

# 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]. Its purely geometric nature, however, is a limiting factor for considering sensing noise and map uncertainty reduction from a probabilistic inference perspective [12]. 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 [13], and subsequently has been developed in many studies [14], [9], including applications to active SLAM [15], [16]. Evaluating information measures accompanies a high computational effort in general, which makes online planning challenging. Efficient computation methods have been proposed in [3] for Cauchy-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 [17] and multi-category semantic maps [18]. Existing methods are, however, limited to discrete control spaces, typically with a finite number of possible control inputs [17], [3], [4], and have not considered

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, aasghari, natanasov}@ucsd.edu.

optimal control formulations of the active mapping problem.

Optimal control has been intensively studied since the work of Bellman [19], 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) [20]. 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 [21]. To mitigate the computational burden in DDP caused by a tensor calculus, an "iterative LQR" (iLQR) method has been proposed in [22], 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 [23] for active target tracking, in [24] 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 [25]. 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 Regularization* (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. The set  $SO(3)$  denotes the special orthogonal group defined by  $SO(3) := \{R \in \mathbb{R}^{3 \times 3} | R^\top R = I, \det R = 1\}$ , and  $SE(3)$  denotes a set of matrix in the special Euclidian group defined by  $SE(3) := \left\{ T := \begin{bmatrix} R & \mathbf{p} \\ \mathbf{0}^\top & 1 \end{bmatrix} \middle| R \in SO(3), \mathbf{p} \in \mathbb{R}^3 \right\} \subset \mathbb{R}^{4 \times 4}$ .

### 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:

$$T_k := \begin{bmatrix} R_k & \mathbf{p}_k \\ \mathbf{0}_{3 \times 1}^\top & 1 \end{bmatrix}, \quad T_{k+1} = T_k \exp(\tau \hat{\mathbf{u}}_k), \quad (1)$$

where  $\mathbf{p}_k \in \mathbb{R}^3$  is position,  $R_k \in SO(3)$  is orientation, and  $\mathbf{u}_k = [\mathbf{v}_k^\top, \boldsymbol{\omega}_k^\top]^\top \in \mathbb{R}^6$  is a control input, consisting of linear velocity  $\mathbf{v}_k \in \mathbb{R}^3$  and angular velocity  $\boldsymbol{\omega}_k \in \mathbb{R}^3$ . The hat operator  $(\hat{\cdot}) : \mathbb{R}^6 \rightarrow se(3)$  maps vectors in  $\mathbb{R}^6$  to the Lie algebra  $se(3)$  associated with the  $SE(3)$  Lie group [26]. The discrete-time kinematic model in (1) is derived from the solution to continuous-time kinematic model  $\dot{T} = T\hat{\mathbf{u}}$ .

The robot evolves in a 3-D environment described by a set  $\Omega \subset \mathbb{R}^3$ . A map  $\mathbf{m} \in \mathbb{R}^n$  of the environment is defined by discretizing  $\Omega$  into  $n \in \mathbb{N}$  tiles, such as voxels or octants [27], and associating each tile  $j \in 1, \dots, n$  with a position  $\mathbf{p}^{(j)} \in \mathbb{R}^3$  and a mapped value  $\mathbf{m}^{(j)} \in \mathbb{R}$ , such as occupancy or signed distance [28].

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

$$\mathbf{z}_k = \mathbf{h}(T_k, \mathbf{m}) + \boldsymbol{\eta}_k, \quad \boldsymbol{\eta}_k \sim \mathcal{N}(0, V(T, \mathbf{m})), \quad (2)$$

where  $\mathbf{h} : SE(3) \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  is the measurement function which depends on both the pose and map states, and  $\boldsymbol{\eta}_k$  is Gaussian sensing noise with zero mean and covariance matrix  $V : SE(3) \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ . 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  $\mathbf{m}$ , we construct its posterior conditioned on a sequence of measurements  $\mathbf{z}_{1:k}$  by means of the Extended Kalman Filter (EKF):

$$\mathbf{m} | \mathbf{z}_{1:k} \sim \mathcal{N}(\boldsymbol{\mu}_k, \Sigma_k), \quad (3)$$

where the mean  $\boldsymbol{\mu}_k \in \mathbb{R}^n$  and covariance  $\Sigma_k \in \mathbb{S}_{>0}^{n \times n}$  satisfy:

$$\begin{aligned} \boldsymbol{\mu}_{k+1} &= \boldsymbol{\mu}_k + \Sigma_k H_{k+1}^\top R_{k+1}^{-1} (\mathbf{z}_{k+1} - h(T_{k+1}, \boldsymbol{\mu}_k)), \\ \Sigma_{k+1} &= (\Sigma_k^{-1} + M_{k+1})^{-1}, \\ R_k &= H_k \Sigma_{k-1} H_k^\top + V_k, \quad M_k := H_k^\top V_k^{-1} H_k, \\ H_{k+1} &= \left. \frac{\partial h(T_{k+1}, \mathbf{m})}{\partial \mathbf{m}} \right|_{\mathbf{m}=\boldsymbol{\mu}_k}, \quad V_{k+1} = V(T_{k+1}, \boldsymbol{\mu}_k). \end{aligned} \quad (4)$$

The Extended Information Filter (EIF) is an equivalent Bayesian filtering approach to the EKF, which parameterizes the Gaussian distributions in terms of an information matrix  $Y_k \in \mathbb{S}_{>0}^{n \times n}$  and information mean  $\boldsymbol{\xi}_k \in \mathbb{R}^n$  as follows:

$$\mathbf{m} | \mathbf{z}_{1:k} \sim \mathcal{N}(Y_k^{-1} \boldsymbol{\xi}_k, Y_k^{-1}). \quad (5)$$

The update equations for  $\boldsymbol{\xi}_k$  and  $Y_k$  are given by:

$$\begin{aligned} \boldsymbol{\xi}_{k+1} &= \boldsymbol{\xi}_k + H_{k+1}^\top R_{k+1}^{-1} (\boldsymbol{\nu}_{k+1} + H_{k+1} Y_k^{-1} \boldsymbol{\xi}_k), \\ \boldsymbol{\nu}_{k+1} &= \mathbf{z}_{k+1} - h(T_{k+1}, Y_k^{-1} \boldsymbol{\xi}_k), \\ Y_{k+1} &= Y_k + M_{k+1}. \end{aligned} \quad (6)$$

We formulate active exploration and mapping as a motion planning problem aiming to minimize the differential entropy<sup>1</sup> in the map state  $\mathbf{m}$ :

$$\min_{\mathbf{u}_0, \dots, \mathbf{u}_{K-1}} \mathbb{H}(\mathbf{m} | \mathbf{z}_{1:K}, T_{1:K}), \quad (7)$$

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  $\mathbf{m} | \mathbf{z}_{1:k}$ , it can be reduced to a deterministic optimal control problem, in which open-loop control policies are optimal [29]. 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:

$$r^{\mathbf{U}} = \log \det(Y_K^{\mathbf{U}}), \quad (8)$$

where  $\mathbf{U} = [\mathbf{u}_0^\top, \mathbf{u}_1^\top, \dots, \mathbf{u}_{K-1}^\top]^\top \in \mathbb{R}^{6K}$  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  $\boldsymbol{\mu}_0 \in \mathbb{R}^n$ , 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  $\mathbf{m} \sim \mathcal{N}(\boldsymbol{\mu}_0, \Sigma_0)$  with a mean  $\boldsymbol{\mu}_0 \in \mathbb{R}^n$  and the covariance matrix  $\Sigma_0 \in \mathbb{S}_{>0}^{n \times n}$ , obtain an open-loop control sequence  $\mathbf{U} = [\mathbf{u}_0^\top, \dots, \mathbf{u}_{K-1}^\top]^\top \in \mathbb{R}^{6K}$  to solve the following deterministic optimal control problem:

$$\max_{\mathbf{U}} \log \det(Y_K^{\mathbf{U}}), \quad (9)$$

subject to

$$T_{k+1} = T_k \exp(\tau \hat{\mathbf{u}}_k), \quad k = 0, \dots, K-1, \quad (10)$$

$$Y_K = \Sigma_0^{-1} + \sum_{k=1}^K M(T_k), \quad (11)$$

$$M(T) = H(T)^\top V(T, \boldsymbol{\mu}_0)^{-1} H(T), \quad (12)$$

$$H(T) = \left. \frac{\partial h(T, \mathbf{m})}{\partial \mathbf{m}} \right|_{\mathbf{m}=\boldsymbol{\mu}_0}. \quad (13)$$

### 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  $\mathbf{p}^{(j)}$  of the  $j$ -th map cell, we consider the robot body-frame coordinates of  $\mathbf{p}^{(j)}$ :

$$\mathbf{q}(T, \mathbf{p}^{(j)}) = QT^{-1} \underline{\mathbf{p}}^{(j)}, \quad (14)$$

<sup>1</sup>The differential entropy of a continuous random variable  $Y$  with probability density function  $p$  is defined as  $\mathbb{H}(Y) := -\int p(y) \log p(y) dy$ .

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

$$Q = [I_3 \quad \mathbf{0}_{3 \times 1}] \in \mathbb{R}^{3 \times 4}, \quad \underline{\mathbf{p}}^{(j)} = \begin{bmatrix} \mathbf{p}^{(j)} \\ 1 \end{bmatrix} \in \mathbb{R}^4. \quad (15)$$

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

$$V(T, \mathbf{m}) = \text{diag}(\{V_{jj}(T, \mathbf{m})\}_{j=1}^n) \in \mathbb{R}^{n \times n}, \quad (16)$$

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

$$V_{jj}(T, \mathbf{m}) \approx \begin{cases} \sigma^2, & \text{if } \mathbf{q}(T, \mathbf{p}^{(j)}) \in \mathcal{F}, \\ \infty, & \text{if } \mathbf{q}(T, \mathbf{p}^{(j)}) \notin \mathcal{F}, \end{cases} \quad (17)$$

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  $\mathcal{F} \subset \mathbb{R}^3$  is:

$$d(\mathbf{q}, \mathcal{F}) = \begin{cases} -\min_{\mathbf{q}^* \in \partial \mathcal{F}} \|\mathbf{q} - \mathbf{q}^*\|, & \text{if } \mathbf{q} \in \mathcal{F}, \\ \min_{\mathbf{q}^* \in \partial \mathcal{F}} \|\mathbf{q} - \mathbf{q}^*\|, & \text{if } \mathbf{q} \notin \mathcal{F}, \end{cases} \quad (18)$$

where  $\partial \mathcal{F}$  is the boundary of  $\mathcal{F}$ .

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

$$V_{jj}^{-1}(T, \mathbf{m}) \approx \frac{1}{\sigma^2} \begin{cases} 1, & \text{if } d(\mathbf{q}(T, \mathbf{p}^{(j)}), \mathcal{F}) \leq 0, \\ 0, & \text{if } d(\mathbf{q}(T, \mathbf{p}^{(j)}), \mathcal{F}) > 0. \end{cases} \quad (19)$$

In order to approximate the right-hand side of (19) by a continuously differentiable function, we rely on a probit function [30], defined by the Gaussian CDF  $\Phi : \mathbb{R} \rightarrow [0, 1]$ :

$$\Phi(x) = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{x}{\sqrt{2\kappa}} - 2 \right) \right], \quad (20)$$

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,  $\Phi(0) = 0.002\dots$ ), for all tuning parameter  $\kappa > 0$  which gives the smoothness of the function. Namely, we have  $\lim_{\kappa \rightarrow +0} \Phi(x) = He(x)$ , where  $He(x)$  is the Heaviside function. Using (20), we formulate (19) as:

$$V_{jj}^{-1}(T, \mathbf{m}) = \frac{1}{\sigma^2} \left( 1 - \Phi(d(\mathbf{q}(T, \mathbf{p}^{(j)}), \mathcal{F})) \right), \quad (21)$$

by which 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  $\mathbf{p}^{(j)}$ . Here, the field of view is set as a cone projected onto 2-D space.

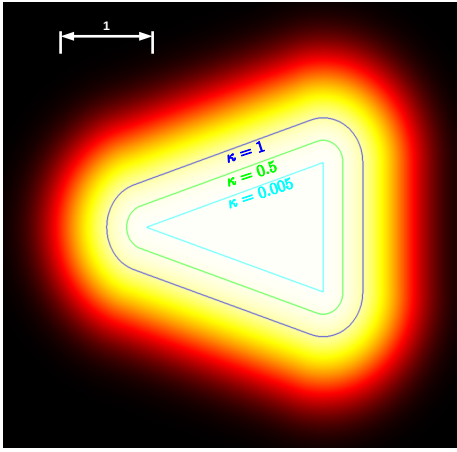


Fig. 1: 2-D plot of a differentiable field of view (21) specified by SDF of a 2-D cone with parameters  $\sigma = 1$  and  $\kappa = 0.5$ . The closed lines depict the level sets of  $V_{jj}^{-1} = 0.98$  for different choices of  $\kappa$ . As  $\kappa$  gets larger,  $V_{jj}^{-1}$  varies smoothly.

### B. Iterative Covariance Regulation

Let  $u_k^{(i)} \in \mathbb{R}$  be the  $i$ -th element of the control vector  $\mathbf{u}_k \in \mathbb{R}^6$  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  $\mathbf{U}$  with respect to the terminal reward  $r^{\mathbf{U}}$ .

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

$$\mathbf{U} \leftarrow \mathbf{U} + \Gamma \frac{\partial r^{\mathbf{U}}}{\partial \mathbf{U}}, \quad (22)$$

where  $\Gamma \in \mathbb{R}^{6K \times 6K}$  is a step size matrix given as  $\Gamma := \text{diag}(\{\{\gamma_k^{(i)}\}_{i=1}^6\}_{k=0}^{K-1})$  with a positive constant  $\gamma_k^{(i)} \in \mathbb{R}_+$ , and each element in the gradient  $\frac{\partial r^{\mathbf{U}}}{\partial \mathbf{U}} := [\{\{\frac{\partial r^{\mathbf{U}}}{\partial u_k^{(i)}}\}_{i=1}^6\}_{k=0}^{K-1}]$  is

$$\frac{\partial r^{\mathbf{U}}}{\partial u_k^{(i)}} = \sum_{s=k+1}^K \text{tr} \left( (Y_K^{\mathbf{U}})^{-1} \frac{\partial M(T_s)}{\partial u_k^{(i)}} \right), \quad (23)$$

$$\frac{\partial M(T_s)}{\partial u_k^{(i)}} = \text{diag} \left( \left\{ w \left( \mathbf{q}(T_s, \mathbf{p}^{(j)}), \frac{\partial \mathbf{q}(T_s, \mathbf{p}^{(j)})}{\partial u_k^{(i)}} \right) \right\}_{j=1}^n \right),$$

$$w \left( \mathbf{q}, \frac{\partial \mathbf{q}}{\partial u} \right) = -\frac{1}{\sigma^2} \Phi'(d(\mathbf{q}, \mathcal{F})) \frac{\partial d(\mathbf{q}, \mathcal{F})}{\partial \mathbf{q}} \frac{\partial \mathbf{q}}{\partial u}, \quad (24)$$

$$\frac{\partial \mathbf{q}(T_s, \mathbf{p}^{(j)})}{\partial u_k^{(i)}} = -Q T_s^{-1} \Lambda_s^{(i)} T_s^{-1} \underline{\mathbf{p}}^{(j)}, \quad (25)$$

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

$$\Lambda_s^{(i)} = \begin{cases} T_k \frac{\partial \exp(\tau \hat{\mathbf{u}}_k)}{\partial u_k^{(i)}}, & \text{if } s = k+1 \\ \Lambda_{s-1}^{(i)} \exp(\tau \hat{\mathbf{u}}_{s-1}), & \text{if } s \in \{k+2, \dots, K\}. \end{cases} \quad (26)$$

*Proof.* Taking the gradient of the reward in (8) with the help of (11) yields (23). Substituting (12), (16), (21) and using

### Algorithm 1: iterative Covariance Regulation (iCR)

**Data:** Initial robot pose  $T_0 \in SE(3)$ , initial map covariance  $\Sigma_0 \in \mathbb{S}_{>0}^{n \times n}$ , map cell positions  $\mathbf{p}^{(j)}$  for  $j \in \{1, \dots, n\}$ , and initial control sequence  $\mathbf{u}_{0:K-1}$

Initialize  $M \leftarrow 0_{n \times n}$ ,  $\frac{\partial M}{\partial u} \leftarrow 0_{n \times n}$ .

**repeat**

Set  $Y \leftarrow \Sigma_0^{-1}$ .

**for**  $k \leftarrow 0$  **to**  $K-1$  **do**

$T_{k+1} \leftarrow T_k \exp(\tau \hat{\mathbf{u}}_k)$ .

**for**  $j \leftarrow 1$  **to**  $n$  **do**

$M_{jj} \leftarrow H(T)^{\top} V^{-1}(T) H(T)$  by RHS of (21) using  $[\mathbf{p}^{(j)}, T_{k+1}]$ .

**end**

Update  $Y \leftarrow Y + M$ .

**end**

**for**  $k \leftarrow K-1$  **to**  $0$  **do**

Initialize  $\frac{\partial r}{\partial u^{(i)}} \leftarrow 0$  for all  $i \in \{1, \dots, 6\}$ .

**for**  $s \leftarrow k+1$  **to**  $K$  **do**

**for**  $i \leftarrow 1$  **to**  $6$  **do**

**if**  $s = k+1$  **then**

Set  $\Lambda \leftarrow T_k \frac{\partial \exp(\tau \hat{\mathbf{u}}_k)}{\partial u_k^{(i)}}$  by  $[T_k, \mathbf{u}_k]$ .

**else**

Update  $\Lambda \leftarrow \Lambda \exp(\tau \hat{\mathbf{u}}_{s-1})$ .

**end**

**for**  $j \leftarrow 1$  **to**  $n$  **do**

$[\mathbf{q}, \frac{\partial \mathbf{q}}{\partial u}] \leftarrow (14), (25)$  by  $[\mathbf{p}^{(j)}, T_k, \Lambda]$

$\frac{\partial M}{\partial u}_{jj} \leftarrow$  RHS of (24) by  $[\mathbf{q}, \frac{\partial \mathbf{q}}{\partial u}]$ .

**end**

$\frac{\partial r}{\partial u^{(i)}} \leftarrow \frac{\partial r}{\partial u^{(i)}} + \text{tr}(Y^{-1} \frac{\partial M}{\partial u})$ .

**end**

$u_k^{(i)} \leftarrow u_k^{(i)} + \alpha^{(i)} \frac{\partial r}{\partial u^{(i)}}$  for all  $i \in \{1, \dots, 6\}$ .

**end**

the chain rule, we get (24). In (24),  $\Phi'(d(\mathbf{q}, \mathcal{F}))$  and  $\frac{\partial d(\mathbf{q}, \mathcal{F})}{\partial \mathbf{q}}$  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  $\frac{\partial \mathbf{q}}{\partial u_k^{(i)}}$ . Taking the gradient of (14) with respect to  $u_k^{(i)}$  yields (25). The initial condition and the update law for  $\Lambda_s^{(i)}$  can be derived by taking the gradient of the pose kinematics in (1), which leads to (26), and hence Proposition 1 is deduced.  $\square$

$\frac{\partial \exp(\tau \hat{\mathbf{u}}_k)}{\partial u_k^{(i)}}$  in (26) is computed by using the following lemma (see [26] for a proof).

**Lemma 1.** For  $\mathbf{u}_k = [u_k^{(1)}, \dots, u_k^{(6)}]^{\top} \in \mathbb{R}^6$ , it holds that

$$\frac{\partial \exp(\tau \hat{\mathbf{u}}_k)}{\partial u_k^{(i)}} = \tau (\mathcal{J}_L(\tau \mathbf{u}_k) \mathbf{e}_i)^{\wedge} \exp(\tau \hat{\mathbf{u}}_k), \quad (27)$$

where  $\mathbf{e}_i \in \mathbb{R}^6$  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  $\mathcal{F}$  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). 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  $\mu_0 = \mathbf{0} \in \mathbb{R}^n$  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  $\mathbf{z}_k \in \mathbb{R}^n$ , 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  $\mu_{k-1}^{(j)}$  for all the  $j$ -th cells outside the field of view. The constructed measurement vector  $\mathbf{z}_k$  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  $g : \mathbb{R} \rightarrow \{-1, 1\}$  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. 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  $k = 0, \dots, K - 1$  and use iCR at each planning in order to optimize the trajectory.

Fig. 2 shows snapshots of the robot’s exploration and mapping via iCR for two environments. Environment A has dimensions 14.46[m]  $\times$  20.04[m]. Environment B has dimensions 17.37[m]  $\times$  17.43[m]. The robot tries to explore the center of both environments initially, where there is an abundance of unobserved map cells. Then, approximately after  $k = 300$ , the optimal trajectories move towards the edges of the environment, in which there remain some patches of unexplored regions. It is worth mentioning that iCR does not explicitly plan to visit unexplored regions; the observed behavior is only a result of minimizing the uncertainty, which leads to either refining the visited regions or discovering unexplored areas of the map.

Fig. 3 illustrates the final occupancy map of Environment A at  $k = 925$  and a heat map showing the diagonal elements in the information matrix obtained via four different exploration strategies: (a) iCR, (b) iCR + frontier, (c) frontier, and (d) random. Fig. 4 shows analogous plots performed in Environment B. As shown in Fig. 3 (a) and 4 (a), via iCR, the center of the information map becomes brighter than the edges, meaning that the cells closer to the center receive more robot observations. Throughout multiple times of planning, we observe that the initial random path (green) brings insignificant uncertainty reduction since the trajectory overlaps the observed region highly. On the other hand, the

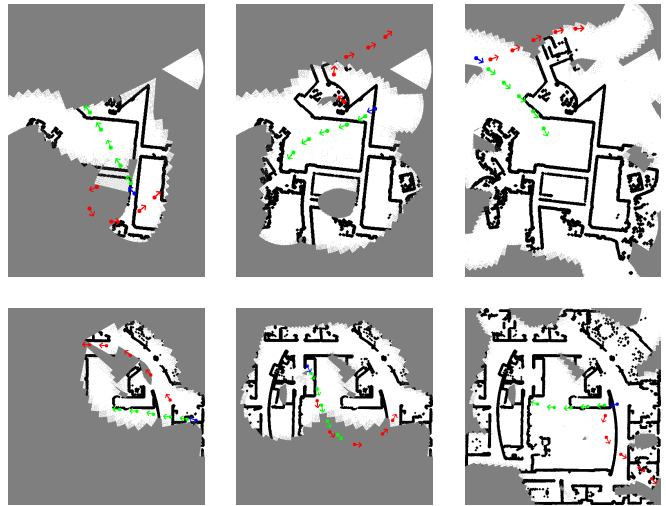


Fig. 2: Snapshots of the robot’s exploration and mapping using iCR in two environments, Environment A (top) and Environment B (bottom), at times  $k = 51$  (left),  $k = 126$  (center), and  $k = 476$  (right), respectively. The occupancy maps are obtained by applying a binary threshold to the mean of 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 [31].

optimal path from iCR (red) tends to visit unexplored areas, by which the information quantity from EKF becomes larger. This shows the advantage of iCR-based path planning which finds the informative trajectory over the continuous control space even by starting from a random control sequence without using any high-level heuristics, such as biasing the trajectory to visit frontiers as in [10], [18], [3]. Furthermore, our simulations show that adding heuristics to the trajectory initialization in iCR can lead to more efficient exploration. For the demonstration, we set the initial trajectory during each planning step to face the boundary between observed map cells and the largest unexplored region. In this case, iCR can be considered as an augmentation over the path from frontier-based exploration [10]. As illustrated in both Fig. 3 (b) and Fig. 4 (b), the exploration via iCR with frontier receives relatively more observations in the edge map cells, compared to the exploration via iCR with random initialization does as shown in Fig. 3 (a) and Fig. 4 (a), by which we can deduce that providing a frontier location to the initial trajectory can enhance the map certainty by enabling more homogeneous exploration.

We test two baseline exploration strategies to compare with iCR. The first baseline is the same as frontier-based exploration [10], where we plan a trajectory facing the frontier between the explored area and the largest unexplored region. As shown in Fig. 3 (c) and Fig. 4 (c), frontier-based exploration leaves most of edge pixels unexplored. By investigating through frontier-based exploration episodes, we observe that the robot shortly starts to oscillate between edges of the map since the largest unexplored patch per-

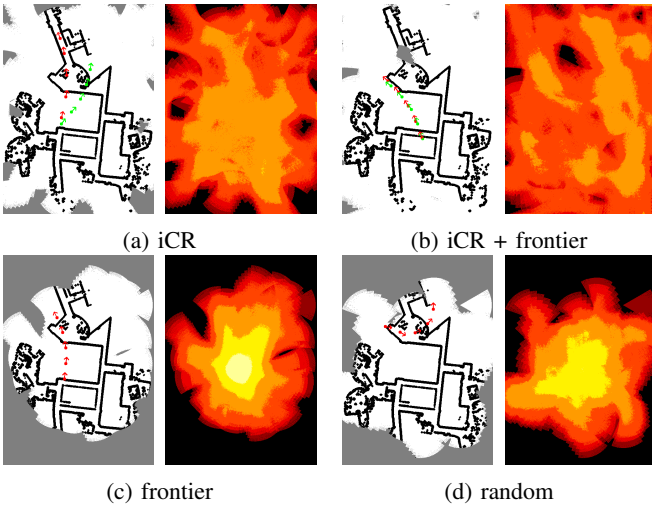


Fig. 3: Final map of Environment A at  $k = 925$  for four different exploration strategies (a)–(d). For each pair in (a)–(d), the left image shows the occupancy map obtained by applying a threshold function to the mean of the EKF-based mapping, and the right image depicts the information map, obtained from the inverse covariance of the EKF.

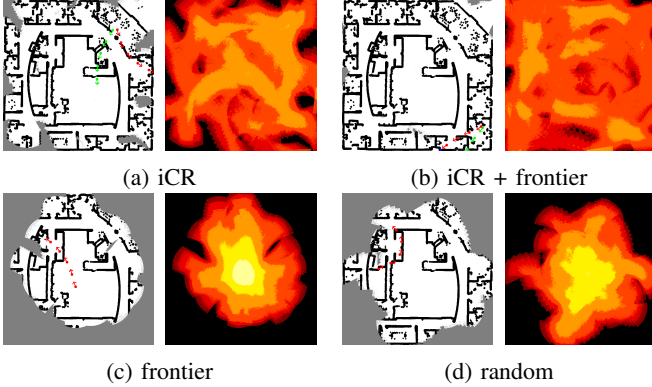


Fig. 4: Analogous demonstrations to Fig. 3 conducted for Environment B.

petually changes among different corners of the simulation environment, and hence the robot fails to utilize time steps in order to efficiently explore the environment. The second baseline is a random policy under constant linear velocity  $v_k = 1.5$  and random angular velocity  $w_k \sim U[-\pi/3, \pi/3]$ . Similar to the frontier-based policy, the resulting final map under the random policy leaves a large part of the environment unexplored, which is expected since the random policy does not aim to minimize the uncertainty of the map  $\mathbf{m}$ .

Fig. 5 shows the reward function  $\log |Y_k|$ , the certainty in the map  $\mathbf{m}$ , over time for the experiments shown in Fig. 3 (left) and 4 (right). In both plots in Fig. 5, we observe that both iCR (blue) and iCR + frontier (green) have a larger increase of the information than the information acquired by frontiers (cyan) and random policy (red) for all time. As mentioned earlier, the robot tends to start exploration by visiting information-rich regions, such as the center of the

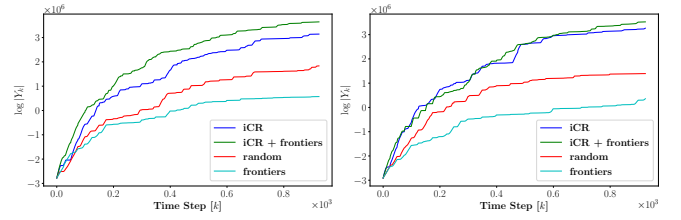


Fig. 5: Evolution of the log-determinant of the information matrix over time steps. The left and right plots correspond to the exploration episodes of Fig. 3 and Fig. 4, respectively.

map. This results in a steep climb in the reward function during the initial phase of exploration ( $k \leq 300$ ). As the robot makes more observations, the slope of the reward curve decreases, which corresponds to the refining phase where there are only a few unexplored map cells and the rest of the map has been discovered with high certainty.

## V. CONCLUSION

This paper developed *iterative Covariance Regulation* (iCR), a new forward-backward gradient descent algorithm for active exploration and mapping over continuous  $SE(3)$  trajectories. 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 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. Comparison to other more recently developed exploration methods, such as CSQMI [3] and FSMI [4], and experimental demonstration will be studied in future work as well.

## 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  $\mathcal{F}_c \subset \mathbb{R}^3$  is given by  $\mathcal{F}_c = \{(x, y, z) \in \Omega | y^2 + z^2 \leq (\tan(\psi)x)^2, x \in [0, h]\}$ . Let  $\mathcal{F}_{c,z=\zeta}$  be the cone region projected onto 2-D space of  $z = \zeta$  for a value  $\zeta \in \mathbb{R}$ . Then, the Signed Distance Function for  $\mathcal{F}_{c,z=0}$  is given in the following lemma.

**Lemma 2.** *Signed Distance Function  $d : \mathbb{R}^2 \times \mathcal{F}_{c,z=0} \rightarrow \mathbb{R}$  of the cone projected onto 2-D space  $z = 0$  is given by*

$$d(\mathbf{q}, \mathcal{F}_{c,z=0}) = \begin{cases} \frac{\mathbf{a}_i^\top \mathbf{q} + b_i}{\|\mathbf{a}_i\|}, & \text{if } \mathbf{q} \in \mathcal{D}_i, \\ \|\mathbf{q} - \mathbf{q}_i\|, & \text{if } \mathbf{q} \in \mathcal{P}_i. \end{cases} \quad (28)$$

where  $p_x^* = \frac{h}{1 + \sin(\psi)}$ ,  $\mathcal{D}_1 = \{(x, y) \in \Omega | y \in (L_1(x), \bar{L}_1(x)), \forall x \in (-\infty, h]\}$ ,  $\mathcal{D}_2 = \{(x, y) \in \Omega | y \in (-\bar{L}_1(x), -L_1(x)), \forall x \in (-\infty, h]\}$ ,  $\mathcal{D}_3 = \{(x, y) \in \Omega | y \in (-L_1(x), L_1(x)), \forall x \in [p_x^*, \infty)\}$ ,  $\mathcal{P}_1 = \{(x, y) \in$

$\Omega|y \in (\bar{l}_1(x), +\infty), \forall x \in \mathbb{R}\}, \mathcal{P}_2 = \{(x, y) \in \Omega|y \in (-\infty, -\bar{l}_1(x)), \forall x \in \mathbb{R}\}, \mathcal{P}_3 = \{(x, y) \in \Omega|y \in (-\bar{l}_1(x), \bar{l}_1(x)), \forall x \in \mathbb{R}_{\leq 0}\},$

$$l_1(x) = \begin{cases} -\frac{1}{\tan(\psi)}x, & \forall x \in (-\infty, 0], \\ 0, & \forall x \in [0, p_x^*], \\ \tan(\pi/4 + \psi/2)x - \frac{h}{\cos(\psi)}, & \forall x \in [p_x^*, h], \\ h \tan(\psi), & \forall x \in [h, \infty), \end{cases}$$

$$\bar{l}_1(x) = \begin{cases} -\frac{x-h}{\tan(\psi)} + h \tan(\psi), & \forall x \in (-\infty, h], \\ h \tan(\psi), & \forall x \in [h, \infty), \end{cases}$$

$$a_1 = \begin{bmatrix} -1 \\ \frac{1}{\tan(\psi)} \end{bmatrix}, a_2 = \begin{bmatrix} -1 \\ -\frac{1}{\tan(\psi)} \end{bmatrix}, a_3 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad (29)$$

$$b_1 = 0, \quad b_2 = 0, \quad b_3 = -h, \quad (30)$$

$$q_1 = \begin{bmatrix} h \\ h \tan(\psi) \end{bmatrix}, q_2 = \begin{bmatrix} h \\ -h \tan(\psi) \end{bmatrix}, q_3 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

*Proof.* The separated domains  $\mathcal{D}_i$  and  $\mathcal{P}_i$  for  $i \in \{1, 2, 3\}$  are defined so that if  $\mathbf{q} \in \mathcal{D}_i$  then the closest point in  $\partial\mathcal{F}$  lies in a line  $\{\mathbf{p} \in \mathbb{R}^2 | \mathbf{a}_i^\top \mathbf{p} + b_i = 0\}$ , and if  $\mathbf{q} \in \mathcal{P}_i$  then the closest point in  $\partial\mathcal{F}$  is a point  $\mathbf{q}_i$  which is an edge of the projected cone. The normed distance  $d^+ : \mathbb{R}^2 \rightarrow \mathbb{R}_+$  between a point  $\mathbf{q}$  and a line  $\mathcal{L} := \{\mathbf{p} \in \mathbb{R}^2 | \mathbf{a}^\top \mathbf{p} + b = 0\}$  is known as  $d^+(\mathbf{q}, \mathcal{L}) = \frac{|\mathbf{a}^\top \mathbf{q} + b|}{\|\mathbf{a}\|}$ . There are two choices for a pair of  $(\mathbf{a}, b)$  for any given line, in which the other pair has opposite sign to one. By setting the pairs  $(\mathbf{a}_i, b_i)$  for  $i \in \{1, 2, 3\}$  as (29)–(30), one can deduce that if  $\mathbf{q} \in \mathcal{F}_c$  then  $\mathbf{a}_i^\top \mathbf{q} + b_i \leq 0$  and if  $\mathbf{q} \notin \mathcal{F}_c$  then  $\mathbf{a}_i^\top \mathbf{q} + b_i > 0$ , which leads to the first line in (28). Indeed, all the other domains  $\mathcal{P}_i$  is outside the cone  $\mathcal{F}_c$ , thus if  $\mathbf{q} \in \mathcal{P}_i$  then the SDF has a positive distance between the two points  $\mathbf{q}$  and  $\mathbf{q}_i$ , which yields the second line of (28).  $\square$

## REFERENCES

- [1] C. Cadena, L. Carlone, H. 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] B. Charrow, S. Liu, V. Kumar, and N. Michael, “Information-theoretic mapping using cauchy-schwarz quadratic 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. Carlone, 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] S. Thrun, “Probabilistic algorithms in robotics,” *Ai Magazine*, vol. 21, no. 4, pp. 93–93, 2000.
- [13] 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.
- [14] 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.
- [15] 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.
- [16] 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.
- [17] 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.
- [18] A. Asgharivaskasi and N. Atanasov, “Active Bayesian multi-class mapping from range and semantic segmentation observation,” *arXiv preprint arXiv:2101.01831*, 2021.
- [19] R. Bellman, “On the theory of dynamic programming,” *National Academy of Sciences of the United States of America*, vol. 38, no. 8, p. 716, 1952.
- [20] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *Mathematical theory of optimal processes*. John Wiley & Sons, 1963.
- [21] D. H. Jacobson and D. Q. Mayne, *Differential dynamic programming*. North-Holland, 1970.
- [22] W. Li and E. Todorov, “Iterative linear quadratic regulator design for nonlinear biological movement systems,” in *International Conference on Informatics in Control, Automation and Robotics*, 2004, pp. 222–229.
- [23] 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.
- [24] E. Cristofalo, E. Montijano, and M. Schwager, “Vision-based control for fast 3-d reconstruction with an aerial robot,” *IEEE Transactions on Control Systems Technology*, vol. 28, no. 4, pp. 1189–1202, 2019.
- [25] 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.
- [26] T. D. Barfoot, *State estimation for robotics*. Cambridge University Press, 2017.
- [27] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, “Octomap: An efficient probabilistic 3D mapping framework based on octrees,” *Autonomous robots*, vol. 34, no. 3, pp. 189–206, 2013.
- [28] H. Oleynikova, Z. Taylor, M. Fehr, R. Siegwart, and J. Nieto, “Voxblox: Incremental 3D Euclidean signed distance fields for on-board MAV planning,” in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, pp. 1366–1373.
- [29] N. Atanasov, J. Le Ny, K. Daniilidis, and G. J. Pappas, “Information acquisition with sensing robots: Algorithms and error bounds,” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2014, pp. 6447–6454.
- [30] C. M. Bishop, *Pattern recognition and machine learning*. springer, 2006.
- [31] C. Stachniss and D. Haehnel, *StachnissLab*. [Online]. Available: <https://www.ipb.uni-bonn.de/datasets/>