Active Exploration and Mapping via Iterative Covariance Regulation over Continuous $SE(3)$ Trajectories

Shumon Koga Arash Asgharivaskasi Nikolay Atanasov

Abstract—This paper develops iterative Covariance Regulation (iCR), a novel method for active exploration and mapping for a mobile robot equipped with on-board sensors. The problem is posed as optimal control over the $SE(3)$ pose kinematics of the robot to minimize the differential entropy of the map conditioned the potential sensor observations. We introduce a differentiable field of view formulation, and derive iCR via the gradient descent method to iteratively update an open-loop control sequence in continuous space so that the covariance of the map estimate is minimized. We demonstrate autonomous exploration and uncertainty reduction in simulated occupancy grid environments.

I. INTRODUCTION

Simultaneous Localization and Mapping (SLAM) is a key research direction that has enabled robots to transition from controlled, structured, and fully known environments to operation in a priori unknown real-world conditions [1]. Many current SLAM techniques, however, remain passive in their utilization of sensor data. Active SLAM [2] is an extension of the SLAM problem which couples perception and control, aiming to acquire more information about the environment and reduce the uncertainty in the localization and mapping process. Active SLAM introduces unique challenges related to keeping the map and location estimation process accurate, and yet computing and propagating uncertainty over many potential sensing trajectories efficiently to select an informative one. Most of the literature in active exploration and mapping focuses on discrete [3], [4] or sampling-based [5] planning techniques. However, the trajectories and information collection process of the robot sensors evolve continuously over the $SE(3)$ space of sensor poses. As evidenced by successful applications of continuous control to exploration in reinforcement learning [6] and active target tracking [7], developing active SLAM techniques for continuous control is expected to reduce uncertainty more effectively and smoothly compared to discrete control techniques.

This paper develops a new forward-backward gradient computation technique to optimize multi-step control input sequences over the $SE(3)$ pose kinematics, leading to maximum uncertainty reduction. The core problem is formalized as $SE(3)$ trajectory optimization to minimize the differential entropy of the map state conditioned on the sequence of measurements obtained by an on-board sensor (e.g., Lidar or RGB-D camera). Assuming a Gaussian prior over the map state and its Bayesian evolution along a sensing trajectory, the differential entropy of the map is proportional to the log determinant of the covariance matrix at the final time. To ensure that the covariance matrix evolution is differentiable with respect to the control sequence, we introduce a new differentiable field of view formulation for the sensing model, providing a smooth transition from unobserved to observed space in the environment. Finally, the gradient of the objective function with respect to the multi-step control input sequences is computed explicitly and the control trajectory is updated via gradient descent.

Throughout this paper, we focus on occlusion-free planning by allowing the robot to pass through the occupied space in the environment, which can be implemented for 2-D mapping using an aerial robot and have been studied in literature [8]. While in [8] the gradient descent is applied via perturbation method to obtain an approximated gradient, our approach develops an explicit gradient formulation. We emphasize that our formulation is different from gradient descent methods that optimize the instantaneous sensing cost [9] because it considers the sensing performance over a long-horizon in an optimal control formulation.

Related Work: One of the earliest approaches for active exploration and mapping [10] is based on detecting and planning a shortest path to map frontiers (boundaries between explored and unexplored space). Frontier-based exploration is an efficient and effective method, prompting its widespread use in robotics [11], [12], [13]. Its purely geometric nature, however, is a limiting factor for considering sensing noise and map uncertainty reduction from a probabilistic inference perspective [14]. Information-theoretic planning is an alternative approach, which utilizes an information measure, such as mutual information or conditional entropy, to quantify and minimize the uncertainty in the map state. Information-theoretic mapping was first introduced by Elfes [15], and subsequently has been developed in many studies [16], [9], [5], [17], including applications to active SLAM [18], [19]. Evaluating information measures accompanies a high computational effort in general, which makes online planning challenging. Efficient computation methods have been proposed in [3] for Causchy-Schwarz quadratic mutual information (CSQMI), and in [4] for fast Shannon mutual information (FSMI). Instead of binary occupancy grid mapping, recent information-based active mapping techniques have considered truncated signed distance field (TSDF) maps [20] and multi-category semantic maps [21]. Existing methods are, however, limited to discrete control spaces, typically.

We gratefully acknowledge support from ARL DCIST CRA W911NF-17-2-0181 and ONR SAI N00014-18-1-2828. The authors are with the Department of Electrical and Computer Engineering, UC San Diego, 9500 Gilman Drive, La Jolla, CA, 92093-0411. {skoga,asghari,natanasov}@ucsd.edu. Supplementary materials can be found at https://shumon0423.github.io/IROS2021_webpage.
with a finite number of possible control inputs [22], [20], [3], [4], and have not considered optimal control formulations of the active mapping problem.

Optimal control has been intensively studied since the work of Bellman [23], which developed the well-known dynamic programming algorithm. Applying dynamic programming to continuous control spaces requires finite-dimensional approximations of the value function. Instead of a globally optimal control policy, a locally optimal state-control trajectory may be obtained via Pontryagin’s Maximum Principle (PMP) [24]. An iterative approach to solve PMP was originally developed by Kelley and Bryson [25], [26], via solving the adjoint system backward in time and updating the new control sequence, known as the adjoint method. As a second-order convergent algorithm, Differential Dynamic Programming (DDP) has been proposed to solve the value function and the control input via iteration of forward and backward path [27], [28]. To mitigate the computational burden in DDP caused by a tensor calculus, an “iterative LQR” (iLQR) method has been proposed in [29], which successfully eliminates the heavy computation by invoking the linearization only at the backward path.

Among the many optimization methods, gradient descent is perhaps one of the most recognized approaches. The gradient-based methods for optimizing an instantaneous cost have been applied to planning tasks of the mobile robots, such as in [30], [31] for active target tracking, in [32] for localizing 3-D features in an environment. For enabling the gradient calculus, a differentiable formulation of a field of view has been proposed in [33]. However, when applying the gradient descent to the optimal control with dynamical systems, the gradient with respect to multi-step control input sequences is needed, unlike optimizations of a static function or an instantaneous cost. The adjoint method has been widely utilized as an implicit gradient descent method for solving such an optimal control under multi-step control input sequences. However, the dynamical system in the adjoint method is supposed to belong to a vector field, which is not directly applicable to the SE(3) pose kinematics we consider. In addition to it, an optimal control approach for an information-based active exploration and mapping remains an open problem.

Contributions: We develop iterative Covariance Regulation (iCR), a new forward-backward gradient descent algorithm for finite-horizon optimal control of the covariance matrix of an active estimation process. To ensure that the Riccati equation governing the covariance evolution can be differentiated with respect to a control sequence, we introduce a new differentiable field of view formulation of the sensing model.

II. PROBLEM STATEMENT

This section formalizes active exploration and mapping as an optimal control problem.

A. Motion and Sensor Models

Consider a robot with pose \( T_k \in SE(3) \) at time \( t = t_k \in \mathbb{R}_+ \), where \( \{t_k\}_{k=0}^{K} \) for some \( K \in \mathbb{N} \) is an increasing sequence. The definition of pose and its discrete-time kinematic model are:

\[
T_k := \begin{bmatrix} R_k & p_k \\ 0_{3 \times 1} & 1 \end{bmatrix}, \quad T_{k+1} = T_k \exp(\tau \hat{u}_k), \tag{1}
\]

where \( p_k \in \mathbb{R}^3 \) is position, \( R_k \in SO(3) \) is orientation, and \( u_k = [v_k^T, \omega_k^T]^T \in \mathbb{R}^6 \) is a control input, consisting of linear velocity \( v_k \in \mathbb{R}^3 \) and angular velocity \( \omega_k \in \mathbb{R}^3 \). The hat operator \( (\cdot) : \mathbb{R}^6 \to se(3) \) maps vectors in \( \mathbb{R}^6 \) to the Lie algebra \( se(3) \) associated with the \( SE(3) \) Lie group [34].

The robot evolves in a 3-D environment described by a set \( \Omega \subset \mathbb{R}^3 \). A map \( m \in \mathbb{R}^n \) of the environment is defined by discretizing \( \Omega \) into \( n \in \mathbb{N} \) tiles, such as voxels or octants [35], and associating each tile \( j \in 1, \ldots, n \) with a position \( p(j) \in \mathbb{R}^3 \) and a mapped value \( m(j) \in \mathbb{R} \), such as occupancy or signed distance [36].

Let \( z_k \in \mathbb{R}^n \) be a measurement of the environment obtained by the robot at time \( t_k \), according to the following sensor model:

\[
z_k = h(T_k, m) + \eta_k, \quad \eta_k \sim \mathcal{N}(0, V(T, m)), \tag{2}
\]

where \( h : SE(3) \times \mathbb{R}^n \to \mathbb{R}^n \) is the measurement function which depends on both the pose and map states, and \( \eta_k \) is Gaussian sensing noise with zero mean and covariance matrix \( V : SE(3) \times \mathbb{R}^n \to \mathbb{R}^{n \times n} \). We present a differentiable formulation of the noise covariance matrix, which allows capturing field of view constraints, in Sec. III.

B. Active Exploration and Mapping

Assuming a Gaussian prior on the map state \( m \), we construct its posterior conditioned on a sequence of measurements \( z_{1:k} \) by means of the Extended Kalman Filter (EKF):

\[
m|z_{1:k} \sim \mathcal{N}(\mu_k, \Sigma_k), \tag{3}
\]

where the mean \( \mu_k \in \mathbb{R}^n \) and covariance \( \Sigma_k \in \mathbb{S}_{++}^{n \times n} \) satisfy:

\[
\begin{align*}
\mu_{k+1} &= \mu_k + \Sigma_k H_{k+1}^{-1} R_{k+1} (z_{k+1} - h(T_{k+1}, \mu_k)), \\
\Sigma_{k+1} &= (\Sigma_k^{-1} + M_k^{-1})^{-1}, \\
R_k &= H_k \Sigma_{k-1} H_k^T + V_k, \\
M_k &= H_k^T V_k^{-1} H_k, \\
H_{k+1} &= \frac{\partial h(T_{k+1}, m)}{\partial m} \bigg|_{m=\mu_k}, V_{k+1} = V(T_{k+1}, \mu_k).
\end{align*}
\]

The Extended Information Filter (EIF) [37], is an equivalent Bayesian filtering approach to the EKF, which parameterizes the Gaussian distributions in terms of an information matrix \( Y_k \in \mathbb{S}_{++}^{n \times n} \) and information mean \( \xi_k \in \mathbb{R}^n \) as follows:

\[
m|z_{1:k} \sim \mathcal{N}(Y_k^{-1} \xi_k, Y_k^{-1}). \tag{5}
\]

The update equations for \( \xi_k \) and \( Y_k \) are given by:

\[
\begin{align*}
\xi_{k+1} &= \xi_k + H_{k+1}^T R_{k+1}^{-1} (\nu_{k+1} + H_{k+1} Y_k^{-1} \xi_k), \\
\nu_{k+1} &= z_{k+1} - h(T_{k+1}, Y_k^{-1} \xi_k), \\
Y_{k+1} &= Y_k + M_k + 1.
\end{align*}
\]
We formulate active exploration and mapping as a motion planning problem aiming to minimize the differential entropy \(^1\) in the map state \(m\):

\[
\min_{u_0, \ldots, u_{K-1}} \mathbb{H}(m|z_{1:K}, T_{1:K}),
\]

subject to the motion model in (1) and the EIF update in (6).

While, in general, the active mapping problem above is a stochastic optimal control problem, owing to the Gaussian distribution of \(m|z_{1:k}\), it can be reduced to a deterministic optimal control problem, in which open-loop control policies are optimal [38]. In particular, the differential entropy in (7) is proportional to \(\log \det(Y_K^{-1})\) for a Gaussian distribution, and thus the problem can be reformulated as maximizing the following terminal reward function:

\[
r^U = \log \det(Y_K^{U}),
\]

where \(U = [u_0^\top, u_1^\top, \ldots, u_{K-1}^\top] \in \mathbb{R}^{6K}\) is a control sequence. With the help of the EIF, computing the information matrix in (6) recursively, the terminal reward function can be expressed with respect to the trajectory of the pose state \(T_{1:K}\). However, since the equations in (4) are evaluated at each updated mean both in the EKF and the EIF, which depend on the stochastic measurements, offline computation of the motion planning is not possible. Instead, we evaluate them at an initial estimate of the map \(\mu_0 \in \mathbb{R}^n\), by which \(H\) and \(V\) can be fixed offline. The problem is stated as follows.

**Problem**

Given a prior Gaussian distribution over the map state \(m \sim \mathcal{N}(\mu_0, \Sigma_0)\) with a mean \(\mu_0 \in \mathbb{R}^n\) and the covariance matrix \(\Sigma_0 \in \mathbb{S}_+^n\), obtain an open-loop control sequence \(U = [u_0^\top, \ldots, u_{K-1}^\top] \in \mathbb{R}^{6K}\) to solve the following deterministic optimal control problem:

\[
\max_{U} \log \det(Y_K^{U}),
\]

subject to

\[
T_{k+1} = T_k \exp(\tau u_k), \quad k = 0, \ldots, K-1,
\]

\[
Y_K = \Sigma_0^{-1} + \sum_{k=1}^{K} M(T_k),
\]

\[
M(T) = H(T)^\top V(T, \mu_0)^{-1} H(T),
\]

\[
H(T) = \frac{\partial h(T, m)}{\partial m} \bigg|_{m=\mu_0}.
\]

### III. Planning Method

This section develops “iterative Covariance Regulation (iCR)”, a new \(SE(3)\) trajectory optimization method to solve the active mapping optimal control problem.

#### A. Differentiable Field of View

For a given robot pose state \(T\) and the position \(p^{(j)}\) of the \(j\)-th map cell, we consider the robot body-frame coordinates of \(p^{(j)}\):

\[
q(T, p^{(j)}) = QT^{-1}p^{(j)},
\]

where the projection matrix \(Q\) and the homogeneous coordinates \(p^{(j)}\) are defined as:

\[
Q = \begin{bmatrix} I_3 & 0_{3 \times 1} \end{bmatrix} \in \mathbb{R}^{3 \times 4}, \quad p^{(j)} = \begin{bmatrix} p^{(j)}_1 \end{bmatrix} \in \mathbb{R}^4.
\]

Let the field of view of the robot be described as a fixed region \(F \subset \mathbb{R}^3\) in the body frame. We formulate the sensor noise matrix \(V\) so that the magnitude of the noise in the unobserved domain \(\Omega/F\) is approximately infinite. Hence, we formulate the measurement matrices as:

\[
V(T, m) = \text{diag}(\{V_{jj}(T, m)\})_{j=1}^{n} \in \mathbb{R}^{n \times n},
\]

where an approximate expression of \(V_{jj}(T, m)\) for \(j \in 1, \ldots, n\) is

\[
V_{jj}(T, m) \approx \begin{cases} \sigma^2, & \text{if } q(T, p^{(j)}) \in F, \\ \infty, & \text{if } q(T, p^{(j)}) \notin F, \end{cases}
\]

where \(\sigma \in \mathbb{R}_+\) is a standard deviation of the sensor noise in the observed domain. For the sake of enabling the gradient descent in the next section, we need the approximate expression in (17) by a continuous and differential function with respect to \(T\). For that reason, we introduce a Signed Distance Function (SDF) defined below for further analysis.

**Definition 1.** The signed distance function \(d : \mathbb{R}^3 \rightarrow \mathbb{R}\) associated with a set \(F \subset \mathbb{R}^3\) is:

\[
d(q, F) = \begin{cases} \min_{q \in F} \|q - q^*\|, & \text{if } q \in F, \\ \min_{q \notin F} \|q - q^*\|, & \text{if } q \notin F, \end{cases}
\]

where \(\partial F\) is the boundary of \(F\).

The conditions \(q(T, p^{(j)}) \in F\) and \(q(T, p^{(j)}) \notin F\) in (17) are then replaced by inequality conditions in terms of the SDF of \(F\). Incorporating this idea, the inverse of (17) is described as:

\[
V^{-1}_{jj}(T, m) = \begin{cases} 1, & \text{if } d(q(T, p^{(j)}), F) \leq 0, \\ 0, & \text{if } d(q(T, p^{(j)}), F) > 0. \end{cases}
\]

In order to approximate the right-hand side of (19) by a continuously differentiable function, we rely on a probit function [39], defined by the Gaussian CDF \(\Phi : \mathbb{R} \rightarrow [0, 1]::

\[
\Phi(x) = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{x}{\sqrt{2\kappa}} \right) \right],
\]

where \(\text{erf}(y) := \frac{2}{\sqrt{\pi}} \int_0^y e^{-t^2} dt\). Then, (20) satisfies \(\Phi(x) \approx 0\) for all \(x < 0\) (indeed, \(\Phi(0) = 0.002...\)), for all tuning parameter \(\kappa > 0\) which gives the smoothness of the function. Namely, we have \(\lim_{x \to -\infty} \Phi(x) = H_e(x)\), where \(H_e(x)\) is the Heaviside function, also known as the unit step function.

Using the function (20), we formulate (19) as:

\[
V^{-1}_{jj}(T, m) = \begin{cases} \frac{1}{\sigma^2} \left( 1 - \Phi(d(q(T, p^{(j)}), F)) \right), & \text{by a differentiable field of view is obtained and we can apply the gradient descent method next. The visualization of the differentiable field of view is given in Fig. 1 which depicts 2-D plot of (21) with respect to } p^{(j)}. \end{cases}
\]

Here, the field of view is set as a cone projected onto 2-D space, SDF of which is derived in Appendix.
B. Iterative Covariance Regulation

Let $u_k^{(i)} \in \mathbb{R}$ be the $i$-th element of the control vector $u_k \in \mathbb{R}^6$ at time $t = t_k$. Our proposed method for active exploration and mapping is presented in Alg. 1. Its derivation is based on gradient descent in the space of control sequences $U$ with respect to the terminal reward $r_U$.

**Proposition 1.** The gradient-descent update for solving active exploration (9)–(13) with differentiable field of view (16), (21) is given by

$$U \leftarrow U + \Gamma \frac{\partial r_U}{\partial U},$$

where $\Gamma \in \mathbb{R}^{6K \times 6K}$ is a step size matrix given as $\Gamma \defeq \text{diag}(\{\gamma_k^{(i)}\}_{i=1}^{K})$ with a positive constant $\gamma_k^{(i)} \in \mathbb{R}^+$, and each element in the gradient $\frac{\partial r_U}{\partial U} := \{\{\frac{\partial u_k^{(j)}}{\partial u_k^{(i)}}\}_{i=1}^{K}\}_{j=1}^{K}$ is

$$\frac{\partial r}{\partial u_k^{(i)}} = \sum_{s=k+1}^{K} \text{tr} \left( (Y_K^{-1})^{-1} \frac{\partial M(T_s)}{\partial u_k^{(i)}} \right), \quad \frac{\partial M(T_s)}{\partial u_k^{(i)}} = \text{diag} \left( \left\{ w \left( q(T_s, p^{(j)}), \frac{\partial q(T_s, p^{(j)})}{\partial u_k^{(i)}} \right) \right\}_{j=1}^{n} \right),$$

$$w \left( q, \frac{\partial q}{\partial u} \right) = -\frac{1}{\sigma^2} \Phi'(d(q, F)) \frac{\partial d(q, F)}{\partial q} \frac{\partial q}{\partial u},$$

$$\frac{\partial q(T_s, p^{(j)})}{\partial u_k^{(i)}} = -QT_s^{-1} \Lambda_s^{(i)} T_s^{-1} p^{(j)},$$

and $\Lambda_s^{(i)} \defeq \frac{\partial r}{\partial u_k^{(i)}}$ is obtained for $s \in \{k + 1, \ldots, K\}$ via:

$$\Lambda_s^{(i)} = \begin{cases} T_k \frac{\partial \exp(\tau u_k)}{\partial u_k^{(i)}}, & \text{if } s = k + 1 \\ \Lambda_s^{(i)} \exp(\tau u_{s-1}), & \text{if } s \in \{k+2, \ldots, K\} \end{cases} \quad (26)$$

**Proof.** Taking the gradient of the reward in (8) with the help of (11), one can obtain

$$\frac{\partial r_U}{\partial u_k^{(i)}} = \text{tr} \left( (Y_K^{-1})^{-1} \frac{\partial Y_U}{\partial u_k^{(i)}} \right) = \text{tr} \left( (Y_U^{-1}) \left( \sum_{s=1}^{K} \frac{\partial M(T_s)}{\partial u_k^{(i)}} \right) \right) = \sum_{s=k+1}^{K} \text{tr} \left( (Y_U^{-1}) \left( \frac{\partial M(T_s)}{\partial u_k^{(i)}} \right) \right) \quad (27)$$

Substituting (12), (16), (21) and using the chain rule, we get (23)–(24).

In (24), $\Phi'(d(q, F))$ and $\frac{\partial d(q, F)}{\partial q}$ can be obtained by the derivative of (20) and the gradient of the SDF for a given
field of view, and thus it remains to compute the gradient \( \frac{\partial p_i}{\partial u_k} \). Taking the gradient of (14) with respect to \( u_k \) yields (25), by defining \( \Lambda_{s_k} := \frac{\partial \tau_k}{\partial u_k} \). The initial condition and the update law for \( \Lambda_k \) can be derived by taking the gradient of the pose kinematics in (1), which leads to (26), and hence we can deduce Proposition 1.

For the computation of \( \frac{\partial \exp(\tau u_k)}{\partial u_k} \) in (26), we use the following lemma (see [34] for a proof).

**Lemma 1.** For \( u_k = [u_k^{(1)}, \ldots, u_k^{(6)}]^T \in \mathbb{R}^6 \), it holds that

\[
\frac{\partial \exp(\tau u_k)}{\partial u_k} = \tau (\mathcal{J}_L(\tau u_k)e_i)^\top \exp(\tau u_k),
\]

where \( e_i \in \mathbb{R}^6 \) is the \( i \)-th unit vector, and \( \mathcal{J}_L(\cdot) \) is the left Jacobian of \( SE(3) \).

### IV. Evaluation

We examine the efficacy of the proposed method in various simulated scenarios using 2-D occupancy maps of real-world environments. The robot moves according to the motion model in (1) as \( SE(2) \) dynamics, and its on-board sensor measures the occupancy of neighboring map cells inside its field of view. The field of view \( F \) is set as an isosceles triangle with height 3 [m] the angle between the two legs equal to \( 60^\circ \) (see Fig. 1 for visualization and Appendix for derivation). The mapping process is performed as follows. First, we define the initial distribution of \( \mathbf{m} \) as a multivariate Gaussian \( \mathcal{N}(\mu_0, \Sigma_0) \), where \( \mu_0 = 0 \in \mathbb{R}^n \) and \( \Sigma_0 = \text{diag}(100) \in \mathbb{R}^{n \times n} \). We encode measurements as \( z^{(j)} = -1 \) for free cells and \( z^{(j)} = 1 \) for occupied cells, both of which are for all \( j \)-th cells inside the field of view. Next, we construct a measurement vector \( z_k \in \mathbb{R}^n \), the \( j \)-th element of which has the measured occupancy value \( z^{(j)} \) for all \( j \)-th cells inside the field of view. All other elements have a prior mean \( \mu_k^{(j)} \) for all the \( j \)-th cells outside the field of view. The constructed measurement vector \( z_k \) is passed to EKF to get the updated mean \( \mu_k \) and the information matrix \( Y_k \).

In order to visualize the occupancy map, we apply a threshold function \( g: \mathbb{R} \to \{-1, 1\} \) which returns \( g(x) = 1 \) for all \( x > 0 \) and \( g(x) = -1 \) for all \( x < 0 \), to each element in the updated mean \( \mu_k \). We set the smoothing factor \( \kappa \) in (20) to 0.5. One should note that a small \( \kappa \) eliminates the impact of the space outside of the field of view on the policy optimization, while selecting a large \( \kappa \) leads to neglecting the influence of the region inside the field of view. Both extreme cases result in vanishing gradients during the optimal control computation. Furthermore, since we seek for an open-loop policy, the planning horizon \( K \) can be chosen arbitrarily large. However, a longer planning horizon tends to require more iterations for iCR in order to ensure an informative trajectory. Therefore, due to computational power constraints, we choose \( K = 5 \), which based on our experiments can reach to an informative trajectory in less than 10 iterations. In each planning phase, we start by initializing a trajectory with constant linear velocity \( v_k = 1.5 \), and random angular velocity \( w_k \sim U[-\pi/10, \pi/10] \) for \( k = 0, \ldots, K - 1 \) and use iCR at each planning in order to optimize the trajectory.

Fig. 2 shows snapshots of the robot’s exploration and mapping using iCR in two environments, Environment A (top) and Environment B (bottom), at times \( k = 51 \) (left), \( k = 126 \) (center), and \( k = 476 \) (right), respectively. The occupancy maps are obtained by applying a binary threshold to the mean of the EKF. Green and red paths represent the initial trajectory and the optimal trajectory from iCR, respectively. The map data is available from C. Stachniss’s lab [40].

**Fig. 2:** Snapshots of the robot’s exploration and mapping using iCR in two environments, Environment A (top) and Environment B (bottom), at times \( k = 51 \) (left), \( k = 126 \) (center), and \( k = 476 \) (right), respectively. The occupancy maps are obtained by applying a binary threshold to the mean of the EKF. Green and red paths represent the initial trajectory and the optimal trajectory from iCR, respectively. The map data is available from C. Stachniss’s lab [40].
This paper developed *iterative Covariance Regulation* (iCR), a new forward-backward gradient descent algorithm for active exploration and mapping over continuous $SE(3)$ trajectories. Active mapping was posed as a finite-horizon deterministic optimal control problem, aiming to maximize the information matrix of the map at the terminal time conditioned on the potential measurement data from the on-
Lemma 2.

focus on accounting for occlusion in the differentiable field $F$ on board sensors. Our approach utilizes a differentiable field of view for the robot sensing model and developed a trajectory optimization approach by computing the gradient of the terminal reward function with respect to the multi-step control sequence over the $SE(3)$ manifold. The proposed method was demonstrated in the context of active occupancy grid mapping in numerical experiments. The paper considered occlusion-free and obstacle-free planning, which is clearly a limitation for real-world deployment. Future work will focus on accounting for occlusion in the differentiable field of view formulation and for obstacles in the iCR trajectory optimization to achieve safe active exploration and mapping.

APPENDIX

Signed Distance Function of a 2-D Cone

Suppose that the field of view is a cone which has a vertex at the origin, the circular base at $x = h$, and the angles $[−\psi, \psi]$ from the vertex at $x = y$ plane. Then, the field $F_c \subset \mathbb{R}^3$ is given by

$$F_c = \{(x, y, z) \in \Omega | y^2 + z^2 \leq (\tan(\psi) x)^2, x \in [0, h]\}.$$

Let $F_{c,x=\psi}$ be the cone region projected onto 2-D space of $z = \psi$ for a value $\psi \in \mathbb{R}$. Then, the Signed Distance Function for $F_{c,x=\psi}$ is given in the following lemma.

Lemma 2. Signed Distance Function $d : \mathbb{R}^2 \times F_{c,x=\psi} \to \mathbb{R}$ of the cone projected onto 2-D space $z = 0$ is given by

$$d(q, F_{c,x=\psi}) = \left\{ \begin{array}{ll} \frac{a_1 q + b_1}{|a_1|}, & \text{if } q \in D_1, \\ \frac{1}{|q - q_1|}, & \text{if } q \in P_1, \end{array} \right. \quad (29)$$

where

$$p_x = \frac{h}{1 + \sin(\psi)},$$

$$D_1 = \{(x, y) \in \Omega | y \in (l_1(x), l_1(x)), \forall x \in (-\infty, h)\},$$

$$D_2 = \{(x, y) \in \Omega | y \in (-l_1(x), -l_1(x)), \forall x \in (-\infty, h)\},$$

$$D_3 = \{(x, y) \in \Omega | y \in (l_1(x), l_1(x)), \forall x \in [p_x, \infty)\},$$

$$P_1 = \{(x, y) \in \Omega | y \in (l_1(x), +\infty), \forall x \in \mathbb{R}\},$$

$$P_2 = \{(x, y) \in \Omega | y \in (-\infty, -l_1(x)), \forall x \in \mathbb{R}\},$$

$$P_3 = \{(x, y) \in \Omega | y \in (-l_1(x), l_1(x)), \forall x \in \mathbb{R} \leq 0\},$$

$$l_1(x) = \left\{ \begin{array}{ll} \frac{1}{\tan(\psi)} x, & \forall x \in (-\infty, 0), \\ 0, & \forall x \in [0, p_x], \end{array} \right. \right.$$  

$$l_2(x) = \left\{ \begin{array}{ll} \tan(\pi/4 + \psi/2) x - \frac{h}{\cos(\psi)}, & \forall x \in [p_x, h], \\ h \tan(\psi), & \forall x \in [h, \infty), \end{array} \right. \right.$$  

$$a_1 = \left[ \begin{array}{c} -1 \\ \frac{1}{\tan(\psi)} \end{array} \right], a_2 = \left[ \begin{array}{c} -1 \\ \frac{1}{\tan(\psi)} \end{array} \right], a_3 = \left[ \begin{array}{c} 1 \\ 0 \end{array} \right], \quad (30)$$

$$b_1 = 0, \quad b_2 = 0, \quad b_3 = -h, \quad (31)$$

$$q_1 = \left[ \begin{array}{c} h \\ h \tan(\psi) \end{array} \right], q_2 = \left[ \begin{array}{c} -h \\ -h \tan(\psi) \end{array} \right], q_3 = \left[ \begin{array}{c} 0 \\ 0 \end{array} \right].$$

Proof. The separated domains $D_i$ and $P_i$ for $i \in \{1, 2, 3\}$ are defined so that if $q \in D_i$ then the closest point in $\partial F$ lies in a line $(p \in \mathbb{R}^2 | p + b_i = 0)$, and if $q \in P_i$ then the closest point in $\partial F$ is a point $q_i$ which is an edge of the projected cone. The normed distance $d^+ : \mathbb{R}^2 \to \mathbb{R}_+$ between a point $q$ and a line $L := \{p \in \mathbb{R}^2 | p + b = 0\}$ is known as

$$d^+(q, L) = \frac{a_1^T q + b_1}{|a_1|}.$$

There are two choices for a pair of $(a, b)$ for any given line, in which the other pair has opposite sign to one. By setting the pairs $(a_i, b_i)$ for $i \in \{1, 2, 3\}$ as (30–31), one can deduce that if $q \in F_c$ then $a_1 q + b_1 \leq 0$ and if $q \notin F_c$ then $a_1^T q + b_1 > 0$, which leads to the first line in (29). Indeed, all the other domains $P_i$ is outside the cone $F_c$, thus if $q \in P_i$ then the SDF has a positive distance between the two points $q$ and $q_i$, which yields the second line of (29). \qed

REFERENCES

[1] C. Cadena, L. Carluccio, C. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, and J. J. Leonard, “Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age,” IEEE Transactions on Robotics, vol. 32, no. 6, pp. 1309–1332, 2016.

[2] J. Delmerico, E. Mueggler, J. Nitsch, and D. Scaramuzza, “Active autonomous aerial exploration for ground robot path planning,” IEEE Robotics and Automation Letters, vol. 2, no. 2, pp. 664–671, 2017.

[3] K. Charrow, S. Liu, V. Kumar, and N. Michael, “Information-theoretic mapping using cauchy-schwarz quadric mutual information,” in IEEE International Conference on Robotics and Automation (ICRA), 2015, pp. 4791–4798.

[4] Z. Zhang, T. Henderson, S. Karaman, and V. Sze, “Fsmi: Fast computation of shannon mutual information for information-theoretic mapping,” The International Journal of Robotics Research, vol. 39, no. 9, pp. 1155–1177, 2020.

[5] G. A. Hollinger and G. S. Sukhatme, “Sampling-based robotic information gathering algorithms,” The International Journal of Robotics Research, vol. 33, no. 9, pp. 1271–1287, 2014.

[6] G. Dalal, K. Dvijotham, M. Vecerik, T. Hester, C. Paduraru, and Y. Tassa, “Safe exploration in continuous action spaces,” arXiv preprint arXiv:1801.08757, 2018.

[7] K. Zhou and S. I. Roumeliotis, “Multirobot active target tracking with combinations of relative observations,” IEEE Transactions on Robotics, vol. 27, no. 4, pp. 678–695, 2011.

[8] V. Indelman, L. Carluccio, and F. Dellaert, “Planning in the continuous domain: A generalized belief space approach for autonomous navigation in unknown environments,” The International Journal of Robotics Research, vol. 34, no. 7, pp. 849–882, 2015.

[9] B. Grocholsky, “Information-theoretic control of multiple sensor platforms,” Ph.D. dissertation, Univ. of Sydney, 2002 [Online]. Available: http://www.acfr.usyd.edu.au, 2002.

[10] B. Yamauchi, “A frontier-based approach for autonomous exploration,” in IEEE International Symposium on Computational Intelligence in Robotics and Automation, 1997, pp. 146–151.

[11] W. Burgard, M. Moors, C. Stachniss, and F. E. Schneider, “Coordinated multi-robot exploration,” IEEE Transactions on robotics, vol. 21, no. 3, pp. 376–386, 2005.

[12] D. Holz, N. Basilio, F. Amigoni, and S. Behnke, “Evaluating the efficiency of frontier-based exploration strategies,” in International Symposium on Robotics, 2010, pp. 1–8.

[13] H. Oleynikova, M. Burri, Z. Taylor, J. Nieto, R. Siegwart, and E. Galceran, “Continuous-time trajectory optimization for online UAV replanning,” in IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2016, pp. 5332–5339.

[14] S. Thrun, “Probabilistic algorithms in robotics,” AI Magazine, vol. 21, no. 4, pp. 93–93, 2000.

[15] A. Elfes, “Robot navigation: Integrating perception, environmental constraints and task execution within a probabilistic framework,” in International Workshop on Reasoning with Uncertainty in Robotics. Springer, 1995, pp. 91–130.
[16] S. J. Moorehead, R. Simmons, and W. L. Whittaker, “Autonomous exploration using multiple sources of information,” in *IEEE International Conference on Robotics and Automation (ICRA)*, vol. 3, 2001, pp. 3098–3103.

[17] E. Sommerville and I. Reid, “Information-theoretic active scene exploration,” in *IEEE Conference on Computer Vision and Pattern Recognition*, 2008, pp. 1–7.

[18] L. Carlone, J. Du, M. K. Ng, B. Bona, and M. Indri, “Active SLAM and exploration with particle filters using Kullback-Leibler divergence,” *Journal of Intelligent & Robotic Systems*, vol. 75, no. 2, pp. 291–311, 2014.

[19] N. Atanasov, J. Le Ny, K. Daniilidis, and G. J. Pappas, “Decentralized active information acquisition: Theory and application to multi-robot SLAM,” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 4775–4782.

[20] K. Saulnier, N. Atanasov, G. J. Pappas, and V. Kumar, “Information theoretic active exploration in signed distance fields,” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 4080–4085.

[21] A. Asgharivaskasi and N. Atanasov, “Active Bayesian multi-class mapping from range and semantic segmentation observation,” *arXiv preprint arXiv:2101.01831*, 2021.

[22] 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 (RSS)*, 2019.

[23] H. J. Kelley, “Gradient theory of optimal flight paths,” *American Rocket Society Journal*, vol. 30, no. 10, pp. 947–954, 1960.

[24] F. Morbidi and G. L. Mariottini, “Active target tracking and cooperative localization for teams of aerial vehicles,” *IEEE Transactions on Control Systems Technology*, vol. 21, no. 5, pp. 1694–1707, 2012.

[25] H. Wei, W. Lu, P. Zhu, G. Huang, J. Leonard, and S. Ferrari, “Optimized visibility motion planning for target tracking and localization,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2014, pp. 76–82.

[26] V. Murali, I. Spasojevic, W. Guerra, and S. Karaman, “Perception-aware trajectory generation for aggressive quadrotor flight using differential flatness,” in *American Control Conference (ACC)*, 2019, pp. 3936–3943.

[27] T. D. Barfoot, *State estimation for robotics*. Cambridge University Press, 2017.

[28] C. M. Bishop, *Pattern recognition and machine learning*. springer, 2006.

[29] C. Stachniss and D. Hahnel, *StachnissLab*. [Online]. Available: https://www.ipb.uni-bonn.de/datasets/