2023 IEEE International Conference on Robotics and Automation (ICRA 2023)
May 29 - June 2, 2023. London, UK

Observability-Aware Active Extrinsic Calibration of Multiple Sensors

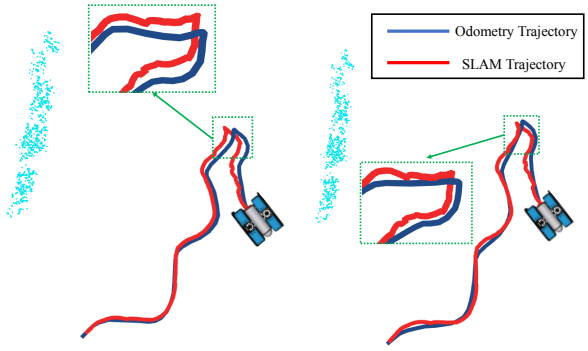
Shida Xu[1], Jonatan Scharff Willners[2], Ziyang Hong[2], Kaicheng Zhang[1], Yvan R. Petillot[2] and Sen Wang[1]

***Abstract*— The extrinsic parameters play a crucial role in multi-sensor fusion, such as visual-inertial Simultaneous Localization and Mapping(SLAM), as they enable the accurate alignment and integration of measurements from different sensors. However, extrinsic calibration is challenging in scenarios, such as underwater, where in-view structures are scanty and visibility is limited, causing incorrect extrinsic calibration due to insufficient motion on all degrees of freedom. In this paper, we propose an entropy-based active extrinsic calibration algorithm leverages observability analysis and information entropy to enhance the accuracy and reliability of extrinsic calibration. It determines the system observability numerically by using singular value decomposition (SVD) of the Fisher Information Matrix (FIM). Furthermore, when the extrinsic parameter is not fully observable, our method actively searches for the next best motion to recover the system's observability via entropy-based optimization. Experimental results on synthetic data, in a simulation, and using an actual underwater vehicle verify that the proposed method is able to avoid the calibration failure while improving the calibration accuracy and reliability.**

I. Introduction

Extrinsic parameters are crucial for multi-sensor fusion and Simultaneous Localisation and Mapping (SLAM). However, calibrating these parameters can be challenging because the non-linearity of the extrinsic transformation, particularly the 3D rotation, can make the optimization problem difficult to converge without a sufficient amount of motion to fully excite all axes. Unfortunately, some scenarios and robot platforms may have confined motion patterns. For instance, the extrinsic calibration involving visual sensors on an underwater vehicle requires the presence of underwater structures in the camera, even suffering from limited visibility. Therefore, it is helpful to identify the motions which degenerate the calibration process and, ideally, also determine the best movement to ensure a good calibration.

Most existing works on extrinsic calibration focus on analyzing the observability of the calibration process. For example, a full observability analysis of a visual-inertial calibration system is studied in [1], showing that at least two-axis rotation and two-axis acceleration of an Inertial Measurement Unit (IMU) is required to calibrate the transformation between a camera and an IMU. However, how the sensor suite moves during the calibration process is very much based on the one who carries out the task. Therefore, it is common for extrinsic calibration to fail to converge to sensible values without some special care. Limited operational requirements may cause insufficient motion. For example, for calibration using a camera, it should always keep facing rich-texture regions to ensure reliable pose estimates. Additionally, even if we can find out that a motion may cause unobservability, we still need a strategy to produce motions that can benefit the calibration.



(a) Transformed trajectory by using [2] (b) Transformed trajectory by using active calibration

Fig. 1. Transformed trajectory comparison between [2] and active aclibration. We transform the SLAM Trajectory of [2] from the camera frame to the odometry frame by using extrinsic parameters. Less misalignment on trajectories means the extrinsic parameter is more accurate. Both results were tested in our wave tank with the BlueRov2 vehicle.

Choosing informative motion segments from all measurements or trying to lock the unobservable direction in the parameter space when systems do not have full observability are promising for extrinsic calibration problems [3], [4]. These methods refuse redundant information or lock some parameters to avoid converging to incorrect calibration parameters during optimization. However, they still mainly rely on data passively received without coupling the motion planning in the extrinsic calibration process.

In this paper, we propose an observability-aware entropy-based active extrinsic calibration that actively searches for the next optimal pose to calibrate the extrinsic parameter. The algorithm first determines the observability of the extrinsic calibration optimization problem by parsing Fisher Information Matrix (FIM). When the calibration parameter is not fully observable, a novel entropy-base optimization method is designed to actively search for the next poses, which a robot can navigate to for reliable calibration. Figure 1 is a comparison between our active calibration and calibration method of [2] with careful manual data selection to cover different motion pattern. We use the calibrated parameter to transform the SLAM trajectory from the camera frame to the odometry frame. It is clear that the active calibration

[1] Department of Electrical and Electronic Engineering, Imperial College London, UK {s.xu23, k.zhang23, sen.wang}@imperial.ac.uk
[2] School of Engineering and Physical Sciences, Heriot-Watt University, UK {J.Scharff_Willners, kz13, Y.R.Petillot}@hw.ac.uk

result has less misalignment, especially near the end of the trajectory, which means the active calibration result is more accurate.

Our main contributions are:

- An entropy-based active method to optimally search for the best next move, which can recover observability of parameters, facilitating convergence and reliability of extrinsic calibration.
- Our proposed method is substantiated by a comprehensive set of experimental results, including evaluations with synthetic data, simulations, and validation using a real-world underwater vehicle. These results demonstrate the effectiveness of our approach in enhancing the convergence and precision of extrinsic calibration.

II. Related Works

Extrinsic calibration is a crucial problem for multi-sensor fusion. Most extrinsic calibration methods are under the state estimation framework and can be mainly divided into two categories: filter-based methods and non-linear optimization-based methods. The filter-based methods normally have fewer requirements for computation and are closely related to filter-based robot localization and SLAM. Therefore, it is widely used on mobile robots; for example, [5] proposes an automatic calibration method for vision and bearing sensors based on Extended Kalman Filter (EKF). [1] proposes an unscented Kalman filter(UKF) based self-calibration method for a visual-inertial system and provides an observability analysis that proves the IMU bias and extrinsic parameters are observable under specific conditions. [6] proposes an EKF-based calibration method for camera and IMU accounting for time correction of IMU measurements. On the other hand, non-linear optimization-based methods require more computation yet tend to achieve better accuracy. [7] calibrates both the intrinsic and extrinsic parameters of a camera and odometry by using Bundle Adjustment (BA). Some works [8], [2] also integrate the calibration problem into graph SLAM systems.

The optimization of the extrinsic calibration can be challenging due to the nonlinearity of the 6 Degree-of-Freedom (DoF) transformation. The choice of calibration data that represents the sensor motion is of importance to obtain a good set of calibration parameters. [9] points out that data sequence may cause degeneration according to observability analysis. [3] uses information theory to detect redundant segments in a dataset and uses an entropy-based method to select informative segments from the dataset for calibration. [4] proposes to detect and lock unobservable directions during updating states by using rank-revealing QR and singular value decomposition of the FIM. However, all these methods focus on using the given data passively to calibrate the extrinsic parameters. In contrast, the method proposed in this paper focuses on how to actively capture data which likely benefits the calibration.

Some of our ideas are also related to the field of Active SLAM. Compared with the classic SLAM approaches, which localize robots and build maps using data received, active SLAM considers the following action, which can be beneficial for reconstruction quality and localization accuracy [10], [11], [12]. For instance, [13] formulates this problem using information entropy, maximizing the information gain (IG), which normally consists of two parts: One is from the map, and the other is from the robot poses. Our work is inspired by information entropy-based active SLAM. However, we not only focus on the fundamental extrinsic calibration problem but also analyze its observability.

III. Proposed Method

A. Problem Formulation

1) State Definition: The extrinsic calibration system aims to estimate the extrinsic parameters, i.e., 6-DoF transformation ${}^{b}\boldsymbol{T}_{c}$, between the sensor frame $\mathcal{F}_b$ and the other sensor frame $\mathcal{F}_c$:

$$ {}^{b}\boldsymbol{T}_{c} = \begin{bmatrix} {}^{b}\boldsymbol{R}_{c} & {}^{b}\boldsymbol{t}_{b\_c} \\ 0 & 1 \end{bmatrix} $$

where ${}^{b}\boldsymbol{R}_{c}$ is the rotation matrix in the Special Orthogonal Group $SO(3)$ and ${}^{b}\boldsymbol{t}_{b\_c} \in \mathbb{R}^3$ is the translation from $\mathcal{F}_b$ to $\mathcal{F}_c$ in $\mathcal{F}_b$. The state vector to estimate is then defined as

$$ \boldsymbol{X} = [\mathrm{Log}({}^{b}\boldsymbol{R}_{c}), {}^{b}\boldsymbol{t}_{b\_c}]^T $$

where $\mathrm{Log}()$ is the logarithmic function to map a quantity from $\mathrm{SO}(3)$ to Lie algebra $\mathfrak{so}(3)$.

2) Sensor Measurements: Assume two trajectories are estimated in the sensor frame $\mathcal{F}_b$ and $\mathcal{F}_c$ using their measurements, respectively. They can be individually represented as a set of relative transformations:

$$ \mathcal{S}_c \doteq \{{}^{c_i}\boldsymbol{T}_{c_j}\}_{i,j\in\mathcal{K}} \quad \text{and} \quad \mathcal{S}_b \doteq \{{}^{b_i}\boldsymbol{T}_{b_j}\}_{i,j\in\mathcal{K}} $$

where $\mathcal{K}$ is the set of timestamps, and ${}^{b_i}\boldsymbol{T}_{b_j}$ and ${}^{c_i}\boldsymbol{T}_{c_j}$ are transformation matrices representing the relative motion from time i to j, respectively. Therefore, the following transformation can be established

$$ {}^{b_i}\boldsymbol{T}_{b_j} = {}^{b}\boldsymbol{T}_{c}\ {}^{c_i}\boldsymbol{T}_{c_j}\ {}^{b}\boldsymbol{T}_{c}^{-1} \tag{1} $$

3) Non-Linear Optimization of Extrinsic Calibration: Given multiple measurements with Gaussian noises, we can use non-linear optimization to solve the calibration problem. According to (1), the transformation can be split into rotational and translational parts, defining rotational and translational residual terms respectively as

$$ \begin{aligned} &\Phi_{\boldsymbol{R}}({}^{b}\boldsymbol{R}_{c}) = \mathrm{Log}({}^{b_i}\boldsymbol{R}_{b_j}^{-1}\ {}^{b}\boldsymbol{R}_{c}\ {}^{c_i}\boldsymbol{R}_{c_j}\ {}^{b}\boldsymbol{R}_{c}^{-1}) \\ &f_{\boldsymbol{t}}({}^{b}\boldsymbol{R}_{c}, {}^{b}\boldsymbol{t}_{b\_c}) = -{}^{b_i}\boldsymbol{R}_{b_j}^{-1}\ {}^{b_i}\boldsymbol{t}_{b_i b_j} + {}^{b_i}\boldsymbol{R}_{b_j}^{-1}\ {}^{b}\boldsymbol{t}_{b\_c} \\ &+{}^{b_i}\boldsymbol{R}_{b_j}^{-1}\ {}^{b}\boldsymbol{R}_{c}\ {}^{c0}\boldsymbol{t}_{c_i c_j} - {}^{b_i}\boldsymbol{R}_{b_j}^{-1}\ {}^{b}\boldsymbol{R}_{c}\ {}^{c_i}\boldsymbol{R}_{c_j}\ {}^{b}\boldsymbol{R}_{c}^{-1}\ {}^{b}\boldsymbol{t}_{b\_c} \end{aligned} \tag{2} $$

Since the rotation residual term $\Phi_{\boldsymbol{R}}({}^{b}\boldsymbol{R}_{c})$ solely depends on the rotation, for sensors that only provide rotational measurements, the translation parameters cannot be calibrated. In contrast, the translation residual, which links both the rotation and translation, can help calibrate the full 6-DoF parameters.

Now, we can formulate the extrinsic calibration as non-linear optimisation problems:

$$\underset{{}^{b}\boldsymbol{R}_{c}}{\operatorname{argmin}} \sum_{\mathcal{K}} ||\Phi_{\boldsymbol{R}}({}^{b}\boldsymbol{R}_{c})||^{2}_{\Sigma} \quad \text{and} \quad \underset{{}^{b}\boldsymbol{R}_{c},{}^{b}\boldsymbol{t}_{b\_c}}{\operatorname{argmin}} \sum_{\mathcal{K}} ||f_{\boldsymbol{t}}({}^{b}\boldsymbol{R}_{c},\ {}^{b}\boldsymbol{t}_{b\_c})||^{2}_{\Sigma}$$

For sensors providing 6-DoF estimates, e.g., IMUs or cameras, the two residual terms can be combined. However, for a sensor that only measures rotation or translation, e.g., a gyroscope or a Doppler Velocity Log (DVL), only one residual term can be established. To cover different sensor configurations, the rotational and translational residuals are defined separately in this paper.

The Gaussian-Newton method can now be used to solve the optimisation problems iteratively. Specifically, the first-order Taylor approximation of (2) is

$$\Phi_{\boldsymbol{R}}({}^{b}\boldsymbol{R}_{c}) \approx \Phi_{\boldsymbol{R}}({}^{b}\boldsymbol{R}_{c}) + \boldsymbol{J}_{\boldsymbol{R}}\boldsymbol{\Delta\phi}$$
$$f_{\boldsymbol{t}}({}^{b}\boldsymbol{R}_{c},\ {}^{b}\boldsymbol{t}_{b\_c}) \approx f_{\boldsymbol{t}}({}^{b}\boldsymbol{R}_{c},\ {}^{b}\boldsymbol{t}_{b\_c}) + \boldsymbol{J}_{\boldsymbol{t}}\boldsymbol{\Delta X}$$

where $\boldsymbol{\Delta\phi}$ is the rotation increment in $\mathfrak{so}(3)$, and $\boldsymbol{J}_{\boldsymbol{R}}$ is the Jacobin matrix of $\Phi_{\boldsymbol{R}}()$:

$$\boldsymbol{J}_{\boldsymbol{R}} = \frac{\partial \Phi_{\boldsymbol{R}}({}^{b}\boldsymbol{R}_{c})}{\partial \boldsymbol{\Delta\phi}} = {}^{b_i}\boldsymbol{R}_{b_j}^{-1}\ {}^{b}\boldsymbol{R}_{c}({}^{c_i}\boldsymbol{R}_{c_j} - \boldsymbol{I}) \tag{3}$$

Also $\boldsymbol{\Delta X}$ is the increment of the full state $\boldsymbol{X}$, and $\boldsymbol{J}_{\boldsymbol{t}}$ is the Jacobin matrix of the translational residual w.r.t. the full state $\boldsymbol{X}$ as

$$\boldsymbol{J}_{\boldsymbol{t}} = \begin{bmatrix} \frac{\partial f_{\boldsymbol{t}}({}^{b}\boldsymbol{R}_{c},\ {}^{b}\boldsymbol{t}_{b\_c})}{\partial \boldsymbol{\Delta\phi}} & \frac{\partial f_{\boldsymbol{t}}({}^{b}\boldsymbol{R}_{c},\ {}^{b}\boldsymbol{t}_{b\_c})}{\partial {}^{b}\boldsymbol{t}_{b\_c}} \end{bmatrix} = \begin{bmatrix} \frac{\mathrm{d}\Omega_{\boldsymbol{t}}({}^{b}\boldsymbol{R}_{c})}{\mathrm{d}\boldsymbol{\Delta\phi}} & \frac{\mathrm{d}\Psi_{\boldsymbol{t}}({}^{b}\boldsymbol{t}_{b\_c})}{\mathrm{d}{}^{b}\boldsymbol{t}_{b\_c}} \end{bmatrix} \tag{4}$$

where $\frac{\partial f_{\boldsymbol{t}}({}^{b}\boldsymbol{R}_{c},\ {}^{b}\boldsymbol{t}_{b\_c})}{\partial \boldsymbol{\Delta\phi}}$ is the Jacobin of translation residual w.r.t. to the rotation it can be written as:

$$\frac{\mathrm{d}\Omega_{\boldsymbol{t}}({}^{b}\boldsymbol{R}_{c})}{\mathrm{d}\boldsymbol{\Delta\phi}} = -{}^{b_i}\boldsymbol{R}_{b_j}^{-1}\ {}^{b}\boldsymbol{R}_{c}\ {}^{c_i}\boldsymbol{t}_{ci\_cj}^{\wedge} - {}^{b_i}\boldsymbol{R}_{b_j}^{-1}\ {}^{b}\boldsymbol{R}_{c}\ [{}^{c1}\boldsymbol{R}_{c2}\ {}^{b}\boldsymbol{R}_{c}^{-1}\ {}^{b}\boldsymbol{t}_{b\_c}]^{\wedge}$$
$$- {}^{b_i}\boldsymbol{R}_{b_j}^{-1}\ {}^{b}\boldsymbol{R}_{c}\ {}^{c1}\boldsymbol{R}_{c2}\ [{}^{b}\boldsymbol{R}_{c}^{-1}\ {}^{b}\boldsymbol{t}_{b\_c}]^{\wedge}$$

where $(\cdot)^{\wedge}$ denotes the skew-symmetric matrix operator. $\frac{\mathrm{d}\Psi_{\boldsymbol{t}}({}^{b}\boldsymbol{t}_{b\_c})}{\mathrm{d}{}^{b}\boldsymbol{t}_{b\_c}}$ is the Jacobin of translation residual w.r.t. translation and it can be calculated as:

$$\frac{\mathrm{d}\Psi_{\boldsymbol{t}}({}^{b}\boldsymbol{t}_{b\_c})}{\mathrm{d}\boldsymbol{\Delta t}} = {}^{b_i}\boldsymbol{R}_{b_j}^{-1} - {}^{b_i}\boldsymbol{R}_{b_j}^{-1}\ {}^{b}\boldsymbol{R}_{c}\ {}^{c_i}\boldsymbol{R}_{c_j}\ {}^{b}\boldsymbol{R}_{c}^{-1}$$

Because of the page limit, we directly give the closed-form of these Jacobin matrices.[14]

For the rotation residual, its *Fisher Information Matrix (FIM)* is

$$I_{\boldsymbol{R}}({}^{b}\boldsymbol{R}_{c}) = \boldsymbol{J}_{\boldsymbol{R}}^{T}\Sigma^{-1}\boldsymbol{J}_{\boldsymbol{R}} \tag{5}$$

where Σ^{-1} is the inverse of the covariance matrix. Similarly, the FIM of the translation residual is

$$I_{\boldsymbol{t}}({}^{b}\boldsymbol{R}_{c},\ {}^{b}\boldsymbol{t}_{b\_c}) = \boldsymbol{J}_{\boldsymbol{t}}^{T}\Sigma^{-1}\boldsymbol{J}_{\boldsymbol{t}} \tag{6}$$

A necessary condition for the Gaussian-Newton method to solve the problem properly is the FIM must be a positive-definite matrix. However, it can be positive-semi-definite due to insufficient motion experienced. A positive-semi-definite or ill-conditioned FIM often causes the optimization converges to incorrect values, failing to solve the problem. Therefore, we propose to solve this by actively planning the next poses, shaping a positive-semi-definite or ill-conditioned FIM into a positive-definite or well-conditioned FIM.

B. Observability Determination

The observability is determined by the rank of the FIM defined by (5) and (6). When the FIM is full rank, the system is observable.[15] Otherwise, it is unobservable, which means the parameters cannot be well constrained, causing calibration errors. Since it is known Σ^{-1} is a positive-definite matrix, according to [16] we have

$$\operatorname{rank}(\boldsymbol{J}_{\boldsymbol{R}}^{T}\Sigma^{-1}\boldsymbol{J}_{\boldsymbol{R}}) = \operatorname{rank}(\boldsymbol{J}_{\boldsymbol{R}}^{T}\boldsymbol{J}_{\boldsymbol{R}}) = \operatorname{rank}(\boldsymbol{J}_{\boldsymbol{R}})$$
$$\operatorname{rank}(\boldsymbol{J}_{\boldsymbol{t}}^{T}\Sigma^{-1}\boldsymbol{J}_{\boldsymbol{t}}) = \operatorname{rank}(\boldsymbol{J}_{\boldsymbol{t}}^{T}\boldsymbol{J}_{\boldsymbol{t}}) = \operatorname{rank}(\boldsymbol{J}_{\boldsymbol{t}})$$

So the observability depends on the ranks of $\boldsymbol{J}_{\boldsymbol{R}}$ and $\boldsymbol{J}_{\boldsymbol{t}}$ in (3) and (4) which can be numerically calculated by Singular-Value-Decomposition (SVD). However, we decompose $\boldsymbol{J}_{\boldsymbol{R}}^{T}\boldsymbol{J}_{\boldsymbol{R}}$ and $\boldsymbol{J}_{\boldsymbol{t}}^{T}\boldsymbol{J}_{\boldsymbol{t}}$ instead. This is because the row numbers of $\boldsymbol{J}_{\boldsymbol{R}}$ and $\boldsymbol{J}_{\boldsymbol{t}}$ increase along with more measurements, but $\boldsymbol{J}_{\boldsymbol{R}}^{T}\boldsymbol{J}_{\boldsymbol{R}}$ and $\boldsymbol{J}_{\boldsymbol{t}}^{T}\boldsymbol{J}_{\boldsymbol{t}}$ have fixed sizes which are 3×3 and 6×6, respectively. Now we have:

$$\operatorname{SVD}(\boldsymbol{J}_{\boldsymbol{R}}^{T}\boldsymbol{J}_{\boldsymbol{R}}) = \boldsymbol{U}_{\boldsymbol{R}}\ \operatorname{Diag}([\sigma_{\boldsymbol{R}1},\sigma_{\boldsymbol{R}2},\sigma_{\boldsymbol{R}3}]^{T})\ \boldsymbol{V}_{\boldsymbol{R}}$$

where Diag() converts a vector to a diagonal matrix, and $\sigma_{\boldsymbol{R}1} \geq \sigma_{\boldsymbol{R}2} \geq \sigma_{\boldsymbol{R}3}$. Theoretically, when a singular value exists, its corresponding direction becomes unobservable.

For the translation residual, we have a similar SVD as

$$\operatorname{SVD}(\boldsymbol{J}_{\boldsymbol{t}}^{T}\boldsymbol{J}_{\boldsymbol{t}}) = \boldsymbol{U}_{\boldsymbol{t}}\ \operatorname{Diag}([\sigma_{\boldsymbol{t}1},\sigma_{\boldsymbol{t}2},\sigma_{\boldsymbol{t}3},\sigma_{\boldsymbol{t}4},\sigma_{\boldsymbol{t}5},\sigma_{\boldsymbol{t}6}]^{T}\ \boldsymbol{V}_{\boldsymbol{t}}$$

where $\sigma_{\boldsymbol{t}1} \geq \sigma_{\boldsymbol{t}2} \geq \sigma_{\boldsymbol{t}3} \geq \sigma_{\boldsymbol{t}4} \geq \sigma_{\boldsymbol{t}5} \geq \sigma_{\boldsymbol{t}6}$.

C. Entropy-Based Active Optimization (EBAO)

The observability determination algorithm can indicate whether the parameters are observable. When they are unobservable, we propose a strategy to recover the observability by actively planning the next motion that maximizes the entropy reduction, aka the *Information Gain (IG)*.

1) Optimal Next Pose Maximising IG: The *Information Entropy (IE)* is defined as $H(\boldsymbol{X}) = -\int_{X} p(x)\log p(x)\,dx$. If $\boldsymbol{X} \sim \mathcal{N}_{D}(\mu,\Sigma)$ as a multi-variate Gaussian, accoridng to [17] its IE can be formulated as

$$H(\boldsymbol{X}) = \frac{1}{2}\log((2\pi e)^{D}|\Sigma|)$$

Since we assume the extrinsic parameters are random variables following Gaussian distribution. Then we can recover the covariance matrices from the FIM $I_{\boldsymbol{R}}({}^{b}\boldsymbol{R}_{c})$ and $I_{\boldsymbol{t}}({}^{b}\boldsymbol{R}_{c},\ {}^{b}\boldsymbol{t}_{b\_c})$.

a) Rotation Residual: Considering (5), the entropy of the rotation parameter is

$$H({}^{b}\boldsymbol{R}_{c}|\mathcal{S}_{c},\mathcal{S}_{b}) = -\frac{1}{2}\log((2\pi e)^{D}) - \frac{1}{2}\log(|\boldsymbol{J}_{\boldsymbol{R}}^{T}\Sigma^{-1}\boldsymbol{J}_{\boldsymbol{R}}|)$$

Once new measurements ${}^{ci}\boldsymbol{T}_{c_j}^{n}$ and ${}^{bi}\boldsymbol{T}_{b_j}^{n}$ for the next timestamp is added to $\mathcal{S}_{c}$ and $\mathcal{S}_{b}$, new rows will be appended in $\boldsymbol{J}_{\boldsymbol{R}}$ to form $\boldsymbol{J}_{\boldsymbol{R}}^{\mathrm{n}}$. Therefore, the entropy is updated as

$$H({}^{b}\boldsymbol{R}_{c}|\mathcal{S}_{c},\mathcal{S}_{b},\ {}^{c0}\boldsymbol{T}_{c_i}^{n},\ {}^{b0}\boldsymbol{T}_{b_i}^{n}) = -\frac{1}{2}\log((2\pi e)^{D}) - \frac{1}{2}\log(|{\boldsymbol{J}_{\boldsymbol{R}}^{\mathrm{n}}}^{T}\Sigma^{-1}\boldsymbol{J}_{\boldsymbol{R}}^{\mathrm{n}}|)$$

Based on these, we formulate the problem to compute next optimal poses ${}^{c0}\boldsymbol{T}_{c_i}^{n}$ and ${}^{b0}\boldsymbol{T}_{b_i}^{n}$ which maximise the IG through

$$\underset{{}^{b0}\boldsymbol{T}_{b_i}^{n}}{\operatorname{argmax}}\ H({}^{b}\boldsymbol{R}_{c}|\mathcal{S}_{c},\mathcal{S}_{b}) - H({}^{b}\boldsymbol{R}_{c}|\mathcal{S}_{c},\mathcal{S}_{b},{}^{c0}\boldsymbol{T}_{c_i}^{n},{}^{b0}\boldsymbol{T}_{b_i}^{n}) \tag{7}$$

Note that only ${}^{b0}\boldsymbol{T}_{b_i}^n$ is derived since ${}^{c0}\boldsymbol{T}_{c_i}^n$ can be calculated from ${}^{b0}\boldsymbol{T}_{b_i}^n$ using the initial extrinsic parameters. (7) can be further simplified as

$$H({}^b\boldsymbol{R}_c|\mathcal{S}_c,\mathcal{S}_b) - H({}^b\boldsymbol{R}_c|\mathcal{S}_c,\mathcal{S}_b,{}^{c0}\boldsymbol{T}_{c_i}^n,{}^{b0}\boldsymbol{T}_{b_i}^n)$$
$$= \frac{1}{2}\log\frac{|\boldsymbol{J}_{\boldsymbol{R}}^{\mathrm{n}\,T}\Sigma^{-1}\boldsymbol{J}_{\boldsymbol{R}}^{\mathrm{n}}|}{|\boldsymbol{J}_{\boldsymbol{R}}^T\Sigma^{-1}\boldsymbol{J}_{\boldsymbol{R}}|} = \frac{1}{2}\log\frac{|I_{\boldsymbol{R}}^n({}^b\boldsymbol{R}_c)|}{|I_{\boldsymbol{R}}({}^b\boldsymbol{R}_c)|}$$

Since $I_{\boldsymbol{R}}({}^b\boldsymbol{R}_c)$ is the initial FIM which is determined by the initial value, it is considered a constant here. Thus, (7) equals

$$\underset{{}^{b0}\boldsymbol{T}_{b_i}^n}{\operatorname{argmin}} \log\frac{1}{|I_{\boldsymbol{R}}^n({}^b\boldsymbol{R}_c)|} \tag{8}$$

where $|I_{\boldsymbol{R}}^n({}^b\boldsymbol{R}_c)| = \sigma_{\boldsymbol{R}1}\sigma_{\boldsymbol{R}2}\sigma_{\boldsymbol{R}3}$ as the determinant of the FIM. To avoid a negative residual and unstable numerical issues, the residual of (8) is refined as following with an auxiliary $\alpha = 2$ considering $|I_{\boldsymbol{R}}^n({}^b\boldsymbol{R}_c)| \geq 0$:

$$f_{\text{IE\_}\boldsymbol{R}}({}^{b0}\boldsymbol{T}_{b_i}^n) = \frac{1}{\log(\alpha + |I_{\boldsymbol{R}}^n({}^b\boldsymbol{R}_c)|)}$$

b) Translation Residual: Similarly, for the translation, we can derive the below residual using (6)

$$f_{\text{IE\_}\boldsymbol{t}}({}^{b0}\boldsymbol{T}_{b_i}^n) = \frac{1}{\log(\alpha + |I_{\boldsymbol{t}}^n({}^b\boldsymbol{R}_c,\ {}^b\boldsymbol{t}_{b\_c})|)}$$

2) Normalized-Singular-Value Entropy: The IE residual maximizes the IG and reduces the IE with new pose measurements added, especially for the cases where the FIM is nearly rank-deficient. However, the FIM might be ill-conditioned, even though it is full rank. Hence, we proposed *Normalized-Singular-Value Entropy (NSVE)* to measure the singularity of the FIM.

For the rotation residual, its NSVE is defined as

$$H_{\text{NSV}}(I_{\boldsymbol{R}}({}^b\boldsymbol{R}_c)) = -\sum_i \frac{\sigma_{\boldsymbol{R}i}}{\text{norm}(\boldsymbol{\sigma}_{\boldsymbol{R}})}\ log(\frac{\sigma_{\boldsymbol{R}i}}{\text{norm}(\boldsymbol{\sigma}_{\boldsymbol{R}})})$$

where $\boldsymbol{\sigma}_{\boldsymbol{R}i} = [\sigma_{\boldsymbol{R}1}, \sigma_{\boldsymbol{R}2}, \sigma_{\boldsymbol{R}3}]^T$. $\frac{\sigma_{\boldsymbol{R}i}}{\text{norm}(\boldsymbol{\sigma}_{\boldsymbol{R}})}$ is each normalized singular value which range from 0 to 1. Similar to the concept of entropy of discrete random variables, NSVE describes the chaos of the singular value distribution. In order to avoid ill-conditioned FIM, NSVE should be as high as possible.

We can reformulate this problem as non-linear optimization, and the residual is defined as

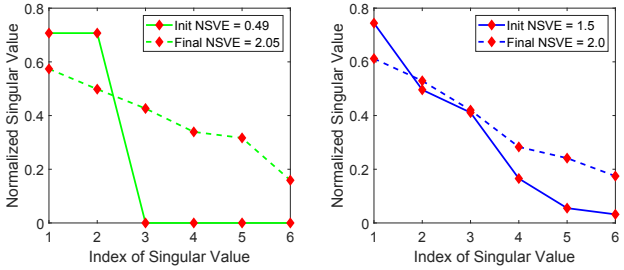
$$f_{\text{NSVE\_}\boldsymbol{R}}({}^{b0}\boldsymbol{T}_{b_i}^n) = \frac{1}{H_{\text{NSV}}(I_{\boldsymbol{R}}^n({}^b\boldsymbol{R}_c))}$$

Similarly, the NSVE residual for the translation is

$$f_{\text{NSVE\_}\boldsymbol{t}}({}^{b0}\boldsymbol{T}_{b_i}^n) = \frac{1}{H_{\text{NSV}}(I_{\boldsymbol{t}}^n({}^b\boldsymbol{R}_c,\ {}^b\boldsymbol{t}_{b\_c}))}$$

3) Non-Linear Optimisation: The IG residuals and NSVE residuals are formulated as a joint non-linear optimization to find the motion which optimally increases the IG while avoiding rank-deficient or ill-condition of the FIM:

$$\underset{{}^{b0}\boldsymbol{T}_{b_i}^n}{\operatorname{argmin}} ||f_{\text{IE\_}\boldsymbol{R}}({}^{b0}\boldsymbol{T}_{b_i}^n)||^2 + ||f_{\text{IE\_}\boldsymbol{t}}({}^{b0}\boldsymbol{T}_{b_i}^n)||^2$$
$$+ ||f_{\text{NSVE\_}\boldsymbol{R}}({}^{b0}\boldsymbol{T}_{b_i}^n)||^2 + ||f_{\text{NSVE\_}\boldsymbol{t}}({}^{b0}\boldsymbol{T}_{b_i}^n)||^2$$



(a) Pure t_x Sequence (b) $\boldsymbol{R}\&\boldsymbol{t}$ Sequence

Fig. 2. Normalized Singular Values and NSVE on two sequences

IV. Experimental Results

We validate the proposed method for multi-sensor extrinsic calibration on synthetic data, in a simulation, and on a real underwater vehicle. Since it is often impossible to obtain ground truth extrinsic parameters in the real world, we use synthetic data and simulation for quantitative evaluation.

A. Experiments on Synthetic Data

1) IE, NSVE and Calibration Errors: Using the proposed method, we first evaluate how the IE, NSVE, and calibration errors evolve. We generate three types of sequences to simulate various settings: 1) Pure rotation whose initial poses only have rotation. 2) Pure translation whose initial poses only have the translation. 3) Rotation-and-translation whose initial poses have full 6-DoF transformation. Each sequence has two initial poses plus ten additional pose measurements searched using the proposed active calibration method. All the poses are corrupt with random Gaussian noises. Table I shows the results before and after the proposed active calibration. It is clear that the sequences with high IE and low NSVE, which indicate the rank-deficiency or ill-conditioned FIMs hindering the optimization converge, often suffer from significant calibration errors. After the active calibration, a considerable reduction in IE and an increase in NSVE are achieved, while the calibration errors considerably decrease.

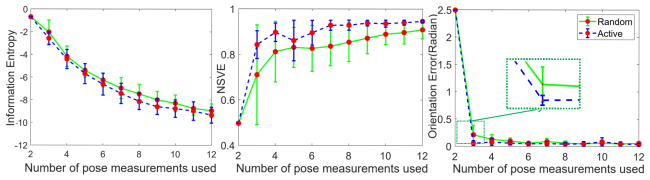
The Normalized Singular Values and NSVE of pure x-translation sequence (t_x) is shown in Figure 2(a). We can see the rank of its FIM is initially only 2 (four singular values are zero) due to its insufficient motion. Its singular values are also distributed apart, producing low NSVE (0.49). After the active calibration, the FIM becomes full-rank, with its singular values being more evenly distributed (NSVE increased to 2.05). Therefore, the calibration errors decrease significantly as shown in Table I. On the other hand, Figure 2(b) shows the initial FIM of the rotation-and-translation sequence ($\boldsymbol{R}\&\boldsymbol{t}$) has full-rank six, and its singular values are already evenly distributed thanks to its sufficient 6-DoF motion. In this case, its NSVE can still have obvious increase which results in a more flat eigenvalue distribution, and the calibration errors reduce slightly.

2) Comparison with Random Motion: The results of ten repeated experiments comparing our method with random motion are given in Figure 3. The two methods share the same initial sets of pose measurements, but the random

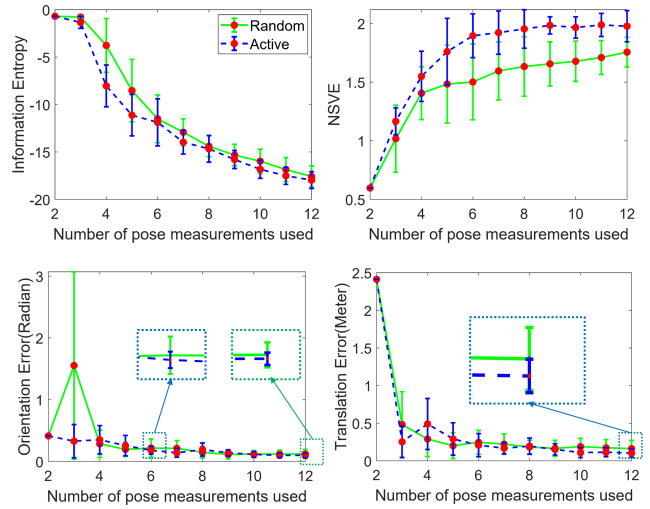
TABLE I: IE, NSVE and Calibration Errors Before and After the Proposed EBAO

	Seq.	Rotation Residual						Translation Residual							
		IE↓		NSVE↑		Rot Error (Radian)↓		IE↓		NSVE↑		Rot Error (Radian)↓		Trans Error (Meter)↓	
		Before	After	Before	After	Before	After	Before	After	Before	After	Before	After	Before	After
Pure $\boldsymbol{R}$	R_x	-0.74	**-7.99**	0.51	**0.93**	1.55	**0.04**	-0.69	**-16.84**	0.61	**1.64**	1.34	**0.0001**	1.4	**0.0001**
	R_xy	-0.72	**-6.14**	0.65	**0.95**	0.29	**0.0006**	-0.69	**-9.12**	1.33	**1.64**	0.83	**0.0002**	0.23	**0.0002**
	R_xyz	-2.73	**-6.65**	0.8	**0.92**	0.17	**0.02**	-0.72	**-16.06**	1.41	**1.99**	0.67	**0.0001**	0.46	**0.0001**
Pure $\boldsymbol{t}$	t_x	-0.69	**-9.93**	N/A	**0.94**	0.78	**0.0003**	-0.69	**-18.91**	0.49	**2.05**	1.98	**0.0001**	3	**9.00E-05**
	t_xy	-0.69	**-8.94**	N/A	**0.95**	0.78	**0.02**	-0.69	**-13.24**	0.89	**2.03**	0.002	0.002	1.37	**0.01**
	t_xyz	-0.69	**-10.15**	N/A	**0.90**	0.78	**0.001**	-0.69	**-18.04**	0.81	**1.98**	0.001	**0.0001**	1.37	**0.0001**
$\boldsymbol{R\&t}$		-0.70	**-9.43**	0.80	**0.95**	0.06	**0.002**	-0.69	**-13.02**	1.5	**2.00**	0.0001	**6.38e-5**	0.0001	**0.0001**

IE: the smaller, the better. NSVE: the bigger, the better. Rotation and Translation Errors: the smaller, the better. In the experiment sention, we used $-\text{Log}(2+I_{\boldsymbol{R}}({}^{b}\boldsymbol{R}_{c}))$ and $-\text{Log}(2+I^{n}_{\boldsymbol{t}}({}^{b}\boldsymbol{R}_{c},\ {}^{b}\boldsymbol{t}_{b\_c}))$ instead of the origial definition of information entropy to avoid infinite numerical issue.



(a) Calibration using Rotation Residual



(b) Calibration using Rotation and Translation Residuals

Fig. 3. Ten repeated experiments comparing random motion and the proposed active method. Error bars denote max and min values of ten runs.

method samples the next ten poses randomly from a Gaussian distribution (variance on translation and rotation are 1 meter and $\frac{\pi}{2}$ radian, respectively). While our method calculates the next ten poses using the proposed EBAO. We can see from the figure that both methods can reduce the IE, while the NSVE of our method is higher eventually. Meanwhile, the proposed active calibration algorithm achieves lower and more stable rotation, and translation calibration errors, especially its max-min errors of ten runs are smaller.

B. Experiments in Simulation

Because of the challenges of extrinsic calibration in underwater scenarios, as motivated in the Introduction, we evaluate the proposed method for underwater vehicles in simulation.

1) Simulation Environment and Setup: The Unmanned Underwater Vehicle (UUV) Simulator [18] is used to perform onboard sensor calibration of the RexROV2 underwater vehicle. Thanks to the autopilot already integrated in the RexROV2 package, a 6-DoF pose can be used as a waypoint for navigating the vehicle. We use pose measurements estimated in its camera frame and IMU frame for extrinsic calibration.

2) Comparison with Manual Control: The calibration system is initialized with two pose measurements, similar to the previous experiments using synthetic data. Then the proposed active calibration algorithm is employed to compute the next pose as a waypoint for the autopilot to control the vehicle. Once the vehicle arrives at the waypoint, its pose is used immediately to perform the extrinsic calibration as a new measurement before using the active calibration algorithm to derive another new pose. This process then repeats until five new pose measurements are added in. For comparison, we also generate five waypoints via random manual control using a joystick. The calibration results of the two methods are shown in Figure 4. It can be seen that the IE of the manual control does not change much due to the limited roll and pitch motion that the manual control cannot achieve, even though we tried to push as much different motion patterns as possible. It is obvious the calibration errors of the proposed active method are much lower than the manual control.

C. Experiments on a Real Underwater Vehicle

Based on the simulation results, we further validate the proposed active method on a real underwater vehicle in a tank environment, as shown in Figure 5. A BlueROV2 underwater vehicle which is equipped with a custom sensor payload consisting of a stereo camera, DVL, and IMU, is used for this experiment. It has an autopilot framework to follow given waypoints [19]. To execute the extrinsic calibration between the camera frame and the inertial frame, we need to ensure that the vision system is able to estimate accurate 6-DoF poses. An underwater visual, acoustic SLAM method [20], [2] is used for this experiment. The poses in the inertial frame are estimated by fusing the DVL, a Gyroscope, and a depth sensor. The extrinsic calibration is initialized as an identity matrix. Similar to the simulation experiment, the proposed active calibration method searches for the optimal next pose, which is taken as a waypoint by the autopilot for navigation. Since at least part of underwater structures needs to be in the camera's view to keep visual tracking,

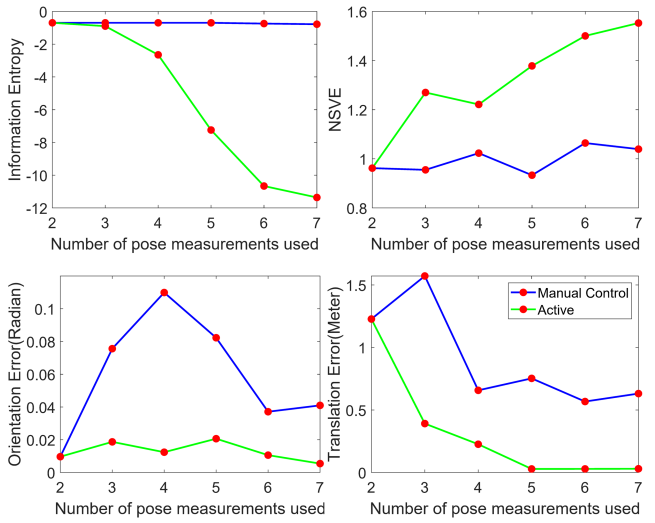
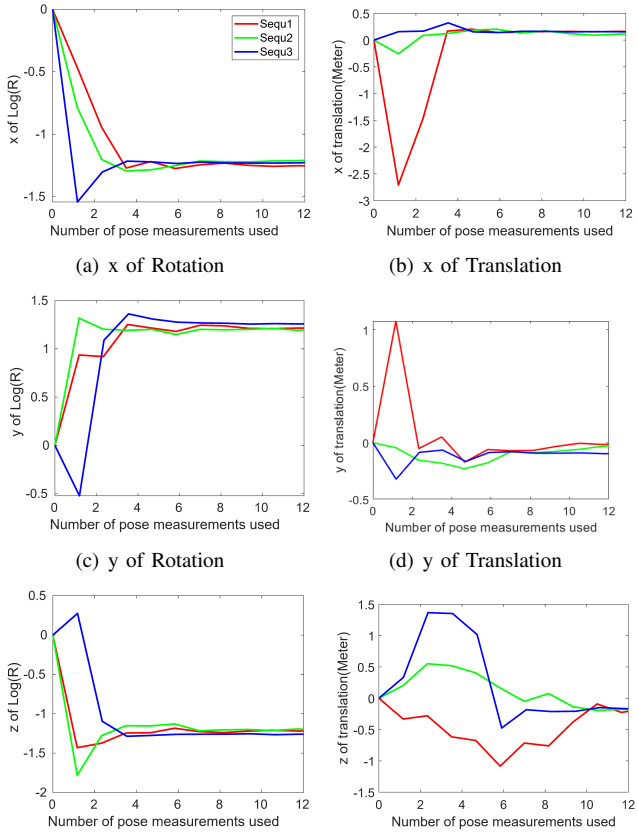


Fig. 4. Comparison between manual control and our active method on calibrating camera-IMU sensor suite in UUV simulator.

(a) BlueROV with sensor frames (b) Tank with underwater structures

Fig. 5. BlueROV robot platform and tank used for real experiments.



(a) x of Rotation (b) x of Translation

(c) y of Rotation (d) y of Translation

(e) z of Rotation (f) z of Translation

Fig. 6. Extrinsic calibration results on a real underwater vehicle. Our experimental results demonstrate that the calibration outcomes from three distinct sequences all converged to the manually measured result.

the waypoint's orientation is refined by ray casting with the structures with a sampling method. The sampling method would sample uniformly around the given waypoint and filter out the orientation that does not have a map in the view of the camera, then chose the orientation sample with the lowest error cost as the next way point.

The three runs of the calibration results are shown in Figure 6. We can see the calibration converges fast, mainly in the first 4-5 measurements, and all three runs converge to similar extrinsic parameters, which are close to the hand-measured values, in 12 measurements. This experiment verifies that the proposed active method is effective for practicing extrinsic calibration.

V. Conclusions

In this paper, we propose a novel observability-aware entropy-based active calibration method. The SVD-based observability detection method first detects the observability of the parameter with a given data sequence by parsing the Fisher Information matrix. Once the parameter is detected as unobservable, the entropy-based active optimization will search for the next waypoint by maximizing the information gain and normalized singular value entropy while recovering the parameter's observability. The synthetic, simulation, and real world tank experiments demonstrated the effective of our algorithm.

Currently, the initial value affects the result a lot. In addition, the information gain and normalized singular value entropy maximization residual is a non-convex function that also is affected by the initial value and might converge to a local minimum. We have further interest in how to find the global minimum for any arbitrary given initial value. In addition, the current algorithm mainly considers searching the motion benefits the calibration in the whole $SE(3)$ space. We are also interested in how to find the solution in a subspace given specific motion constrain. And another interesting further work is to integrate the algorithm into a SLAM framework so we can consider reducing the uncertainty of map, pose estimation, and extrinsic parameters simultaneously.

Acknowledgment

This work is partially supported by EU H2020 DeepField (857339) and METRICS (871252) projects. The authors would like to thank Joshua Roe and Sean Katagiri for their technical support on data collection and trials.

REFERENCES

[1] J. Kelly and G. S. Sukhatme, "Visual-inertial sensor fusion: Localization, mapping and sensor-to-sensor self-calibration," *The International Journal of Robotics Research*, vol. 30, no. 1, pp. 56–79, 2011.

[2] S. Xu, T. Luczynski, J. S. Willners, Z. Hong, K. Zhang, Y. R. Petillot, and S. Wang, "Underwater visual acoustic slam with extrinsic calibration," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2021, pp. 7647–7652.

[3] E. Dexheimer, P. Peluse, J. Chen, J. Pritts, and M. Kaess, "Information-theoretic online multi-camera extrinsic calibration," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 4757–4764, 2022.

[4] J. Maye, H. Sommer, G. Agamennoni, R. Siegwart, and P. Furgale, "Online self-calibration for robotic systems," *The International Journal of Robotics Research*, vol. 35, no. 4, pp. 357–380, 2016.

[5] A. Martinelli, D. Scaramuzza, and R. Siegwart, "Automatic self-calibration of a vision system during robot motion," in *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006.* IEEE, 2006, pp. 43–48.

[6] F. M. Mirzaei and S. I. Roumeliotis, "A kalman filter-based algorithm for imu-camera calibration: Observability analysis and performance evaluation," *IEEE transactions on robotics*, vol. 24, no. 5, pp. 1143–1156, 2008.

[7] L. Heng, B. Li, and M. Pollefeys, "Camodocal: Automatic intrinsic and extrinsic calibration of a rig with multiple generic cameras and odometry," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2013, pp. 1793–1800.

[8] R. Kümmerle, G. Grisetti, and W. Burgard, "Simultaneous calibration, localization, and mapping," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2011, pp. 3716–3721.

[9] J. Brookshire and S. Teller, "Extrinsic calibration from per-sensor egomotion," *Robotics: Science and Systems VIII*, pp. 504–512, 2013.

[10] B. Mu, M. Giamou, L. Paull, A.-a. Agha-mohammadi, J. Leonard, and J. How, "Information-based active slam via topological feature graphs," in *2016 IEEE 55th Conference on decision and control (Cdc)*. IEEE, 2016, pp. 5583–5590.

[11] I. Maurović, M. Seder, K. Lenac, and I. Petrović, "Path planning for active slam based on the d* algorithm with negative edge weights," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 48, no. 8, pp. 1321–1331, 2017.

[12] R. Valencia and J. Andrade-Cetto, "Active pose slam," in *Mapping, Planning and Exploration with Pose SLAM*. Springer, 2018, pp. 89–108.

[13] F. Bourgault, A. A. Makarenko, S. B. Williams, B. Grocholsky, and H. F. Durrant-Whyte, "Information based adaptive robotic exploration," in *IEEE/RSJ international conference on intelligent robots and systems*, vol. 1. IEEE, 2002, pp. 540–545.

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

[15] Z. Wang and G. Dissanayake, "Observability analysis of slam using fisher information matrix," in *2008 10th International Conference on Control, Automation, Robotics and Vision*. IEEE, 2008, pp. 1242–1247.

[16] C. Jauffret, "Observability and fisher information matrix in nonlinear regression," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 2, pp. 756–759, 2007.

[17] T. A. Vidal-Calleja, A. Sanfeliu, and J. Andrade-Cetto, "Action selection for single-camera slam," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 40, no. 6, pp. 1567–1581, 2010.

[18] M. M. M. Manhães, S. A. Scherer, M. Voss, L. R. Douat, and T. Rauschenbach, "Uuv simulator: A gazebo-based package for underwater intervention and multi-robot simulation," in *OCEANS 2016 MTS/IEEE Monterey*. IEEE, 2016, pp. 1–8.

[19] J. S. Willners, I. Carlucho, S. Katagiri, C. Lemoine, J. Roe, D. Stephens, T. Łuczyński, S. Xu, Y. Carreno, È. Pairet, *et al.*, "From market-ready rovs to low-cost auvs," in *OCEANS 2021: San Diego–Porto*. IEEE, 2021, pp. 1–7.

[20] E. Vargas, R. Scona, J. S. Willners, T. Luczynski, Y. Cao, S. Wang, and Y. R. Petillot, "Robust underwater visual slam fusing acoustic sensing," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 2140–2146.