environment is left unexplored in quality-constrained exploration, implying that the total gain is essentially constant. The rich body of exploration methods that maximize gain must therefore be said to implicitly assume a budget-constrained scenario, because to maximize gain per sensor scan is to assume that some gain will be left unexplored, and that total gain is not constant.

The two paradigms are sometimes mixed up, such as by using a method suited for budget-constrained exploration but evaluating its quality-constrained properties, e.g., evaluating gain maximization by path length. In many cases, authors report counterintuitive results such as longer total path lengths and greedy behavior with improved gain estimation, owing to unnecessary backtracking as exploration draws to a finish [9, 15, 17]. Figure 1 shows that path length increases with stronger affinity to maximize gain, and that gain should in fact be *minimized*, if at all considered. This incongruence ultimately stems from the mix-up of the budget-constrained and quality-constrained paradigms: in budget-constrained exploration, obtaining gain is the priority, so greediness is, in fact, the desired outcome. Relatively little prior work can therefore be said to be targeted at the quality-constrained paradigm directly, which is the aim of this paper.

While some backtracking is inevitable, there can also be unnecessary backtracking. [17] show that unnecessary backtracking is incurred if and only if a given region is not yet explored by the *last* time that the optimal plan passes it. In other words, the robot typically has multiple opportunities

to explore a given region, and unnecessary backtracking is guaranteed only once it has missed its last chance.

The optimal plan is, of course, not available to the robot, and it cannot know when it has its last opportunity to explore a region. A cue that this might be the case, however, is that the robot is closer to that region now than it is expected to be later; it has a rare opportunity to explore the region with low path cost. We dub this heuristic *distance advantage*.

The phenomenon where gain maximization causes worse performance was first identified in works aiming to improve exploration by leveraging additional knowledge [17, 8]; consequently, in this paper, we propose a method for autonomous exploration that does not maximize gain, and instead centers on reducing backtracking by leveraging prior knowledge through our distance advantage heuristic.

The paper is outlined as follows: first, related work is analyzed through the lens of budget contra quality-constrained exploration in Section II, the quality-constrained exploration problem is then formally defined in Section III. Next, our proposed method is presented in Section IV, the experimental setup and results in Sections V to VII, and finally, our conclusions and limitations are presented in Sections VIII and IX. To summarize, our main contributions are:

- the relationship between autonomous exploration and gain maximization is disentangled once and for all, explaining and justifying counterintuitive findings reported in existing literature;
- a novel method for autonomous exploration, distance advantage, is proposed, which directly aims to minimize future potential backtracking; and
- its performance is evaluated in simulation, producing significantly shorter paths in the quality-constrained paradigm than nearest frontier [28] or gain maximization [10].

The reader is encouraged to review the supplementary video material for this paper, as it offers insights that are hard to convey with text and images. An implementation of the proposed method is available at lericson.se/da.

II. RELATED WORK

Autonomous exploration was initially proposed in the context of active perception [1], with the term being coined by Whaite and Ferry [26]. In many tasks, a single measurement is insufficient, either due to the nature of the sensor or due to uncertainty. Therefore, it is necessary to plan where to place the sensor in order to collect measurements that most reduce uncertainty. Since the focus of active perception was on object reconstruction in a small workspace, the cost of the path necessary to move the sensor was not relevant. Therefore, these first works were direct extensions of the next-best-view [5] methods from vision to robotics, by maximizing the predicted gain of the next measurement [1, 16, 19, 26].

The subtle confusion between budget-constrained exploration and quality-constrained exploration was already present in the active perception community. Whaite and Ferry [26]

defined the goal of autonomous exploration as that of determining representations of acceptable fidelity, i.e., quality-constrained exploration. However, they proposed to approach autonomous exploration through the lens of gain maximization, implicitly adopting the budget-constrained paradigm.

Carrying over to mobile robotics, where the cost of moving the robot cannot be neglected, the focus became to maximize gain with respect to the distance traveled [10, 28], remaining in the implicit budget-constrained paradigm. Finding the best trade-off between gathering information and moving efficiently has proven to be a challenging problem, leading to extensive research and the very term ‘autonomous exploration’ being appropriated by the mobile robotics community.

Yamauchi [28] argued that to efficiently observe the environment, the robot should plan to visit states that are predicted to have high gain while minimizing the path cost. The concept of frontiers, the border between known and unknown space, was introduced and it was proposed that exploration be done by navigating to the nearest frontier. Directly extending the next-best-view approaches, [10] proposed to explicitly optimize for measurement gain, weighted inversely to the distance necessary to collect it. Already when introducing gain maximization for mobile robotics, [10] observed that prioritizing gain led to fast short-term exploration, but ultimately made completing exploration take longer.

A wide body of literature exists building on [10, 28], extending it to more complex scenarios and improving upon its assumptions. RH-NBV [2] evaluates the objective along a *path*, instead of a single step. While searching over paths scales combinatorially, when compared to single decisions, this can be dealt with through sampling-based planning. AEP [22] combines frontiers and RH-NBV to mitigate the effect of greediness due to gain maximization in the global path, while preserving the efficient local coverage performance of RH-NBV. Some works [11, 31], most notably FUEL [31], attempt to plan optimal tours for a Traveling Salesman Problem (TSP) that visits every frontier cluster. The gain of each frontier cluster can also be locally optimized in a refinement step [31]. RH-NBV and AEP require extensive gain estimation, which can take up to 95% of planning time [20], while FUEL computes solutions to a TSP, which can be computationally demanding. Returning to single-step planning, UFOExplorer [7] shows that determining frontiers to be states with a minimum amount of gain, the nearest frontier exploration strategy produces shorter paths than RH-NBV, AEP and FUEL, while being computationally cheaper. ECHO [29] combines [7] and [31], choosing the nearest gain-having frontier and then optimizing the viewpoint to improve gain. More works extend these ideas, by improving upon the sampling-based planner and the path cost function [13, 20, 27], or considering improved estimates of gain through learning-based methods [6, 21, 23, 25], among others.

Importantly, most works aim to efficiently complete coverage of unknown environments and therefore evaluate quality-constrained exploration metrics, such as time or path length. However, they are implicitly performing budget-constrained

exploration since they maximize gain. Several counterintuitive results have been reported, such as that maximizing gain ultimately leads to longer coverage paths [10, 17], or that improved estimates of gain lead to longer paths [9].

Few works have explicitly addressed the mismatch between quality-constrained exploration goals and the budget-constrained optimization objectives. Li *et al* [12] attempted to determine the optimal exploration path offline by using A^* with an admissible heuristic, estimated from a complete map of the environment. The determined path is used to estimate the competitive ratio of other exploration strategies, showing that pure gain maximization is far from optimal. When an abstract topological map of the environment is available, a high-level global exploration path can be determined by solving a TSP [17]. The solution to the TSP can be used to determine the optimal order in which to visit frontiers, potentially using information gain to locally prioritize frontiers, similar to [22]. Ultimately, it is shown that using information gain leads to longer paths, and that strictly prioritizing the order in which frontiers are visited based on the TSP solution is better, since exploration is finished only when there are no more frontiers. Since more information about the environment does not lead to shorter paths for information gain, [9] proposed a heuristic that prioritizes visiting frontiers that observe elements of the environment that cannot be observed together with other elements. This is shown to lead to shorter paths than gain maximization and to improve with access to more information about the environment.

Combining information gain with quality-driven exploration is possible, and, in fact, paramount to tasks like active SLAM. There, exploration must compete with active localization: the goal is to produce a sufficiently high-quality map, but that requires maintaining an accurate state estimate. Works like [24, 30] address this scenario, where the role of information gain is not to drive exploration, but instead as a form of exploitation for minimizing state uncertainty.

III. PROBLEM STATEMENT

A plan π is a sequence of states, i.e., $\pi = (s_0, \dots, s_T)$, with s connected to its successor s' by the shortest path, with length $d(s, s')$. As the robot follows a plan, it builds a map M_π of the environment. The problem of determining a plan for quality-constrained exploration can then be formulated as

$$\min_{\pi} d(\pi) \quad s.t. \text{DesiredCoverage}(M_\pi), \quad (1)$$

where $d(\pi)$ is the total length of the plan. However, it is typically not possible to evaluate the coverage or feasibility of a complete plan, since at best the environment is partially known. Therefore, in most cases an exploration plan cannot be obtained offline by solving Eq. (1).

When performing exploration online, the exploration plan at any given time can be decomposed into three components: the plan already followed up to state s , the next state t to be determined, and the unknown optimal plan π^* that follows it.

Then, Eq. (1) can be reformulated as the sequential decision problem of determining the optimal next state

$$t^* = \arg \min_{t \in C} (d(s, t) + d(\pi^*)), \quad (2)$$

where C are the states that will improve the coverage of the map. While evaluating C and $d(\pi^*)$ is still not possible without knowing the environment, we now discuss how to address these problems when only the current map and, possibly, local predictions are available.

Like coverage, it is not possible to determine all states C that can improve coverage without knowing the whole environment. However, it is enough to consider the states F that advance coverage and are nearest to the current state, since any path to C must pass through one of these boundary states. F can be determined by finding the states that collect at least a minimum gain [7] or, more commonly, approximated by frontiers [28].

Since the states that improve coverage depend on the map and, through it, on the already followed plan, the decision problem suffers from the curse of history and is intractable to solve. Therefore, determining a good heuristic estimate of $d(\pi^*)$ is one of the fundamental problems of autonomous exploration planning. Predictive ability of the environment for autonomous exploration should reflect in better heuristic estimates and, consequently, better exploration plans. However, it has been reported in the literature that existing heuristics do *not* improve with better predictions [9], indicating there is room for improvement in the planning heuristic.

A. Traditional Exploration Heuristics

Having identified the sequential decision problem corresponding to quality-constrained exploration, it is now possible to analyze the implicit assumptions of the traditional autonomous exploration heuristics: nearest frontier and information gain.

Nearest frontier [7, 28]: The exploration plan is obtained through

$$t^* = \arg \min_{t \in F} d(s, t), \quad (3)$$

implicitly assuming every frontier has an equal-length optimal plan it can follow afterwards.

Gain maximization [2, 10]: The gain $G(s, t)$ is the increase in coverage obtained by following the shortest path from s to t . Its estimate is typically an upper bound [2, 10], but can be improved using learning-based approaches [8, 21, 23]. The exploration plan is obtained through

$$t^* = \arg \max_{t \in F} (\lambda \log G(s, t) - d(s, t)), \quad (4)$$

where λ determines the affinity of gain maximization relative to path cost. Determining an exploration path by maximizing gain corresponds to assuming that collecting more information on the path to t will lead to a shorter optimal path afterward.

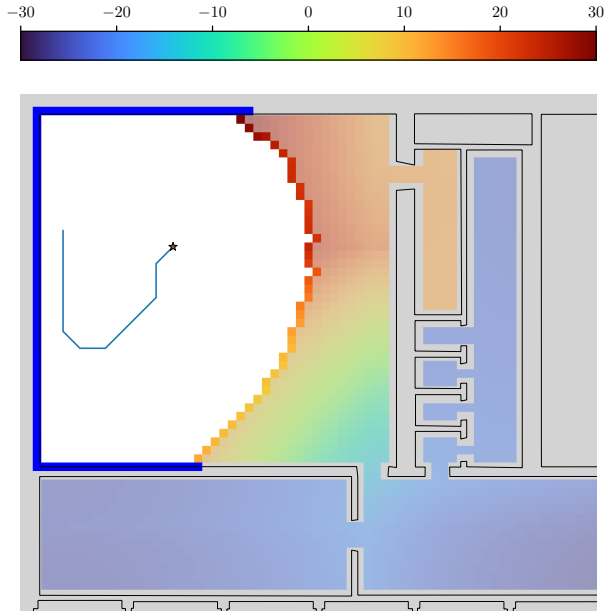


Fig. 2: Illustration of distance advantage in the beginning of exploration. The robot (star) preferentially explores frontiers (solid coloring) with higher distance advantage. It is heading towards a closed off room because it is nearer that region than it would be from most other places. By contrast, its distance to the corridor is higher than it would be elsewhere, repelling it from that region.

IV. DISTANCE ADVANTAGE

In order to design a quality-constrained exploration planner, a suitable heuristic has to be found. It has already been pointed out in [17] that unnecessary backtracking happens when a frontier state is not explored by the *last* time that the optimal plan comes close to it. Therefore, a suitable heuristic for quality-constrained exploration is one that minimizes backtracking by determining which frontiers are unlikely to be revisited.

An indication that a frontier might not be revisited is that it is isolated, i.e., it has high average distance to other reachable states R . Therefore, we propose to determine an exploration plan through

$$t^* = \arg \max_{t \in F} \left(\frac{1}{|R|} \sum_{s' \in R} d(t, s') - d(s, t) \right), \quad (5)$$

where $|R|$ is the number of reachable states. This heuristic prefers to visit frontiers that are near the robot, but on average isolated from reachable space, i.e., frontiers that the robot has a *distance advantage* to visit.

The reachable states are determined using the map and, if available, map predictions, which we show lead to an improvement in the estimate. Determining an exploration path through Eq. (5) corresponds to assuming that the cost of the optimal path after the frontier is lower if the frontier is more isolated, since leaving it behind would cause backtracking.

In Fig. 2, an example exploration state in an occupancy grid environment is shown, with unexplored cells colored according to their distance advantage and frontier cells highlighted with solid coloring.

Computing the distance advantage requires determining the shortest path distance from every frontier to every reachable state. In a graph-like environment, computing shortest path distances requires as many single-source shortest path problems as there are frontiers, and the computational cost for each scales with the size of the graph, i.e., the number of reachable states. Therefore, in order to keep the computational cost bounded regardless of the environment size, only frontiers and reachable states that are in a local window centered around the current state are considered. If the environment is a weighted graph, $d(t, s')$ is determined using Dijkstra's algorithm, while breadth-first search is used for unweighted graphs.

Frontiers that are outside the local window are not considered unless the local window has been fully explored, in which case the fallback planning chooses the nearest frontier outside the local window.

V. EXPERIMENTAL SETUP

In Sections VI and VII, experiments are conducted to evaluate distance advantage, nearest frontier, and information gain with respect to their quality-constrained exploration performance, and their sensitivity to predictions. The experiments are conducted in simulation in order to fairly compare the planning objectives. This section describes the simulation environment and other relevant implementation details.

A. Sensor & Mapping

The robot is simulated as a point-like sensor, with holonomic motion capabilities. The simulated sensor is a 360° laser range sensor, with 720 evenly spaced rays and a maximum range of 4.5m. As the robot moves and collects sensor scans, these are accumulated into an occupancy grid map that discretizes the environment into $25 \times 25 \text{ cm}^2$ cells and marks them as either free, occupied, or unknown. The scans are accumulated conservatively, such that 8-connected paths through free space in the map are guaranteed to be collision-free.

B. Localization & Path Execution

When the next state is chosen by the planner, a shortest 8-connected path is found through exhaustive search in the map using Dijkstra's algorithm. Since there are often many such shortest paths, the one which keeps the most distance from walls is chosen. As the path is executed, at each step in the map, a new sensor scan is obtained. The path terminates when the goal state is reached or when the sensor scan causes a map cell that was unknown to be marked as occupied or free.

Since the path is executed deterministically, the robot is perfectly localized with respect to the starting state.

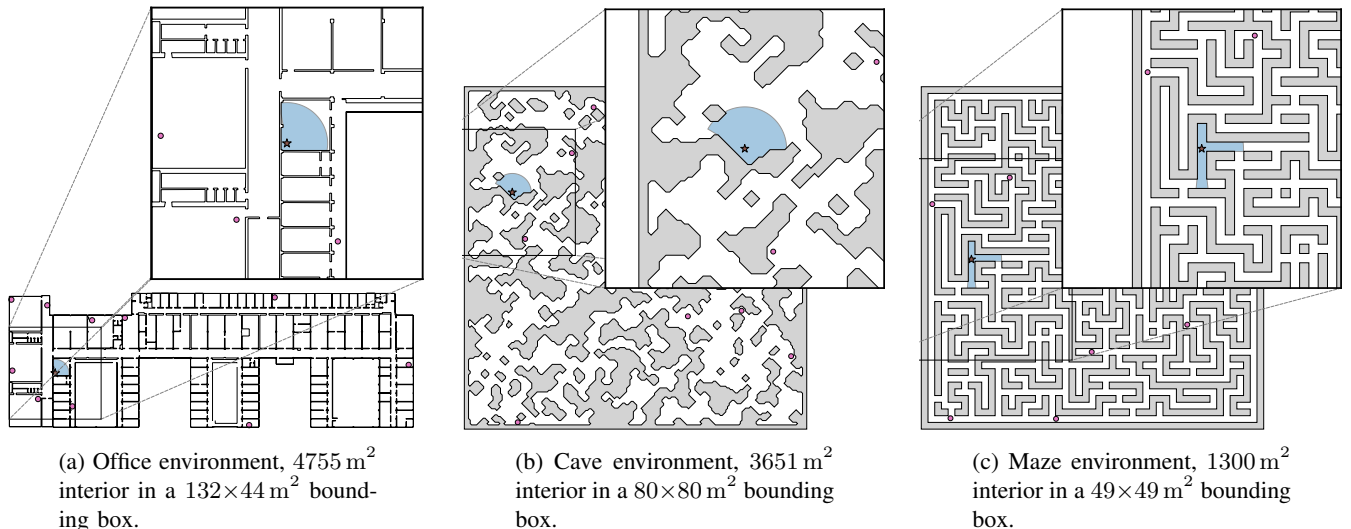


Fig. 3: Data is collected in three diverse environments: a large office from a real-world floor plan with both small cubicles and large lecture halls, a non-rectilinear cave environment with many small pockets, and a labyrinth-like maze with both shallow and deep dead-ends. Pink circles indicate starting locations, the light blue region depicts a sensor scan from the point of view of an example starting location indicated by the brown star polygon. The zoomed region is the same size as a local window for the planner.

C. Predictions

A map predictor, like the ones from [8, 14, 23], is assumed to be available and capable of providing map completions in a local window centered around the current state. The map predictor extends the map with the real environment and the local window size is 30×30 m².

VI. QUALITY-CONSTRAINED EVALUATION

For the purposes of autonomous exploration, environments can be characterized by their connectivity. One end of the spectrum is a perfectly tree-like environment, e.g., a maze without loops. In this case, it is practically irrelevant which frontier is chosen, as long as it is explored to completion, since the robot must always return to the branching point to pursue the next frontier. On the other end of the spectrum is a fully connected environment, where the unexplored space beyond every frontier eventually connects. Real environments are somewhere in the middle of this spectrum, with some branches interconnected and some branches dead ends, e.g., an office environment or a cave system.

The three heuristics are evaluated in three environments with different characteristics, presented in Fig. 3: a large office, with a wide variety of rooms-within-rooms, looping corridors, and wide-open spaces; a cave-like environment with high global connectivity, but low local connectivity; and a maze with a relatively small amount of loops, mostly consisting of dead-end branches of varying depth. The distance at completion for all heuristics in each environment is shown in Table I. Since nearest frontier is a special case of information gain (IG), with no preference for information, it is used as a reference.

In the office environment, which has the most complex connectivity of the three, distance advantage obtains an im-

TABLE I: Distance at completion for each method in each environment of Fig. 3. Data collected across 10 runs for each method/environment pair from different starting locations. The difference to nearest frontier is computed per starting location.

Method	Distance at Completion d_T (m)		
	Office	Cave	Maze
Nearest Frontier	1892.5 ± 50.5	1854.3 ± 106.1	1377.8 ± 36.0
Dist. Adv.	1578.1 ± 43.9	1652.1 ± 93.4	1249.5 ± 27.5
Δ Nearest Frontier	-303.6 ± 35.0	-194.8 ± 68.8	-128.3 ± 45.3
IG Max.	2330.3 ± 71.8	2313.2 ± 133.0	1514.1 ± 40.9
Δ Nearest Frontier	437.8 ± 89.8	407.5 ± 102.1	136.3 ± 60.9

provement of 16% over nearest frontier, while information gain is 23% worse. As the connectivity of the environments simplifies, the margin for improvement or degradation over nearest frontier decreases, since the simpler connectivity leads to there being more opportunities for sub-optimal policies to correct mistakes at a low cost. These results show that distance advantage consistently obtains shorter paths, with the difference being more noticeable in environments with more complex connectivity. Information gain is consistently outperformed by nearest frontier and, by extension, distance advantage.

A. Greediness

Information gain has previously been shown to lead to greedy behaviors, due to sacrificing long-term exploration performance for early short-term gains [9, 17]. In order to evaluate how greediness affects performance, the coverage $c(d)$ and frontier size $f(d)$ as functions of the distance traveled

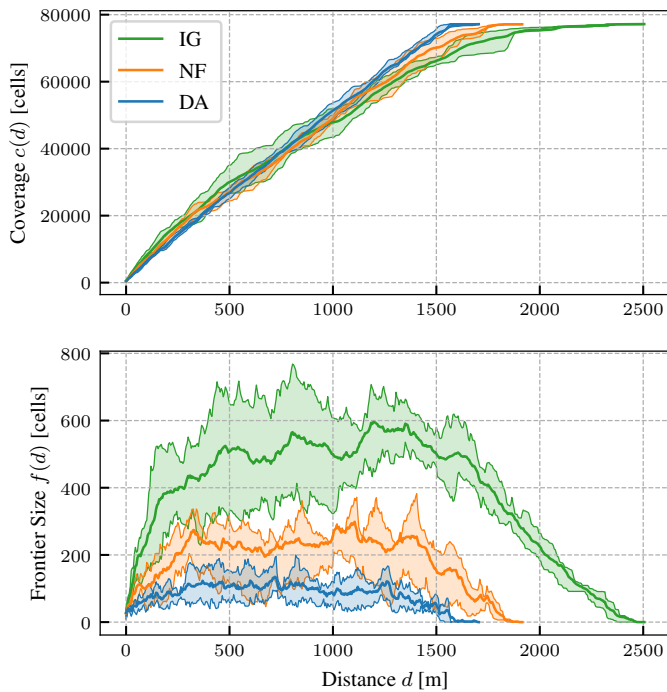


Fig. 4: Comparison of coverage $c(d)$ and total frontier size $f(d)$ as functions of distance traveled d . The shaded areas indicates an 80 % confidence interval, the solid line indicates the mean. Data collected across 10 runs for each method from different starting locations in the office environment.

d were considered, as is shown in Fig. 4.

There are nearly no differences in the coverage rate early on, indicating that the information gain maximization strategy is not successful in doing so. However, there is a large difference in the number of the frontiers, with nearest frontier quickly growing to double the amount of distance advantage, and gain maximization more than doubling nearest frontier. In accordance with [9], these outstanding frontiers represent a kind of debt to be paid, in the form of travel distance at the end of the run. This outstanding frontier debt explains why the coverage rate for information gain, and to a lesser extent nearest frontier, decreases as exploration progresses and ultimately leads to a longer path. By contrast, the frontier size for distance advantage remains approximately constant during most of the exploration run, allowing it to keep an almost constant coverage rate until the end of exploration.

B. Gain Maximization Affinity

When first introducing gain maximization for autonomous exploration, [10] already highlighted that lower affinity λ led to a more meticulous covering of the environment. We evaluate the effect of gain maximization affinity on path length with two gain estimators: naive, which assumes unknown space is non-occluding, and true, which has access to the true gain. The resulting completion distances are shown in Figure 1, clearly showing that higher affinity leads to longer paths and that the effect is worsened with more accurate estimates.

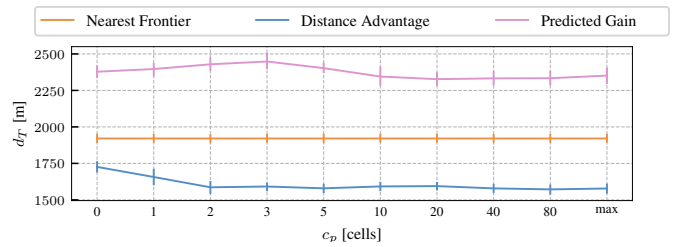


Fig. 5: The effect of prediction range c_p on completion distance d_T in the office environment. Data collected across 10 runs from different starting locations, for each method and prediction range. Error bars represent one standard deviation.

An interesting effect that, to the best of the authors' knowledge, has not been reported is that *negative* affinity leads to shorter paths than nearest frontier. The fact that preference for lower gain improves performance can be understood through the lens of isolation, since a low predicted gain amounts to predicting that a frontier region is soon to terminate. In that way, the preference for low gain is an *ad hoc* heuristic for preferring shallow frontiers, that are unlikely to be revisited by the optimal path since they do not continue deeper into unknown space.

VII. SENSITIVITY TO PREDICTIONS

The previous experiments all consider perfect predictions, given by an oracle that knows the true environment. However, in a real exploration scenario, these predictions would come either from prior knowledge of the environment, e.g., a floor plan, or from learning-based models [8, 14, 23]. In those cases, the predictions will not perfectly reflect the environment, and it is important to assess the sensitivity of each method to the predictions, with respect to both the amount of predicted information and its accuracy. Nearest frontier is completely insensitive to predictions and is presented as a baseline.

A. Prediction Range

The amount of information made available through the predictor could have a large influence in the exploration plan. While [8] showed that it is possible to predict $30 \times 30 \text{ m}^2$ windows from the occupancy map built by the robot in office environments, this might not be possible in more complex environments or too computationally demanding in some situations. Therefore, we examine the effect of the prediction range beyond the frontier, parameterized by the maximum straight line distance c_p in number of cells.

Figure 5 illustrates how different settings of c_p affect performance for all methods in the office environment. It can be seen that the ranking of the methods does not change in the absence of predictions ($c_p = 0$), and distance advantage continues to outperform nearest frontier. Most of the gain that distance advantage gets from the predictions is already realized at $c_p = 2$, which corresponds to a prediction range of only $\sim 50 \text{ cm}$ beyond the frontier. In agreement with the results reported by [9], information gain fails to take advantage of predictions.

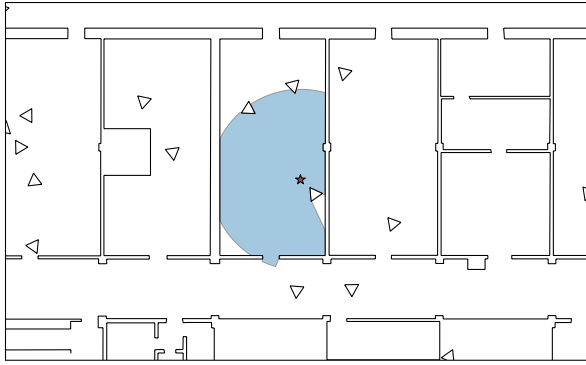


Fig. 6: An example of clutter for the office environment. The clutter consists of polygons, roughly 1 m, randomly placed in the reachable environment. A section of the environment does not get disconnected from the rest due to the clutter closing a passage, like the clutter of actual human environments.

TABLE II: Distance at completion when there is mismatch between the predictions and the environment, due to clutter. The environment and prediction clutter are independently sampled. Data collected across 10 runs for each method/environment/prediction tuple from different starting locations.

Method	Environment	Prediction	d_T (m)
Nearest Frontier	Clean	Clean	1892.5 ± 50.5
	Noise	Clean	2041.7 ± 57.4
	Noise	Noise	2054.0 ± 32.1
Dist. Adv.	Clean	Clean	1578.1 ± 43.9
	Noise	Clean	1782.3 ± 52.1
	Noise	Noise	1812.0 ± 14.8
IG Max.	Clean	Clean	2351.6 ± 85.2
	Noise	Clean	2567.8 ± 76.8
	Noise	Noise	2563.0 ± 85.6

B. Prediction Accuracy

A common mode of failure for predictions is non-structural elements of the environment such as furniture, human occupants, other robots, etc., since those are likely to move or be moved around the environment. In the office environment, the floor plan is fixed and predictions of it might be accessible, but elements like desks, chairs, etc., can be unpredictable. A different source of prior knowledge could also be an old map, in which case the mismatch could go the other way; the clutter in the map might no longer reflect the true state of the environment. To assess the impact of this kind of mismatch between predictions and the actual environment, clutter was generated for the office environment in Fig. 3a by randomly sampling triangles, illustrated in Fig. 6. The same set of triangles is used for every method, but different sets are used between starting locations, predictions and the true environment. The results of this evaluation are presented in Table II.

Clutter is a source of occlusion, so when the environment is cluttered it is harder to observe the environment from far away, causing all methods to produce longer paths. While nearest

frontier produces ~ 150 m longer paths, distance advantage and information gain degrade by ~ 200 m. This suggests that the 150 m increase can be explained due to the cluttering of the environment making observing it harder, with the inaccuracy in predictions accounting for the remaining 50 m. Importantly, the way in which the predictions are wrong, i.e., whether they do not contain clutter or they contain mismatched clutter, does not have a significant influence on the performance.

VIII. LIMITATIONS

The main focus of this work is to highlight how the current state-of-the-art autonomous exploration planning methods ultimately do not optimize the correct objectives. Although the proposed objective is well-motivated both by intuition and by empirical results, a more thorough theoretical analysis has the potential to yield even better formulations. Future work should attempt to characterize the fundamental limits of the proposed objective and improve upon it. Another aspect that warrants additional work is the computational scalability of estimating the optimization objective. Since the proposed objective requires solving a multi-source shortest path problem, the computational time scales with the size of the explored map. This issue was addressed by limiting the computation to a fixed-size local map, but future work should investigate whether it is possible to improve upon this solution.

The experimental results are limited to the minimal case, with no uncertainty and perfect mapping.

IX. CONCLUSION

In this work, we highlight the important differences between budget- and quality-constrained exploration, and address the inconsistencies observed in the autonomous exploration community. We investigate why traditional heuristics, such as information gain and nearest frontier, perform poorly in the quality-constrained paradigm, and propose a new heuristic for quality-constrained exploration. Since quality-constrained exploration is defined to be completed when the map is of sufficient quality, total gain is fixed; maximizing information gain of individual frontiers is therefore ultimately irrelevant as all the gain will be collected eventually. Therefore, the central problem in quality-constrained exploration planning is to determine the correct order in which to explore frontiers, such that the length of unnecessary detours is minimized.

We propose a heuristic, named *distance advantage*, that attempts to identify which frontiers have higher opportunity cost if missed, i.e., those more likely to require a detour later in exploration if they are not explored now, by estimating their average distance to other states. This heuristic is compared to nearest frontier and information gain, and it is shown to consistently explore the environments with shorter paths. Perhaps most importantly, among the evaluated heuristics, distance advantage is the only one to show improvements in performance as access to predictions improves, and is able to handle imperfect predictions. We believe these results clearly show the different nature of budget- and quality-constrained exploration, and indicate that further work should be done

in understanding the correct objective for quality-constrained exploration, and exploring the proposed objective.

REFERENCES

- [1] R. Bajcsy. Active perception. *Proceedings of the IEEE*, 1988.
- [2] Andreas Bircher, Mina Kamel, Kostas Alexis, Helen Oleynikova, and Roland Siegwart. Receding Horizon "Next-Best-View" Planner for 3D Exploration. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 2016.
- [3] Daniele Calisi, Alessandro Farinelli, Luca Iocchi, and Daniele Nardi. Autonomous exploration for search and rescue robots. *WIT Transactions on the Built Environment*, 2007.
- [4] Francis Colas, Srivatsa Mahesh, François Pomerleau, Ming Liu, and Roland Siegwart. 3d path planning and execution for search and rescue ground robots. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2013.
- [5] C. Connolly. The determination of next best views. In *1985 IEEE International Conference on Robotics and Automation Proceedings*, 1985.
- [6] Di Deng, Runlin Duan, Jiahong Liu, Kuangjie Sheng, and Kenji Shimada. Robotic exploration of unknown 2d environment using a frontier-based automatic-differentiable information gain measure. In *2020 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*. IEEE, 2020.
- [7] Daniel Duberg and Patric Jensfelt. UFOExplorer: Fast and Scalable Sampling-Based Exploration With a Graph-Based Planning Structure. *IEEE Robotics and Automation Letters*, 2022.
- [8] Ludvig Ericson and Patric Jensfelt. Beyond the frontier: Predicting unseen walls from occupancy grids by learning from floor plans. *IEEE Robotics and Automation Letters*, 2024.
- [9] Ludvig Ericson, Daniel Duberg, and Patric Jensfelt. Understanding Greediness in Map-Predictive Exploration Planning. In *2021 European Conference on Mobile Robots (ECMR)*. IEEE, 2021.
- [10] Héctor H. González-Baños and Jean-Claude Latombe. Navigation Strategies for Exploring Indoor Environments. *The International Journal of Robotics Research*, 2002.
- [11] Miroslav Kulich, Jan Faigl, and Libor Přebušil. On distance utility in the exploration task. In *2011 IEEE International Conference on Robotics and Automation*. IEEE, 2011.
- [12] Alberto Quattrini Li, Francesco Amigoni, and Nicola Basilico. Searching for Optimal Off-Line Exploration Paths in Grid Environments for a Robot with Limited Visibility. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2012.
- [13] Björn Lindqvist, Ali-Akbar Agha-Mohammadi, and George Nikolakopoulos. Exploration-rrt: A multi-objective path planning and exploration framework for unknown and unstructured environments. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2021.
- [14] Matteo Luperto, Federico Amadelli, Moreno Di Bernardino, and Francesco Amigoni. Mapping beyond what you can see: Predicting the layout of rooms behind closed doors. *Robotics and Autonomous Systems*, 2023.
- [15] Matteo Luperto, Marco Maria Ferrara, Giacomo Boracchi, and Francesco Amigoni. Estimating Map Completeness in Robot Exploration, 2024.
- [16] J. Maver and R. Bajcsy. Occlusions as a guide for planning the next view. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1993.
- [17] Stefan Obwald, Maren Bennewitz, Wolfram Burgard, and Cyrill Stachniss. Speeding-Up Robot Exploration by Exploiting Background Information. *IEEE Robotics and Automation Letters*, 2016.
- [18] Sammy Omari, Pascal Gohl, Michael Burri, Markus Achtelik, and Roland Siegwart. Visual industrial inspection using aerial robots. In *Proceedings of the 2014 3rd International Conference on Applied Robotics for the Power Industry*. IEEE, 2014.
- [19] R. Pito. A sensor-based solution to the "next best view" problem. In *Proceedings of 13th International Conference on Pattern Recognition*. IEEE, 1996.
- [20] Lukas Schmid, Michael Pantic, Raghav Khanna, Lionel Ott, Roland Siegwart, and Juan Nieto. An Efficient Sampling-based Method for Online Informative Path Planning in Unknown Environments. *IEEE Robotics and Automation Letters*, 2020.
- [21] Lukas Schmid, Chao Ni, Yuliang Zhong, Roland Siegwart, and Olov Andersson. Fast and Compute-Efficient Sampling-Based Local Exploration Planning via Distribution Learning. *IEEE Robotics and Automation Letters*, 2022.
- [22] Magnus Selin, Mattias Tiger, Daniel Duberg, Fredrik Heintz, and Patric Jensfelt. Efficient Autonomous Exploration Planning of Large-Scale 3-D Environments. *IEEE Robotics and Automation Letters*, 4, 2019.
- [23] Rakesh Shrestha, Fei-Peng Tian, Wei Feng, Ping Tan, and Richard Vaughan. Learned Map Prediction for Enhanced Mobile Robot Exploration. In *2019 International Conference on Robotics and Automation (ICRA)*, 2019.
- [24] Cyrill Stachniss, Giorgio Grisetti, and Wolfram Burgard. Information Gain-based Exploration Using Rao-Blackwellized Particle Filters. In *Robotics: Science and Systems I*. Robotics: Science and Systems Foundation, 2005.
- [25] Yuezhan Tao, Yuwei Wu, Beiming Li, Fernando Cladera, Alex Zhou, Dinesh Thakur, and Vijay Kumar. Seer: Safe efficient exploration for aerial robots using learning to predict information gain. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023.
- [26] Peter Whaite and Frank P Ferrie. Autonomous explo-

- ration: Driven by uncertainty. *Transactions on Pattern Analysis and Machine Intelligence*, 19(3):193–205, 1997.
- [27] Christian Witting, Marius Fehr, Rik Bähmann, Helen Oleynikova, and Roland Siegwart. History-aware autonomous exploration in confined environments using mavs. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018.
 - [28] B. Yamauchi. A frontier-based approach for autonomous exploration. In *Proceedings 1997 IEEE International Symposium on Computational Intelligence in Robotics and Automation CIRA'97. 'Towards New Computational Principles for Robotics and Automation'*, 1997.
 - [29] Jiajie Yu, Hao Shen, Jianyu Xu, and Tong Zhang. Echo: An efficient heuristic viewpoint determination method on frontier-based autonomous exploration for quadrotors. *IEEE Robotics and Automation Letters*, 2023.
 - [30] Yichen Zhang, Boyu Zhou, Luqi Wang, and Shaojie Shen. Exploration with global consistency using real-time re-integration and active loop closure. In *2022 International Conference on Robotics and Automation (ICRA)*, 2022.
 - [31] Boyu Zhou, Yichen Zhang, Xinyi Chen, and Shaojie Shen. FUEL: Fast UAV Exploration Using Incremental Frontier Structure and Hierarchical Planning. *IEEE Robotics and Automation Letters*, 2021.