Cramér-Rao bounds and optimal design metrics for pose-graph SLAM

Yongbo Chen, *Student Member, IEEE,* Kasra Khosoussi, Shoudong Huang, *Member, IEEE,* Liang Zhao, *Member, IEEE,* and Gamini Dissanayake, *Member, IEEE*

***Abstract*—2D/3D Pose-graph simultaneous localization and mapping (SLAM) is a problem of estimating a set of poses based on noisy measurements of relative rotations and translations. This paper focuses on the relation between the graphical structure of the pose-graph SLAM and its Fisher information matrix (FIM), Cramér-Rao lower bounds (CRLB) and its optimality design metrics (T-optimality and D-optimality). As a main contribution, based on the assumption of isotropic Langevin noise for rotation and block-isotropic Gaussian noise for translation, the FIM and CRLB are derived and shown to be closely related to the graph structure, in particular, the weighted Laplacian matrix. We also prove that total node degrees and weighted number of spanning trees, as two graph connectivity metrics, are respectively closely related to the trace and determinant of FIM. The discussions show that, compared with the D-optimality, the T-optimality metric is easy-computed but with a general effectiveness. We also present upper and lower bounds for the D-optimality metric, which can be efficiently computed and are almost independent of the estimation results. The proposed conclusions are verified with several well-known datasets, such as Intel, KITTI, sphere and so on.**

***Index Terms*—Pose-graph SLAM, Fisher information matrix (FIM), weighted Laplacian matrix, Cramér-Rao lower bounds (CRLB), optimality design metrics**

I. Introduction

SYNCHRONIZATION on the group of rigid body motions in two-dimensional (2D) plane and three-dimensional (3D) space, $\mathbb{R}^2 \times SO(2)$ and $\mathbb{R}^3 \times SO(3)$, is to estimate a set of poses based on noisy measurements of relative rotations and translations [1]. Multiple estimation problems, including pose-graph SLAM, fall into this category [2]. These synchronization problems in general give rise to a weighted graph representation. In essence, there's a correlation between the graphical structure of the 2D/3D pose-graph SLAM problem and its corresponding measurement network.

Given a pose-graph SLAM problem, assume we can obtain its "optimal" solution using graph-based optimization method, the next question to ask is how *reliable* that solution will be. Both the covariance matrix and the FIM can be used to assess the quality of the measurements network. However, compared with the FIM, updating/storing the dense full covariance matrix is prohibitively expensive with the growth of the state vector. Thus, the FIM is the top-priority index to evaluate the accuracy of the SLAM solution in the maximum likelihood (ML) estimate. In the state-of-the-art algorithms of SLAM, the SE-Sync [3] and Cartan-Sync [4] show the outstanding computational efficiency (more than an order of magnitude faster) compared with the Gauss-Newton based approach under the assumption that the rotation noise following the isotropic Langevin distribution on $SO(2)$ and $SO(3)$. Nevertheless, the FIM of the solution using these two methods is not presented in the literature. Beyond that, from the graphical point of view, we know that adding the number of relative measurements between the poses, which is equivalent to introducing new edges to the corresponding graph, helps to reduce the uncertainty of the estimator. In 2D pose-graph SLAM, the FIM has been shown to be closely related to the graph structure of the measurements network, in particular, the weighted Laplacian matrix [5] [6] assuming the Gaussian noise on the relative pose orientation and ignoring the wraparound issue. In this paper, as a main contribution, we will first define the FIM based on the assumption of zero-mean isotropic Langevin noise for orientation and block-isotropic Guassian noise for translation in 2D and 3D pose-graph SLAM, and then extend the results in our earlier work [5], [7] and [6].

It is known that the classical CRLB result for a flat Euclidean space provides us with a simple but strong relation between the covariance matrix $\boldsymbol{C}$ and the FIM $\boldsymbol{F}$: $\boldsymbol{C} \succeq \boldsymbol{F}^{-1}$ [8]. Because of the non-flat property of the parameter space of 3D pose-graph SLAM, its CRLB does not follow this simple expression. The curvature terms of the space will be introduced to show the strict CRLB. As another contribution, we derive the CRLB for 2D and 3D pose-graph SLAM based on the isotropic Langevin noise and block-isotropic Gaussian noise.

Because of the sparseness advantage of the FIM, the Theory of Optimal Experimental Design (TOED), including A-optimality, D-optimality, E-optimality and T-optimality on the FIM are widely used in decision making under uncertainty and belief space planning, with applications including autonomous driving, surveillance and active SLAM [2]. For example, in [9], the D-optimality, E-optimality and T-optimality metrics are used in the active visual object reconstruction, and the T-optimality metric is applied to solve the sensor selection problem in Large Sensor Networks [10] [11]. The TOED is closely related to the graphical structure of the block design [12]. In 2D pose/feature-graph SLAM with the block-isotropic Gaussian noise, the D-optimality metric can be bounded by the weighted number of the spanning trees [6]. Following this

Yongbo Chen, Shoudong Huang, Liang Zhao and Gamini Dissanayake are with Faculty of Engineering and Information Technology, University of Technology, Sydney, Ultimo, NSW, 2007 Australia, e-mail: Yongbo.Chen@student.uts.edu.au, Shoudong.Huang@uts.edu.au, Liang.Zhao@uts.edu.au, Gamini.Dissanayake@uts.edu.au.

Yongbo Chen is also with School of Aerospace Engineering, Beijing Institute of Technology, Beijing, 100081 China

Kasra Khosoussi is with Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology, Cambridge, MA, USA, email: kasra@mit.edu

idea, we extend it into 3D SLAM with $\mathbb{R}^3 \times SO(3)$ relative-pose measurements and also make a comparison between the T-optimality metric and the D-optimality metric from the point of view of the graphical structure of the measurement network.

A. Contributions

The main contributions of this paper are listed below.

- Deriving the formula of the FIM and the CRLB of the synchronization on the group of rigid body motions, $\mathbb{R}^2 \times SO(2)$ and $\mathbb{R}^3 \times SO(3)$, based on the assumption of the isotropic Lagevin noise (for rotation) and the block-isotropic Guassian noise (for translation);
- Exploring the relationship between the FIM of the 3D pose-graph SLAM and the weighted Laplacian matrix;
- Extending the analysis results of the D-optimality metric in [5] and [6] into 3D SLAM with $\mathbb{R}^3 \times SO(3)$ relative-pose measurements;
- Comparing the D-optimality and T-optimality metrics of 2D and 3D pose-graph SLAM from the perspective of the graph theory

B. Outline

In Section II, we review the related works. In Section III, we provide a mathematical formulation of the synchronization on the group of rigid body motions in 2D plane and 3D space, $\mathbb{R}^n \times SO(n),\ n = 2, 3$ (Pose-graph SLAM). In Section IV, we derive the FIM for the 2D Pose-graph SLAM and show its strong relationship with the weighted Laplacian matrix of the measurement graph. Furthermore, we obtain the FIM for the 3D situation and extend our relationship with the weighted Laplacian matrix into the $\mathbb{R}^3 \times SO(3)$ situation in Section V. In Section VI, the CRLB of the 2D and 3D pose-graph SLAM are derived. The TOED metrics, including D-optimality and T-optimality, of the FIM of the 3D pose-graph SLAM are discussed focusing on the graph topology in Section VII. The simulation results obtained using well-known datasets are provided in Section VIII. This is followed by the conclusion in Section IX. Finally, some details about the derivation and proofs are provided in the Appendix. [1]

C. Notation

Throughout this paper, bold lowercase and uppercase letters are reserved for vectors and matrices, respectively. Sets are shown by bold uppercase letters. $S_1 \succeq S_2$ means $S_1 - S_2$ is positive semidefinite. The Kronecker product is denoted by $\otimes$. trace$(\star)$ and $\det(\star)$ represent the trace and determinant value of the matrix $\star$. We denote by diag$(\boldsymbol{M}_1, ..., \boldsymbol{M}_k)$ the block-diagonal matrix with matrices $\boldsymbol{M}_1, ..., \boldsymbol{M}_k$ as blocks on its main diagonal. The squared Frobenius norm is $\| \star \|_F^2 =$ trace$(\star^\top \cdot \star)$. $\| \star \|_\infty$ means the biggest eigenvalue of $\star$. $dist(\star, \bullet) = \| \log(\star^\top \cdot \bullet) \|_F$ is the geodesic distance between $\star$ and $\bullet$ in $SO(n)$. The Euclidean norm is $\| \star \|^2 = \star^\top \cdot \star$, for a vector $\star$. $|\star|$ means the cardinality of the set $\star$. $\mathbb{E}\{\star\}$ means the mathematical expectation of the variable $\star$. $[n]$ denotes the set $\{i \in \mathbb{N}^+ : i \leq n\}$. $\nabla_\star \bullet$ means the partial derivative of $\bullet$ for the parameter $\star$, where $\bullet$ is a function with parameter $\star$. $\star \ltimes \bullet$ means the semi-product group of the group $\star$ and the group $\bullet$.

Finally, $SO(n)$ (special orthogonal group) is defined as:

$$SO(n) \triangleq \{\boldsymbol{R} \in \mathbb{R}^{n \times n} : \boldsymbol{R}^\top \boldsymbol{R} = \boldsymbol{I}_{n \times n}, \det(\boldsymbol{R}) = 1\}, \quad (1)$$

II. Related work

The pose-graph SLAM, which leads to a non-convex optimization problem, has gained considerable attention lately. Its globally optimal solution is the maximum likelihood (ML) estimate for the unknown poses. The well-known back-end methods roughly fall into two categories: The first one is to use the high-efficiency iterative nonlinear optimization methods based on the Gauss-Newton (GN) method, Levenberg-Marquardt (LM) method and Powells dogleg method, such as: g2o [13], iSAM2 [14], SLAM++ [15] and ceres [16], to obtain the locally optimal solutions. Because of the non-convex property of the pose-graph SLAM problem, starting from a poor initial solution, the iterative techniques may be trapped in a local minimum, which correspond to wrong estimates. The other one is to compute global optimal solutions via convex relaxations [3] [17] [18]. Some references have shown that the duality gap of the general pose-graph SLAM in practical applications is close to zero [19], which means that the solution of the relaxed problem using convex relaxations is equal to the solution of the original problem. Beside the pose-graph SLAM, the similar technologies have been widely used in the other robotics and computer vision applications, such as: 3D registration [20] and Structure from motion (SfM) [21].

Bounal [1] proposes the FIM for the estimation problems when the actual parameter space is a Riemannian submanifolds or a Riemannian quotient manifold. In [1], Bounal also shows two simple examples based on the isotropic Gaussian noise: the synchronization on the group of translation $\mathbb{R}^n$ and simple version of the synchronization on the group of rotation $SO(3)$. In his later work [23], the conclusions are extended into the general rotation group $SO(n)$ based on several kinds of Gaussian-like, but non-Gaussian, noise. Especially for the isotropic Langevin noise, this noise has attracted some robotics researchers' attention. As the state-of-the-art back-end algorithms, the convex relaxation based SE-Sync [3] and Cartan-Sync methods [4] are built based on the assumption of the Gaussian noise (for translation) and the isotropic Langevin noise (for rotation). To the best of our knowledge, the FIM for the pose-graph SLAM based on these noises, whose parameter space is a product manifold $\mathbb{R}^n \times SO(n),\ n = 2, 3$, hasn't been analyzed.

CRLB, as a classical tool in estimation theory [24], provides a lower bound on the variance of any unbiased estimator for an estimated problem [23]. The traditional CRLB is defined in a flat Euclidean space. Simith [25] extends the theory of CRLB into the general non-flat manifold. Because the FIM will become singular when no anchor is provided in the estimation problem, Xavier and Barroso [26] use the pseudoinverse of the FIM for this anchor-free case. Based on these new extended

[1] Our results can be straightforwardly generalized to 2D/3D feature-based SLAM problems.

tools, Boumal presents the CRLB for the synchronization of rotations $SO(n)$ in both anchored and anchor-free cases. The pose-graph SLAM is an anchored estimation problem in a product manifold $\mathbb{R}^n \times SO(n),\ n = 2, 3$ commonly. As an extension, we will present its CRLB in this paper.

In fact, there are two different Lie group representations corresponding to the 2D/3D pose-graph SLAM: $\mathbb{R}^n \times SO(n)$ and $SE(n) = \mathbb{R}^n \ltimes SO(n)$. The direct product of $SO(n)$ and $\mathbb{R}^n$ manifolds $\mathbb{R}^n \times SO(n)$ can be represented as a $(2n+1) \times (2n+1)$ matrix with $n + \frac{n(n-1)}{2}$-dimensional minimal representation. The $\mathbb{R}^n$ and $SO(n)$ group can be considered as two separated parts easily, because of the separated Riemannian structure (decided by inner product) [27]. The special Euclidean group $SE(n)$ is isomorphic to $\mathbb{R}^n \times SO(n)$ but with different Riemannian structure (complex and coupled inner product) [28]. It can be represented as a $(n+1) \times (n+1)$ homogeneous transformation matrix with $n + \frac{n(n-1)}{2}$-dimensional minimal representation. Because of the different Riemannian structure, these two groups have the different tangent spaces and gradient forms, which leads to different FIM and CRLB. Compared with $SE(n)$ group, these concepts on the direct product group $\mathbb{R}^n \times SO(n)$ are easier to analyze. Meanwhile, the objective function of the SE-sync method is built based on the direct product group $\mathbb{R}^n \times SO(n)$. Hence, in this paper, we only consider our problem on the direct product group $\mathbb{R}^n \times SO(n)$.

For the TOED based on the FIM, a comparison of these optimality criteria is presented in [29] [30]. The related work in [29] shows that only the D-optimality metric retains the monotonicity during the exploration phase of the robot for the linearized framework. In recent work, the authors show that the monotonicity of all optimality criteria, A-optimality, D-optimality, and E-optimality, and Shannons entropy is greatly affected by the uncertainty representation [31]. As the most common representation, the uncertainty representations on the Lie group SE(3) based on the Gaussian noise have also been considered [32]. In these metrics, the D-optimality metric, which is the most popular metric with a good performance [29], and T-optimality metrics do not need to finish the inverse operation with a lower computational complexity. However, from a computational complexity perspective, the T-optimality metric still has its obvious advantage compared with the D-optimality metric, even using the state-of-the-art incremental technology for the D-optimality metric [33]. So both two metrics have great significance to the different mission requirements. We will investigate them based on the graphical structure of the SLAM.

In our previous work [5] [6] and [7], we explore the impact of the graphical structure of SLAM on some of the desirable attributes of some estimation problems: difference-network, compass-SLAM and 2D pose-graph SLAM with block-isotropic Gaussian noise. In the difference-network and the compass-SLAM, the FIM is proportional to the reduced Laplacian matrix of the corresponding graph, which helps to directly connect the optimal design of the FIM with the TOED of the measurement graph. For the 2D pose-graph SLAM with the block-isotropic Gaussian noise, we state that the D-optimality metric of the FIM can be bounded by the weighted number of spanning trees of the measurement graph(weighted tree-connectivity). Based on the lower bounds, we put forward a new near-t-optimal graph synthesis frameworks for the measurement selection, the pose-graph pruning problems and the D-optimality-aware SLAM front-end. In this paper, we extend our analysis and conclusions into the 2D and 3D pose-SLAM with $\mathbb{R}^n \times SO(n),\ n = 2, 3$ relative-pose measurements by the rigorous mathematical derivation.

III. Synchronization on $\mathbb{R}^n \times SO(n)$, pose-graph SLAM

A. Graph Preliminaries

A directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, which is weakly connected, is used to represent the synchronization on pose-graph SLAM problem naturally. In this paper, $\mathcal{V} = \{1, 2, \cdots, n_p\} \bigcup \{0\}$, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ and $|\mathcal{E}| = m$. Each node denotes a robot pose, and each edge $(i_k, j_k) \in \mathcal{E}$ represents a relative measurement between two robot poses. An un-directed graph $\mathcal{G}_1 = (\mathcal{V}, \mathcal{F})$, corresponding to $\mathcal{G}$, has the same nodes and the same un-directed edges with the directed graph $\mathcal{G}$.

We denote the i-th node degree by d_i. For i-th node, we can define three node sets V_i^+, V_i^- and V_i satisfying $(i,j) \in \mathcal{E} \Leftrightarrow j \in V_i^+$, $(j,i) \in \mathcal{E} \Leftrightarrow j \in V_i^-$ and $V_i = V_i^- \bigcup V_i^+$, so we have $|V_i^+| + |V_i^-| = |V_i| = d_i$. Due to the relative nature of measurements, we can define a global coordinates system and *anchor* one of the nodes to it. Without loss of generality, the first pose is assumed to be the origin of our global coordinates system.

The incidence matrix of $\mathcal{G}$ is denoted by $\boldsymbol{A}_0 \in \{-1, 0, 1\}^{(n_p+1)\times m}$. $a_{ik} = -1$ and $a_{jk} = 1$ (the (i,k)-th and (j,k)-th element of $\boldsymbol{A}_0$) are non-zero, if there is an edge $e = (i,j) \in \mathcal{E}$. The incidence matrix after anchoring $\boldsymbol{x}_0$ to the origin, $\boldsymbol{A} \in \{-1, 0, 1\}^{n_p \times m}$, is obtained simply by removing the row corresponding to the first node in $\boldsymbol{A}_0$. The Laplacian matrix and the reduced Laplacian matrix of $\mathcal{G}$ are respectively defined as $\boldsymbol{L}_0 \triangleq \boldsymbol{A}_0\boldsymbol{A}_0^\top$ and $\boldsymbol{L} \triangleq \boldsymbol{A}\boldsymbol{A}^\top$. It can be shown that $\boldsymbol{L}_0$ and $\boldsymbol{L}$ are respectively positive semi-definite and positive definite, if $\mathcal{G}$ is (weakly) connected. The Laplacian matrix and the reduced Laplacian matrix of $\mathcal{G}$ can be written as $\boldsymbol{L}_0 = \boldsymbol{D}_0 - \boldsymbol{W}_0$ and $\boldsymbol{L} = \boldsymbol{D} - \boldsymbol{W}$, where $\boldsymbol{D}_0 \triangleq \mathrm{diag}(d_0, d_1, \ldots, d_n)$, $\boldsymbol{D} \triangleq \mathrm{diag}(d_1, \ldots, d_n)$, $\boldsymbol{W}_0$ and $\boldsymbol{W}$ are respectively the original adjacency matrix and the adjacency matrix of the graph after removing the row and column corresponding to $\boldsymbol{x}_0$. The weighted Laplacian matrix and the weighted reduced Laplacian matrix are defined as $\boldsymbol{L}_\omega^0 \triangleq \boldsymbol{A}_0\boldsymbol{\Sigma}_0\boldsymbol{A}_0^\top$ and $\boldsymbol{L}_\omega \triangleq \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}^\top$, where $\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}_0$ are diagonal matrices whose diagonal elements are the weighted value of the graph edge.

B. Synchronization on $\mathbb{R}^n \times SO(n)$

Synchronization on the group of the rigid body motions in 2D plane and 3D space, $\mathbb{R}^n \times SO(n),\ n = 2, 3$, is the problem of estimating a set of positions $\boldsymbol{x}_0,\ \boldsymbol{x}_1, \cdots,\ \boldsymbol{x}_{n_p} \in \mathbb{R}^n$ and rotations $\boldsymbol{R}_0,\ \boldsymbol{R}_1, \cdots,\ \boldsymbol{R}_{n_p} \in SO(n)$ from the noisy measurements of some relative rotations $\boldsymbol{R}_j\boldsymbol{R}_i^\top$ and the relative coordinate transformations $\boldsymbol{R}_i^\top(\boldsymbol{x}_j - \boldsymbol{x}_i)$.

In 2D/3D pose-graph SLAM, considering the anchored pose, the parameter space is: $\mathcal{P} = \{\mathbb{R}^n \times \cdots \times \mathbb{R}^n\}_{n_p+1} \times \{SO(n) \times \cdots \times SO(n)\}_{n_p+1}$. We can create a new rotation graph $\mathcal{G}_1 = (\mathcal{V}, \mathcal{F})$, whose nodes only represent the rotations of the poses, by the pose graph $\mathcal{G}$. The rotation graph $\mathcal{G}_1$ is un-directed, because the measurements (i, j) introduce the same information for i-th node and j-th node.

For the pose-graph edge, $(i, j) \in \mathcal{E}$, we have a measurement $\boldsymbol{p}_{ij} \in \mathbb{R}^n$ of the relative noisy measurement between i-th and j-th poses:

$$\boldsymbol{p}_{ij} = \boldsymbol{R}_i^\top (\boldsymbol{x}_j - \boldsymbol{x}_i) + \boldsymbol{y}_{ij} \tag{2}$$

where $\boldsymbol{y}_{ij}$ is a random vector whose distributed function $f_{ij} : \mathbb{R}^n \to \mathbb{R}^+$ meets an isotropic Gaussian distribution:

$$\begin{aligned} f_{ij}(\boldsymbol{y}_{ij}) =& \frac{1}{2\pi \ \det(\boldsymbol{\Sigma}_{ij})^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{y}_{ij}^\top \boldsymbol{\Sigma}_{ij}^{-1} \boldsymbol{y}_{ij})\right) \\ \boldsymbol{\Sigma}_{ij} =& \delta_{ij}^2 \boldsymbol{I}_{n\times n} \end{aligned} \tag{3}$$

where δ_{ij} means the isotropic variance value. We write $\boldsymbol{y}_{ij} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_{ij})$ to mean that $\boldsymbol{y}_{ij}$ is a random vector with probability density functions (PDF) (3).

For the edge in the rotation graph $(i, j) \in \mathcal{F}$, we have the relative noisy rotation $\boldsymbol{H}_{ij} \in SO(n)$ between $\boldsymbol{R}_i$ and $\boldsymbol{R}_j$:

$$\boldsymbol{H}_{ij} = \boldsymbol{Z}_{ij} \boldsymbol{R}_j \boldsymbol{R}_i^\top \tag{4}$$

where $\boldsymbol{Z}_{ij}$ is a random rotation whose distributed function $\widehat{f}_{ij} : SO(n) \to \mathbb{R}^+$ meets an isotropic Langevin distribution with mean $\boldsymbol{I}_{n\times n}$ and concentration $\kappa_{ij} \geq 0$ [23]:

$$\begin{aligned} \widehat{f}_{ij}(\boldsymbol{Z}_{ij}) =& \frac{1}{c_n(\kappa_{ij})} \exp\left(\kappa_{ij} \ \text{trace}(\boldsymbol{Z}_{ij})\right) \\ c_2(\kappa_{ij}) =& I_0(2\kappa_{ij}) \\ c_3(\kappa_{ij}) =& \exp(\kappa_{ij}) \left(I_0(2\kappa_{ij}) - I_1(2\kappa_{ij})\right) \\ I_v(2\kappa_{ij}) =& \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp\left(2\kappa_{ij} \cos(\theta)\right) \cos(v\theta) d\theta \end{aligned} \tag{5}$$

where $c_n(\kappa_{ij}),\ n = 2, 3$ is a normalization constant such that $\widehat{f}_{ij}$ has unit mass. $I_v(2\kappa_{ij}),\ v = \{0, 1, 2, \cdots\} \in \mathbb{Z}$ means the modified Bessel functions [34]. We write $\boldsymbol{Z}_{ij} \sim Lang(\boldsymbol{I}_{n\times n}, \kappa_{ij})$ to mean that $\boldsymbol{Z}_{ij}$ is a random variable with PDF (5). This PDF meets the assumptions shown in [23]: 1. smoothness and support; 2. independence; 3. invariance assumptions.

Under the independence assumption, the log-likelihood of an estimator $\boldsymbol{\theta} = \boldsymbol{x} \times \boldsymbol{R} = (\boldsymbol{x}_0, \cdots, \boldsymbol{x}_{n_p}, \boldsymbol{R}_0, \cdots, \boldsymbol{R}_{n_p}) \in \mathcal{P}$, given the measurements $\boldsymbol{p}_{ij}$ and $\boldsymbol{H}_{ij}$, is given by:

$$\begin{aligned} & L(\boldsymbol{x}, \boldsymbol{R}) \\ =& \sum_{j \in V_i^+} \log f_{ij}(\boldsymbol{y}_{ij}) + \frac{1}{2} \sum_{(i,j)\in\mathcal{F}} \log \widehat{f}_{ij}(\boldsymbol{Z}_{ij}) \\ +& \sum_{j \in V_i^-} \log f_{ji}(\boldsymbol{y}_{ji}) \end{aligned} \tag{6}$$

The coefficient $\frac{1}{2}$ is used to balance the information in the un-directed rotation graph $\mathcal{F}$ and the directed pose-graph $\mathcal{G}$.

C. Geometry of the parameter spaces

The FIM is a classical tool for estimation problems on Euclidean spaces. In order to define the FIM on a manifold $\mathbb{R}^n \times SO(n),\ n = 2, 3$, we need to define some notions and tools to describe the parameter spaces for the synchronization.

1) Tangent space on $SO(n)$: As a Lie group, the dimension of its least representation is $d = \frac{n(n-1)}{2}$. We can admit a tangent space $\mathcal{T}_{\boldsymbol{Q}} SO(n),\ n = 2, 3$ for each rotation:

$$\mathcal{T}_{\boldsymbol{Q}} SO(n) = \boldsymbol{Q} so(n) \triangleq \{\boldsymbol{Q}\boldsymbol{\Omega} : \boldsymbol{\Omega} \in \mathbb{R}^{n\times n}, \boldsymbol{\Omega}^\top + \boldsymbol{\Omega} = 0\} \tag{7}$$

where $\boldsymbol{Q} \in SO(n) \in \mathbb{R}^{n\times n}$.

2) Inner product on $\mathbb{R}^n$ and $SO(n)$: Based the Riemannian metric of the manifold $\mathbb{R}^n \times SO(n)$, we define the inner products on the tangent space on $\mathbb{R}^n$ and $SO(n)$ respectively are:

$$< \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2 >_{\boldsymbol{\theta}} = \begin{cases} \boldsymbol{\Omega}_1^\top \boldsymbol{\Omega}_2, & \boldsymbol{\theta} = \boldsymbol{X} \\ \text{trace}(\boldsymbol{\Omega}_1^\top \boldsymbol{\Omega}_2), & \boldsymbol{\theta} = \boldsymbol{R} \end{cases} \tag{8}$$

The Riemannian metric of $\mathbb{R}^n \times SO(n)$ is the sum of the Riemannian metrics of $\mathbb{R}^n$ and $SO(n)$.

3) Gradient on $SO(n)$: Let $h : \mathbb{R}^{n\times n} \to \mathbb{R}$ is a differential function, we can define the gradient of h by:

$$\begin{aligned} \text{grad } h(\boldsymbol{Q}) &= \boldsymbol{Q}\text{skew}(\boldsymbol{Q}^\top \nabla h(\boldsymbol{Q})) \\ \text{skew}(\star) &\triangleq (\star - \star^\top)/2 \end{aligned} \tag{9}$$

where $\nabla h(\boldsymbol{Q})$ means the gradient of h seen as a Euclidean function in $\mathbb{R}^{n\times n}$.

The directional derivative of h at $\boldsymbol{Q}$ along $\boldsymbol{Q}\boldsymbol{\Omega}$ can be written as:

$$< \text{grad } h(\boldsymbol{Q}), \boldsymbol{Q}\boldsymbol{\Omega} >_{\boldsymbol{R}} \tag{10}$$

where $\boldsymbol{Q}\boldsymbol{\Omega}$ is a tangent vector in the tangent space.

D. FIM

Definition [23]: Let $\boldsymbol{\theta} \in \mathcal{P}$ be unknown parameter and $h(y; \boldsymbol{\theta})$ be the PDF of the measurement y conditioned by $\boldsymbol{\theta}$, shown in (3) and (5). Based on the log-likelihood function $L(\boldsymbol{\theta}) = \log\ h(y; \boldsymbol{\theta})$ shown in (6) and the orthonormal basis, the (i, i_1)-th block of the FIM is defined as:

$$\boldsymbol{F}_{i,i_1} = \mathbb{E}\left\{< \text{grad} L(\boldsymbol{\theta}), e_i >_{\boldsymbol{\theta}} \cdot < \text{grad} L(\boldsymbol{\theta}), e_{i_1} >_{\boldsymbol{\theta}}\right\} \tag{11}$$

where e_i and e_{i_1} are the bases of the i-th and i_1-th parameters. They will be defined based on the parameter space [1]. It is noted that the FIM is directly decided by the bases. For a parameter space, there may exist more than one kind of basis.

IV. FIM FOR 2D POSE-GRAPH SLAM

A. Geometry of the parameter spaces

As shown in Section III-C, the tools, including the tangent space, the inner product and the gradient, are completely suitable for both the 2D ($\mathbb{R}^2 \times SO(2)$) and 3D ($\mathbb{R}^3 \times SO(3)$) parameter space. In this part, we will show the orthonormal basis for the group $\mathbb{R}^2 \times SO(2)$.

1) Orthonormal basis for $\mathbb{R}^2 \times SO(2)$: The orthonormal basis $\boldsymbol{E}^{\boldsymbol{x},\boldsymbol{R}} = (\boldsymbol{E}_0^{\boldsymbol{x}}, \cdots, \boldsymbol{E}_{n_p}^{\boldsymbol{x}}, \boldsymbol{E}_0^{\boldsymbol{R}}, \cdots, \boldsymbol{E}_{n_p}^{\boldsymbol{R}})$ of the tangent space $\mathbb{R}^2 \times \mathcal{T}_{\boldsymbol{R}}SO(2)$ is:

$$\begin{aligned}
\boldsymbol{E}_i^{\boldsymbol{x}} &= (\boldsymbol{E}_i^{X\top}, \mathbf{0}_{2\times 2(n_p+1)}^\top)^\top,\ i \in [n_p] \bigcup \{0\} \\
\boldsymbol{E}_j^{\boldsymbol{R}} &= (\mathbf{0}_{2\times 2(n_p+1)}^\top, \boldsymbol{E}_j^{R\top})^\top,\ j \in [n_p] \bigcup \{0\} \\
\boldsymbol{E}_i^X &= (\mathbf{0}, \cdots, \mathbf{0}, \underbrace{\boldsymbol{I}_{2\times 2}}_{i-\text{th}}, \mathbf{0}, \cdots, \mathbf{0})_{2\times 2(n_p+1)} \\
\boldsymbol{E}_j^R &= (\mathbf{0}, \cdots, \mathbf{0}, \underbrace{\boldsymbol{R}_j\boldsymbol{E}}_{j-\text{th}}, \mathbf{0}, \cdots, \mathbf{0})_{2\times 2(n_p+1)} \\
\boldsymbol{E} &= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}
\end{aligned} \tag{12}$$

B. FIM

Because of the group $\mathbb{R}^2 \times SO(2)$, there are four different parts in the FIM: One sub-matrix corresponding to the Euclidean space $\mathbb{R}^2$. One sub-matrix corresponding to $SO(2)$ Lie group, and other two transposition relation sub-matrices combined the Euclidean space $\mathbb{R}^2$ and $SO(2)$ Lie group. They are shown as follows:

Lemma 1. *X and Y are random variables, when they are independent, we have $\mathbb{E}\{XY\} = \mathbb{E}\{X\}\mathbb{E}\{Y\}$ [22].*

1) $\mathbb{R}^2$ sub-matrix: Based on the definition (11), we have:

$$\begin{aligned}
\boldsymbol{F}_{i,i_1} =& \mathbb{E}\{< \text{grad}L(\boldsymbol{\theta}), \boldsymbol{E}_i^{\boldsymbol{x}} >_{\boldsymbol{X}} \cdot < \text{grad}L(\boldsymbol{\theta}), \boldsymbol{E}_{i_1}^{\boldsymbol{x}} >_{\boldsymbol{X}}^\top\} \\
=& \mathbb{E}\{\text{grad}_i L(\boldsymbol{x})^\top \cdot \boldsymbol{I}_{2\times 2} \cdot \boldsymbol{I}_{2\times 2}^\top \cdot \text{grad}_{i_1} L(\boldsymbol{x})\} \\
=& \mathbb{E}\{\text{grad}_i L(\boldsymbol{x})^\top \cdot \text{grad}_{i_1} L(\boldsymbol{x})\}
\end{aligned} \tag{13}$$

where $\text{grad}_i L(\boldsymbol{x})$ means the gradient for parameter $\boldsymbol{x}_i$.

Based on (6), we have:

$$\begin{aligned}
&\text{grad}_i L(\boldsymbol{x}) \\
=& \text{grad}_i \sum_{j\in V_i^+} \log f_{ij}(\boldsymbol{y}_{ij}) + \text{grad}_i \sum_{j\in V_i^-} \log f_{ji}(\boldsymbol{y}_{ji}) \\
=& \sum_{j\in V_i^+} \nabla_{\boldsymbol{y}_{ij}} f_{ij}(\boldsymbol{y}_{ij}) \nabla_{\boldsymbol{x}_i} \boldsymbol{y}_{ij} + \sum_{j\in V_i^-} \nabla_{\boldsymbol{y}_{ji}} f_{ji}(\boldsymbol{y}_{ji}) \nabla_{\boldsymbol{x}_i} \boldsymbol{y}_{ji} \\
=& \sum_{j\in V_i^+} -\delta_{ij}^{-2} \boldsymbol{y}_{ij}^\top \boldsymbol{R}_i^\top + \sum_{j\in V_i^-} \delta_{ji}^{-2} \boldsymbol{y}_{ji}^\top \boldsymbol{R}_j^\top
\end{aligned} \tag{14}$$

When $i = i_1$, based on Lemma.1 and $\mathbb{E}\{\boldsymbol{y}_{ij}\} = \mathbb{E}\{\boldsymbol{y}_{ji}\} = \mathbf{0}$, substitute (14) into (13), we have:

$$\begin{aligned}
\boldsymbol{F}_{i,i_1} =& \mathbb{E}\{ \sum_{j\in V_i^+} \delta_{ij}^{-4} \boldsymbol{R}_i \boldsymbol{y}_{ij} \boldsymbol{y}_{ij}^\top \boldsymbol{R}_i^\top + \mathbf{0} \\
&+ \mathbf{0} + \sum_{j\in V_i^-} \delta_{ji}^{-4} \boldsymbol{R}_j \boldsymbol{y}_{ji} \boldsymbol{y}_{ji}^\top \boldsymbol{R}_j^\top \}
\end{aligned} \tag{15}$$

Because $\mathbb{E}\{\boldsymbol{y}_{ij}\boldsymbol{y}_{ij}^\top\} = \delta_{ij}^{-2}\boldsymbol{I}_{2\times 2}$ and $\mathbb{E}\{\boldsymbol{y}_{ji}\boldsymbol{y}_{ji}^\top\} = \delta_{ji}^{-2}\boldsymbol{I}_{2\times 2}$, we have:

$$\boldsymbol{F}_{i,i_1} = \sum_{j\in V_i^+} \delta_{ij}^{-2} \boldsymbol{I}_{2\times 2} + \sum_{j\in V_i^-} \delta_{ij}^{-2} \boldsymbol{I}_{2\times 2} \tag{16}$$

When $(i, i_1) \in \mathcal{E}$, based on Lemma.1 and $\mathbb{E}\{\boldsymbol{y}_{ij}\} = \mathbb{E}\{\boldsymbol{y}_{ji}\} = \mathbf{0}$, substitute (14) into (13), we have:

$$\begin{aligned}
\boldsymbol{F}_{i,i_1} =& \mathbb{E}\{\left(-\delta_{ii_1}^{-2} \boldsymbol{y}_{ii_1}^\top \boldsymbol{R}_i^\top\right) \cdot \left(\delta_{ii_1}^{-2} \boldsymbol{y}_{ii_1}^\top \boldsymbol{R}_i^\top\right)^\top\} \\
=& -\delta_{ii_1}^{-2} \boldsymbol{I}_{2\times 2}
\end{aligned} \tag{17}$$

When $(i_1, i) \in \mathcal{E}$, based on (17), we have

$$\boldsymbol{F}_{i_1,i} = \boldsymbol{F}_{i,i_1}^\top = -\delta_{i_1 i}^{-2} \boldsymbol{I}_{2\times 2} \tag{18}$$

Combine (16), (17) and (18), we have:

$$\boldsymbol{F}_{i,i_1} = \begin{cases} \sum_{j_1\in V_i^+} \sigma_{ij_1}^{-2} \boldsymbol{I}_{2\times 2} + \sum_{j_2\in V_i^-} \sigma_{j_2 i}^{-2} \boldsymbol{I}_{2\times 2}, & i = i_1 \\ -\sigma_{ii_1}^{-2} \boldsymbol{I}_{2\times 2}, & (i, i_1) \in \mathcal{E} \\ -\sigma_{i_1 i}^{-2} \boldsymbol{I}_{2\times 2}, & (i_1, i) \in \mathcal{E} \\ \mathbf{0}, & else \end{cases} \tag{19}$$

It is easy to find that (19) equals to the weighted Laplacian matrix after using the Kronecker product operation for $\boldsymbol{I}_{2\times 2}$. We can denote this sub-FIM corresponding to the Euclidean space $\mathbb{R}^2$ as $\boldsymbol{L}_w^{\mathbb{R}^2}$.

2) $SO(2) \times \mathbb{R}^2$ sub-matrix: Based on the geometry of the parameter space and the definition in (11), we can get the block of the $SO(2) \times \mathbb{R}^2$ sub-matrix:

$$\begin{aligned}
&\boldsymbol{F}_{n_p+1+i,i_1} \\
=& \mathbb{E}\{< \text{grad}L(\boldsymbol{\theta}), \boldsymbol{E}_i^{\boldsymbol{R}} >_{\boldsymbol{R}} \cdot < \text{grad}L(\boldsymbol{\theta}), \boldsymbol{E}_{i_1}^{\boldsymbol{x}} >_{\boldsymbol{X}}^\top\} \\
=& \mathbb{E}\{\text{trace}(\text{grad}_i L(\boldsymbol{R})^\top \cdot \boldsymbol{R}_i \boldsymbol{E}) \cdot \boldsymbol{I}_{2\times 2}^\top \cdot \text{grad}_{i_1} L(\boldsymbol{x})\}
\end{aligned} \tag{20}$$

where $\text{grad}_i L(\boldsymbol{R})$ means the gradient for parameter $\boldsymbol{R}_i$.

We have:

$$\begin{aligned}
&\text{grad}_i L(\boldsymbol{R}) \\
=& \text{grad}_i \left(\frac{1}{2} \sum_{(i,j)\in\mathcal{F}} \log \widehat{f}_{ij}(\boldsymbol{Z}_{ij}) + \sum_{j\in V_i^+} \log f_{ij}(\boldsymbol{y}_{ij}) \right)
\end{aligned} \tag{21}$$

For the first part $\text{grad}_i \frac{1}{2} \sum_{(i,j)\in\mathcal{F}} \log \widehat{f}_{ij}(\boldsymbol{Z}_{ij})$, Based on (5), we have:

$$\begin{aligned}
&\text{grad}_i \frac{1}{2} \sum_{(i,j)\in\mathcal{F}} \log \widehat{f}_{ij}(\boldsymbol{Z}_{ij}) \\
=& \text{grad}_i \sum_{j\in V_i} \log \left(\frac{1}{c_2(\kappa_{ij})} \exp(\kappa_{ij} \text{trace}(\boldsymbol{Z}_{ij})) \right)
\end{aligned} \tag{22}$$

It is noted that, because of the un-directed graph $\mathcal{F}$ and the coefficient $\frac{1}{2}$, all nodes in set V_i can be seen as a new directed graph meeting $(i, j) \in \mathcal{E} \Leftrightarrow j \in V_i$.

Set $f(\star) = \frac{1}{c_2(\kappa_{ij})} \exp(\kappa_{ij}\ \text{trace}(\star))$ and $\boldsymbol{H} = \boldsymbol{Z}_{ij}\boldsymbol{R}_i^\top$, we have:

$$\text{grad}_i \sum_{j\in V_i} \log \widehat{f}_{ij}(\boldsymbol{Z}_{ij}) = \sum_{j\in V_i} \text{grad}_i \log\left(f(\boldsymbol{H}\boldsymbol{R}_i)\right) \tag{23}$$

Based on Appendix A, the gradient of $log\left(f(\boldsymbol{H}\boldsymbol{R}_i)\right)$ can be computed as:

$$\begin{aligned}
&\text{grad}_i\ \log f(\boldsymbol{H}\boldsymbol{R}_i) \\
=& \boldsymbol{R}_i \text{skew}(\boldsymbol{R}_i^\top (f(\boldsymbol{H}\boldsymbol{R}_i))^{-1} \boldsymbol{H}^\top \bigtriangledown_{\boldsymbol{H}\boldsymbol{R}_i} f(\boldsymbol{H}\boldsymbol{R}_i))
\end{aligned} \tag{24}$$

Substitute $f(\star) = \frac{1}{c_2(\kappa_{ij})}\exp(\kappa_{ij}\ \text{trace}(\star))$ and $\boldsymbol{H} = \boldsymbol{Z}_{ij}\boldsymbol{R}_i^\top$ into (24) and (23), we have:

$$\text{grad}_i \sum_{j\in V_i} \log \widehat{f}_{ij}(\boldsymbol{Z}_{ij}) = \sum_{j\in V_i} -\kappa_{ij}\boldsymbol{R}_i \text{skew}(\boldsymbol{Z}_{ij}) \tag{25}$$

Let's consider the second part in (21), we can first get the usual gradient of $f_{ij}(\boldsymbol{y}_{ij})$ in $\mathbb{R}^{n\times n}$:

$$\begin{aligned}
&\nabla\ log\ f_{ij}(\boldsymbol{y}_{ij}) \\
=&\nabla_{\boldsymbol{R}_i\in\mathbb{R}^{n\times n}} \log \frac{1}{2\pi|\boldsymbol{\Sigma}|^{1/2}}\exp(-\frac{1}{2}\boldsymbol{y}_{ij}(\boldsymbol{R}_i)^\top\boldsymbol{\Sigma}_{ij}^{-1}\boldsymbol{y}_{ij}(\boldsymbol{R}_i))) \\
=&\nabla_{\boldsymbol{R}_i\in\mathbb{R}^{n\times n}} \frac{(\boldsymbol{p}_{ij}-\boldsymbol{R}_i^\top(\boldsymbol{x}_j-\boldsymbol{x}_i))^\top\boldsymbol{\Sigma}_{ij}^{-1}(\boldsymbol{p}_{ij}-\boldsymbol{R}_i^\top(\boldsymbol{x}_j-\boldsymbol{x}_i))}{2} \\
=&\nabla_{\boldsymbol{R}_i\in\mathbb{R}^{n\times n}} \delta_{ij}^{-2}\frac{1}{2}\boldsymbol{y}_{ij}(\boldsymbol{R}_i)\boldsymbol{y}_{ij}(\boldsymbol{R}_i) \\
=&\delta_{ij}^{-2}(\boldsymbol{x}_j-\boldsymbol{x}_i)\boldsymbol{y}_{ij}^\top
\end{aligned} \tag{26}$$

Based on the general definition of the gradient on $SO(2)$ shown in (9), let $\boldsymbol{Q} = \boldsymbol{R}_i$, we have:

$$\begin{aligned}
&\text{grad}\ log f_{ij}(\boldsymbol{R}_i) \\
=&\boldsymbol{R}_i skew(\boldsymbol{R}_i^\top\cdot\nabla f_{ij}(\boldsymbol{y}_{ij}(\boldsymbol{R}_i))) \\
=&\delta_{ij}^{-2}((\boldsymbol{x}_j-\boldsymbol{x}_i)\boldsymbol{y}_{ij}^\top - \boldsymbol{R}_i\boldsymbol{y}_{ij}(\boldsymbol{x}_j-\boldsymbol{x}_i)^\top\boldsymbol{R}_i)/2
\end{aligned} \tag{27}$$

Re-write the above equation (27) based on the basis $\boldsymbol{R}_i\boldsymbol{E}$, we have [35]:

$$\begin{aligned}
&\sum_{j\in V_i^+}\text{grad}_i \log f_{ij}(\boldsymbol{y}_{ij}) \\
=&\sum_{j\in V_i^+}\frac{1}{2}\boldsymbol{R}_i\boldsymbol{E}\nabla_{\boldsymbol{R}_i}\log f_{ij}(\boldsymbol{y}_{ij}) \\
=&\sum_{j\in V_i^+}\frac{1}{2}\boldsymbol{R}_i\boldsymbol{E}\delta_{ij}^{-2}\boldsymbol{y}_{ij}^\top\nabla_{\boldsymbol{R}_i}\boldsymbol{y}_{ij} \\
=&\sum_{j\in V_i^+}\frac{1}{2}\delta_{ij}^{-2}\boldsymbol{R}_i\boldsymbol{E}\boldsymbol{y}_{ji}^\top\boldsymbol{E}^\top\boldsymbol{R}_i^\top(\boldsymbol{x}_i-\boldsymbol{x}_j)
\end{aligned} \tag{28}$$

For $\text{grad}_{i_1}L(\boldsymbol{x})$, based on (14), we have:

$$\begin{aligned}
&\text{grad}_{i_1}L(\boldsymbol{x}) \\
=&\sum_{j_1\in V_{i_1}^+}\delta_{i_1j_1}^{-2}\boldsymbol{y}_{i_1j_1}^\top\boldsymbol{R}_{i_1}^\top - \sum_{j_1\in V_{i_1}^-}\delta_{j_1i_1}^{-2}\boldsymbol{y}_{j_1i_1}^\top\boldsymbol{R}_{j_1}^\top
\end{aligned} \tag{29}$$

Substitute (28), (29) and (25) into (20), let $g_{ij} = (\boldsymbol{x}_i-\boldsymbol{x}_j)^\top\boldsymbol{R}_i\boldsymbol{E}^\top\boldsymbol{y}_{ij} \in \mathbb{R}$, based on Lemma 1 and $\mathbb{E}\{\boldsymbol{Z}_{ij}\} = \boldsymbol{0}$, we have:

$$\begin{aligned}
&\boldsymbol{F}_{n_p+1+i,i_1} \\
=&\mathbb{E}\{\text{trace}(\text{grad}_i L(\boldsymbol{R})^\top\cdot\boldsymbol{R}_i\boldsymbol{E})\cdot\boldsymbol{I}_{2\times2}^\top\cdot\text{grad}_{i_1}L(\boldsymbol{x})\} \\
=&\boldsymbol{0} + \sum_{j\in V_i^+}\sum_{j_1\in V_{i_1}^+}\mathbb{E}\{\delta_{ij}^{-2}\delta_{i_1j_1}^{-2}\text{trace}\,(g_{ij})\,\boldsymbol{y}_{i_1j_1}^\top\boldsymbol{R}_{i_1}^\top\} \\
&-\sum_{j\in V_i^+}\sum_{j_1\in V_{i_1}^-}\mathbb{E}\{\delta_{ij}^{-2}\delta_{j_1i_1}^{-2}\text{trace}\,(g_{ij})\,\boldsymbol{y}_{j_1i_1}^\top\boldsymbol{R}_{j_1}^\top\}
\end{aligned} \tag{30}$$

a) : When $(i,j) = (i_1,j_1) \in \mathcal{E}$, we have:

$$\begin{aligned}
&\boldsymbol{F}_{n_p+1+i,i_1} \\
=&\sum_{j\in V_i^+}\sum_{j_1\in V_{i_1}^+}\mathbb{E}\{\delta_{ij}^{-2}\delta_{i_1j_1}^{-2}\text{trace}\,(g_{ij})\,\boldsymbol{y}_{i_1j_1}^\top\boldsymbol{R}_{i_1}^\top\} + \boldsymbol{0} \\
=&\sum_{j\in V_i^+}\delta_{ij}^{-4}(\boldsymbol{x}_i-\boldsymbol{x}_j)^\top\boldsymbol{E}\boldsymbol{R}_i\mathbb{E}\{\boldsymbol{y}_{ij}\boldsymbol{y}_{ij}^\top\}\boldsymbol{R}_i^\top \\
=&\sum_{j\in V_i^+}\delta_{ij}^{-2}(\boldsymbol{x}_i-\boldsymbol{x}_j)^\top\boldsymbol{E}
\end{aligned} \tag{31}$$

b) : When $(i,j) = (j_1,i_1) \in \mathcal{E}$, we have:

$$\begin{aligned}
&\boldsymbol{F}_{n_p+1+i,i_1} \\
=&\boldsymbol{0} - \sum_{j\in V_i^+}\sum_{j_1\in V_{i_1}^-}\mathbb{E}\{\delta_{ij}^{-2}\delta_{j_1i_1}^{-2}\text{trace}\,(g_{ij})\,\boldsymbol{y}_{j_1i_1}^\top\boldsymbol{R}_{j_1}^\top\} \\
=&\delta_{ij}^{-4}(\boldsymbol{x}_j-\boldsymbol{x}_i)^\top\boldsymbol{E}\boldsymbol{R}_i\mathbb{E}\{\boldsymbol{y}_{ij}\boldsymbol{y}_{j_1i_1}^\top\}\boldsymbol{R}_{j_1}^\top \\
=&\delta_{ij}^{-2}(\boldsymbol{x}_j-\boldsymbol{x}_i)^\top\boldsymbol{E}
\end{aligned} \tag{32}$$

So we have:

$$\boldsymbol{F}_{n_p+1+i,i_1} = \begin{cases}\sum_{j\in V_i^+}\delta_{ij}^{-2}(\boldsymbol{x}_i-\boldsymbol{x}_j)^\top\boldsymbol{E} & i=i_1 \\ \delta_{ii_1}^{-2}(\boldsymbol{x}_{i_1}-\boldsymbol{x}_i)^\top\boldsymbol{E} & (i,i_1)\in\mathcal{E} \\ \boldsymbol{0} & else\end{cases}. \tag{33}$$

This sub-FIM corresponding to the parameter space $SO(2)\times\mathbb{R}^2$ is denoted as: $\triangle_w$.

3) $\mathbb{R}^2\times SO(2)$ *sub-matrix:* Because the FIM is symmetrical, we can get the block of the $\mathbb{R}^2\times SO(2)$ sub-matrix:

$$\boldsymbol{F}_{i_1,n_p+1+i} = \boldsymbol{F}_{n_p+1+i,i_1}^\top \tag{34}$$

So this sub-FIM is $\triangle_w^\top$.

4) $SO(2)$ *sub-matrix:* Using the definition in (11), we can get the block of the $SO(2)\times SO(2)$ sub-matrix:

$$\begin{aligned}
&\boldsymbol{F}_{n_p+1+i,n_p+1+i_1} \\
=&\mathbb{E}\{<\text{grad}L(\boldsymbol{\theta}),\boldsymbol{E}_i^{\boldsymbol{R}}>_{\boldsymbol{R}}\cdot<\text{grad}L(\boldsymbol{\theta}),\boldsymbol{E}_{i_1}^{\boldsymbol{R}}>_{\boldsymbol{R}}^\top\} \\
=&\mathbb{E}\{\text{trace}(\text{grad}_i L(\boldsymbol{R})^\top\boldsymbol{R}_i\boldsymbol{E})\cdot\text{trace}(\text{grad}_{i_1}L(\boldsymbol{R})^\top\boldsymbol{R}_{i_1}\boldsymbol{E})^\top\}
\end{aligned} \tag{35}$$

Based on (25), (28), Lemma 1 and $\mathbb{E}\{g_{ij}\} = 0$, we have:

$$\begin{aligned}
&\boldsymbol{F}_{n_p+1+i,n_p+1+i_1} = \sum_{j\in V_i^+}\sum_{j_1\in V_{i_1}^+}\mathbb{E}\{\delta_{ij}^{-2}\delta_{i_1j_1}^{-2}g_{ij}g_{i_1j_1}^\top\}+ \\
&\mathbb{E}\{\sum_{j\in V_i}\kappa_{ij}\text{trace}(\text{skew}(\boldsymbol{Z}_{ij})\boldsymbol{E})\cdot \\
&\sum_{j_1\in V_{i_1}}\kappa_{i_1j_1}\text{trace}(\text{skew}(\boldsymbol{Z}_{i_1j_1})\boldsymbol{E})\}
\end{aligned} \tag{36}$$

Set $\boldsymbol{Z}_{ij} = \begin{bmatrix} c_{ij} & -s_{ij} \\ s_{ij} & c_{ij}\end{bmatrix}$, where $c_{ij} = \cos(\alpha_{ij})$ and $s_{ij} = \sin(\alpha_{ij})$, $\alpha_{ij}\in so(2)$ [35], we have:

$$\text{trace}(\text{skew}(\boldsymbol{Z}_{ij})\boldsymbol{E}) = -2\kappa_{ij}s_{ij} \tag{37}$$

Substitute (37) into (36), we have:

$$\begin{aligned}
&\boldsymbol{F}_{n_p+1+i,n_p+1+i_1} = \sum_{j\in V_i^+}\sum_{j_1\in V_{i_1}^+}\delta_{ij}^{-2}\delta_{i_1j_1}^{-2}\mathbb{E}\{g_{ij}g_{i_1j_1}^\top\} \\
&+\sum_{j\in V_i}\sum_{j\in V_{i_1}}4\kappa_{ij}\kappa_{i_1j_1}\mathbb{E}\{s_{ij}s_{i_1j_1}\}
\end{aligned} \tag{38}$$

Because when the minimal representation α_{ij} is small, it meets $\alpha_{ij} \approx s_{ij}$. In the 2D plane, the isotropic Langevin distribution is similar to Gaussian distribution, we can regard it as: $s_{ij} \sim \mathcal{N}(0, \frac{1}{2\kappa_{ij}})$ [3]. In fact, the accurate value of $\mathbb{E}\{s_{ij}s_{i_1j_1}\}$ is $\frac{I_1(2\kappa_{ij})}{2\kappa_{ij}I_0(2\kappa_{ij})}$ using the similar derivation process in Appendix B and Appendix C for 3D case. When κ_{ij} becomes large, we will have: $\frac{I_1(2\kappa_{ij})}{2\kappa_{ij}I_0(2\kappa_{ij})} \to \frac{1}{2\kappa_{ij}}$.

a) : When $i = i_1$, we have:

$$\begin{aligned}
&\boldsymbol{F}_{n_p+1+i,n_p+1+i_1} \\
&= \sum_{j\in V_i^+} \delta_{ij}^{-4}(\boldsymbol{x}_i - \boldsymbol{x}_j)^\top \mathbb{E}\{\boldsymbol{y}_{ij}\boldsymbol{y}_{ij}^\top\}(\boldsymbol{x}_i - \boldsymbol{x}_j) + \sum_{j\in V_i} 4\kappa_{ij}^2 \frac{1}{2\kappa_{ij}} \\
&= \sum_{j\in V_i} 2\kappa_{ij} + \sum_{j\in V_i^+} \delta_{ij}^{-2}\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2
\end{aligned} \tag{39}$$

b) : When $i \neq i_1$ and $(i, i_1) \in \mathcal{E}$ or $(i_1, i) \in \mathcal{E}$, use $s_{ii_1} = -s_{i_1i}$, we have:

$$\boldsymbol{F}_{n_p+1+i,n_p+1+i_1} = \begin{cases} -2\kappa_{ii_1}, & (i, i_1) \in \mathcal{E} \\ -2\kappa_{i_1i}, & (i_1, i) \in \mathcal{E} \end{cases} \tag{40}$$

Finally, we have:

$$\boldsymbol{F}_{n_p+1+i,n_p+1+i_1} = \begin{cases} \sum_{j\in V_i} 2\kappa_{ij} + \psi_i, & i = i_1 \\ -2\kappa_{ii_1}, & (i, i_1) \in \mathcal{E} \\ -2\kappa_{i_1i}, & (i_1, i) \in \mathcal{E} \\ 0, & else \end{cases} \tag{41}$$

$$\psi_i = \sum_{j\in V_i^+} \delta_{ij}^{-2}\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2$$

Except ψ_i, the rest sub-matrix is a weighted Laplacian matrix. This weighted Laplacian matrix is denoted as $\boldsymbol{L}_w^{SO(2)}$.

5) The whole FIM: Combine (19), (33), (34) and (41), we can get the complete FIM:

$$\boldsymbol{\mathcal{I}}_{2D} = \left[\begin{array}{c:c} \boldsymbol{L}_w^{\mathbb{R}^2} & \triangle_w^\top \\ \hdashline \triangle_w & \boldsymbol{L}_w^{SO(2)} + \mathrm{diag}\{\psi_1, \cdots, \psi_{n_p}\} \end{array}\right] \tag{42}$$

It is noted that, for the 2D pose-graph SLAM, if the first node $\boldsymbol{x}_0$ is anchored, the corresponding rows and columns (node $\boldsymbol{x}_0$) of the FIM need to be deleted. So the normal SLAM problem is related to the weighted reduced Laplacian matrix.

C. Discussion about the FIM for 2D pose-graph SLAM

In the previous 2D work [5], [6] and [7], based on the block-isotropic Gaussian noise and the Euler angle, we get a similar formulation of the FIM by the equation:

$$\boldsymbol{\mathcal{I}}_{2D} = \boldsymbol{J}^\top \boldsymbol{\Sigma} \boldsymbol{J} \tag{43}$$

where $\boldsymbol{J}$ is the Jacobian matrix of the factors of the pose-graph SLAM. In fact, there are still three important differences between this work with our previous work:

a) Different noise assumptions: The noise for the rotation part used in this work is different from the Gaussian noise. For the Gaussian noise, we have following result: If $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\delta})$, we can get $\boldsymbol{G}\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{G}\boldsymbol{\mu}, \boldsymbol{G}\boldsymbol{\delta}\boldsymbol{G}^\top)$. So the FIM can be computed by (43). This equation may not be suitable for other kinds of noise, so we need to compute the FIM based on its general definition shown in (11).

b) Lie group representation instead of Euler angle: The previous work is built based on the Euler angle representation for the orientation of the poses. The measurement of the rotation part is simply written as the subtraction of the orientation. Because of the non-uniqueness of the Euler angle and the periodicity of the trigonometric functions, the wraparound problem will lead the measurement function to be complex instead of a simple subtraction. In this work, the formulation is built based on the rotation group, so there is no wraparound issues.

c) Expandability of conclusions from 2D to 3D: Based on the Gaussian noise on the Euler angle, for the 3D case, the relative pose measurement can not be written as a simple subtraction formulation, which is greatly different from the 2D case, so the expandability of the original method is limited in 2D case. Based on the rigorous treatment of uncertainty on $\mathbb{R}^n \times SO(n)$ (isotropic Langevin noise for rotation and block-isotropic Gaussian noise for translation), the similar but more complex conclusions are presented rigorously in the 2D case and further extended into the 3D case in the following section.

V. FIM for 3D pose-graph SLAM

A. Geometry of the parameter spaces

As shown in Section III-C, the tools, including the tangent space, the inner product and the gradient, are also suitable for the 3D ($\mathbb{R}^3 \times SO(3)$) parameter space. For the orthonormal basis, we will show its difference with (12) because of the higher dimensions.

1) Orthonormal basis for $\mathbb{R}^3 \times SO(3)$: The orthonormal basis $\boldsymbol{E} = (\boldsymbol{E}_0^{\boldsymbol{x}}, \cdots, \boldsymbol{E}_{n_p}^{\boldsymbol{x}}, \boldsymbol{E}_0^{\boldsymbol{R}}, \cdots, \boldsymbol{E}_{n_p}^{\boldsymbol{R}})$ of the tangent space $\mathbb{R}^3 \times \mathcal{T}_{\boldsymbol{R}}SO(3)$ is:

$$\begin{aligned}
\boldsymbol{E}_i^{\boldsymbol{x}} &= ({\boldsymbol{E}_i^X}^\top, \boldsymbol{0}_{3\times 9(n_p+1)}{}^\top)^\top, \ i \in [n_p]\bigcup\{0\} \\
\boldsymbol{E}_{j,k}^{\boldsymbol{R}} &= (\boldsymbol{0}_{3\times 3(n_p+1)}, \boldsymbol{E}_{j,k}^R), \ j \in [n_p]\bigcup\{0\}, k \in [3] \\
\boldsymbol{E}_i^X &= (\boldsymbol{0}, \cdots, \boldsymbol{0}, \underbrace{\boldsymbol{I}_{3\times 3}}_{i-\text{th}}, \boldsymbol{0}, \cdots, \boldsymbol{0})_{3\times 3(n_p+1)} \\
\boldsymbol{E}_{j,k}^R &= (\boldsymbol{0}, \cdots, \boldsymbol{0}, \underbrace{\boldsymbol{R}_j\boldsymbol{E}_k}_{3j+k-\text{th}}, \boldsymbol{0}, \cdots, \boldsymbol{0})_{3\times 3(n_p+1)} \\
\boldsymbol{E}_1 &= \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \boldsymbol{E}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \\
\boldsymbol{E}_3 &= \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}
\end{aligned} \tag{44}$$

It is noted that, in many applications, they use $\sqrt{2}\boldsymbol{E}_1$, $\sqrt{2}\boldsymbol{E}_2$, $\sqrt{2}\boldsymbol{E}_3$ as the bases of the manifold $SO(3)$. By this way, the weight for rotation group part in the finial FIM will be $2\omega_{ij}$, where ω_{ij} will be shown in (59).

B. FIM

Similar to the FIM in Section III-D, there are four parts in the FIM for the group $\mathbb{R}^3 \times SO(3)$ of the rigid body motions in 3D space. Their derivation processes are shown as follows:

1) $\mathbb{R}^3$ sub-matrix: The derivation of this sub-matrix is similar to Section IV-B1. Based on (19), we can directly obtain the sub-FIM corresponding to Euclidean space $\mathbb{R}^3$:

$$\boldsymbol{F}_{i,i_1} = \begin{cases} \sum_{j_1 \in V_i^+} \sigma_{ij_1}^{-2} \boldsymbol{I}_{3\times 3} + \sum_{j_2 \in V_i^-} \sigma_{j_2 i}^{-2} \boldsymbol{I}_{3\times 3}, & i = i_1 \\ -\sigma_{ii_1}^{-2} \boldsymbol{I}_{3\times 3}, & (i,i_1) \in \mathcal{E} \\ -\sigma_{i_1 i}^{-2} \boldsymbol{I}_{3\times 3}, & (i_1,i) \in \mathcal{E} \\ \boldsymbol{0}, & else \end{cases} \tag{45}$$

We can also find that it is a weighted Laplacian matrix after using the Kronecker product operation. We denote it as $\boldsymbol{L}_w^{\mathbb{R}^3}$.

2) $SO(3)\times\mathbb{R}^3$ sub-matrix: Based on the geometry of the parameter space and the definition in (11), we can get the block of the $SO(3)\times\mathbb{R}^3$ sub-matrices:

$$\begin{aligned} &\boldsymbol{F}_{n_p+1+i,i_1}(k) \\ =&\mathbb{E}\{< \text{grad}L(\boldsymbol{\theta}), \boldsymbol{E}_{i,k}^{\boldsymbol{R}} >_{\boldsymbol{R}} \cdot < \text{grad}L(\boldsymbol{\theta}), \boldsymbol{E}_{i_1}^{\boldsymbol{X}} >_{\boldsymbol{X}}^{\top}\} \\ =&\mathbb{E}\{\text{trace}(\text{grad}_i L(\boldsymbol{R})^{\top} \cdot \boldsymbol{R}_i \boldsymbol{E}_k) \cdot \boldsymbol{I}_{3\times 3}^{\top} \cdot \text{grad}_{i_1} L(\boldsymbol{x})\} \end{aligned} \tag{46}$$

where $\boldsymbol{F}_{n_p+1+i,i_1}(k) \in \mathbb{R}_{3\times 1}$ means the k-th row of the n_p+1+i,i_1-th block of the FIM.

We have:

$$\begin{aligned} &\text{grad}_i L(\boldsymbol{R}) \\ =&\text{grad}_i \left(\frac{1}{2} \sum_{(i,j)\in\mathcal{F}} \log \widehat{f}_{ij}(\boldsymbol{Z}_{ij}) + \sum_{j\in V_i^+} \log f_{ij}(\boldsymbol{y}_{ij}) \right) \end{aligned} \tag{47}$$

Based on (25), we can obtain the first part $\text{grad}_i \frac{1}{2} \sum_{(i,j)\in\mathcal{F}} \log \widehat{f}_{ij}(\boldsymbol{Z}_{ij})$ is:

$$\text{grad}_i \frac{1}{2} \sum_{(i,j)\in\mathcal{F}} \log \widehat{f}_{ij}(\boldsymbol{Z}_{ij}) = \sum_{j\in V_i} -\kappa_{ij} \boldsymbol{R}_i \text{skew}(\boldsymbol{Z}_{ij}) \tag{48}$$

Let's consider the second part, based on the basis $\boldsymbol{R}_i\boldsymbol{E}$ [35]:

$$\begin{aligned} &\sum_{j\in V_i^+} \text{grad}_i \log f_{ij}(\boldsymbol{y}_{ij}) \\ =& \sum_{j\in V_i^+} \sum_{k=1}^{3} \boldsymbol{R}_i \boldsymbol{E}_k \nabla_{\boldsymbol{R}_i\boldsymbol{E}_k} \log f_{ij}(\boldsymbol{y}_{ij}) \\ =& \sum_{j\in V_i^+} \delta_{ij}^{-2} \sum_{k=1}^{3} \boldsymbol{R}_i \boldsymbol{E}_k (\boldsymbol{y}_{ij}^{\top} \boldsymbol{E}_k^{\top} \boldsymbol{R}_i^{\top} (\boldsymbol{x}_i - \boldsymbol{x}_j)) \end{aligned} \tag{49}$$

For $\text{grad}_{i_1} L(\boldsymbol{x})$, based on (29), we have:

$$\begin{aligned} &\text{grad}_{i_1} L(\boldsymbol{x}) \\ =& \sum_{j_1\in V_{i_1}^+} \delta_{i_1 j_1}^{-2} \boldsymbol{y}_{i_1 j_1}^{\top} \boldsymbol{R}_{i_1}^{\top} - \sum_{j_1\in V_{i_1}^-} \delta_{j_1 i_1}^{-2} \boldsymbol{y}_{j_1 i_1}^{\top} \boldsymbol{R}_{j_1}^{\top} \end{aligned} \tag{50}$$

Substitute (48), (49) and (50) into (46), let $\boldsymbol{P}_{ij} = \left(\sum_{k=1}^{3} \boldsymbol{R}_i \boldsymbol{E}_k \boldsymbol{y}_{ij}^{\top} \boldsymbol{E}_k^{\top} \boldsymbol{R}_i^{\top} (\boldsymbol{x}_i - \boldsymbol{x}_j) \right)^{\top} \boldsymbol{R}_i \boldsymbol{E}_k \in \mathbb{R}_{3\times 3}$, based on Lemma 1 and $\mathbb{E}\{\boldsymbol{Z}_{ij}\} = \boldsymbol{0}$, we have:

$$\begin{aligned} &\boldsymbol{F}_{n_p+1+i,i_1}(k) \\ =&\mathbb{E}\{\text{trace}(\text{grad}_i L(\boldsymbol{R})^{\top} \cdot \boldsymbol{R}_i \boldsymbol{E}_k) \cdot \boldsymbol{I}_{2\times 2}^{\top} \cdot \text{grad}_{i_1} L(\boldsymbol{x})\} \\ =&\boldsymbol{0} + \sum_{j\in V_i^+} \sum_{j_1\in V_{i_1}^+} \mathbb{E}\{\delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \text{trace}\,(\boldsymbol{P}_{ij})\, \boldsymbol{y}_{i_1 j_1}^{\top} \boldsymbol{R}_{i_1}^{\top}\} \\ &- \sum_{j\in V_i^+} \sum_{j_1\in V_{i_1}^-} \mathbb{E}\{\delta_{ij}^{-2} \delta_{j_1 i_1}^{-2} \text{trace}\,(\boldsymbol{P}_{ij})\, \boldsymbol{y}_{j_1 i_1}^{\top} \boldsymbol{R}_{j_1}^{\top}\} \end{aligned} \tag{51}$$

a) : When $(i,j) = (i_1,j_1) \in \mathcal{E}$, we have:

$$\begin{aligned} &\boldsymbol{F}_{n_p+1+i,i_1}(k) \\ =& \sum_{j\in V_i^+} \sum_{j_1\in V_{i_1}^+} \mathbb{E}\{\delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \text{trace}\,(\boldsymbol{P}_{ij})\, \boldsymbol{y}_{i_1 j_1}^{\top} \boldsymbol{R}_{i_1}^{\top}\} + \boldsymbol{0} \\ =& \sum_{j\in V_i^+} \delta_{ij}^{-4} (\boldsymbol{x}_i - \boldsymbol{x}_j)^{\top} \boldsymbol{R}_i \boldsymbol{E}_k \mathbb{E}\{\boldsymbol{y}_{ij} \boldsymbol{y}_{ij}^{\top}\} \boldsymbol{R}_i^{\top} \\ =& \sum_{j\in V_i^+} \delta_{ij}^{-2} (\boldsymbol{x}_i - \boldsymbol{x}_j)^{\top} \boldsymbol{R}_i \boldsymbol{E}_k \boldsymbol{R}_i^{\top} \end{aligned} \tag{52}$$

b) : When $(i,j) = (j_1,i_1) \in \mathcal{E}$, we have:

$$\begin{aligned} &\boldsymbol{F}_{n_p+1+i,i_1}(k) \\ =&\boldsymbol{0} - \sum_{j\in V_i^+} \sum_{j_1\in V_{i_1}^-} \mathbb{E}\{\delta_{ij}^{-2} \delta_{j_1 i_1}^{-2} \text{trace}\,(\boldsymbol{P}_{ij})\, \boldsymbol{y}_{j_1 i_1}^{\top} \boldsymbol{R}_{j_1}^{\top}\} \\ =&\delta_{ij}^{-4} (\boldsymbol{x}_j - \boldsymbol{x}_i)^{\top} \boldsymbol{R}_i \boldsymbol{E}_k \mathbb{E}\{\boldsymbol{y}_{ij} \boldsymbol{y}_{j_1 i_1}^{\top}\} \boldsymbol{R}_{j_1}^{\top} \\ =&\delta_{ij}^{-2} (\boldsymbol{x}_j - \boldsymbol{x}_i)^{\top} \boldsymbol{R}_i \boldsymbol{E}_k \boldsymbol{R}_i^{\top} \end{aligned} \tag{53}$$

So we have:

$$\boldsymbol{F}_{n_p+1+i,i_1} = \begin{cases} \begin{bmatrix} \sum_{j\in V_i^+} \delta_{ij}^{-2} (\boldsymbol{x}_j - \boldsymbol{x}_i)^{\top} \boldsymbol{R}_i \boldsymbol{E}_1 \boldsymbol{R}_i^{\top} \\ \sum_{j\in V_i^+} \delta_{ij}^{-2} (\boldsymbol{x}_j - \boldsymbol{x}_i)^{\top} \boldsymbol{R}_i \boldsymbol{E}_2 \boldsymbol{R}_i^{\top} \\ \sum_{j\in V_i^+} \delta_{ij}^{-2} (\boldsymbol{x}_j - \boldsymbol{x}_i)^{\top} \boldsymbol{R}_i \boldsymbol{E}_3 \boldsymbol{R}_i^{\top} \end{bmatrix} & i = i_1 \\ \begin{bmatrix} \delta_{ii_1}^{-2} (\boldsymbol{x}_{i_1} - \boldsymbol{x}_i)^{\top} \boldsymbol{R}_i \boldsymbol{E}_1 \boldsymbol{R}_i^{\top} \\ \delta_{ii_1}^{-2} (\boldsymbol{x}_{i_1} - \boldsymbol{x}_i)^{\top} \boldsymbol{R}_i \boldsymbol{E}_2 \boldsymbol{R}_i^{\top} \\ \delta_{ii_1}^{-2} (\boldsymbol{x}_{i_1} - \boldsymbol{x}_i)^{\top} \boldsymbol{R}_i \boldsymbol{E}_3 \boldsymbol{R}_i^{\top} \end{bmatrix} & (i,i_1) \in \mathcal{E} \\ \boldsymbol{0} & else \end{cases}. \tag{54}$$

This sub-FIM corresponding to the parameter space $SO(3)\times\mathbb{R}^3$ is denoted as $\triangle_w^{3D}$.

3) $\mathbb{R}^3\times SO(3)$ sub-matrix: Because the FIM is symmetrical, we can get the block of the $\mathbb{R}^3\times SO(3)$ sub-matrix:

$$\boldsymbol{F}_{i_1,n_p+1+i} = \boldsymbol{F}_{n_p+1+i,i_1}^{\top} \tag{55}$$

So this sub-FIM is $\triangle_w^{3D\top}$.

4) $SO(3)$ sub-matrix: Using the definition in (11), we can get the block of the $SO(3)\times SO(3)$ sub-matrix:

$$\begin{aligned} &\boldsymbol{F}_{n_p+1+i,n_p+1+i_1}(k,l) \\ =&\mathbb{E}\{< \text{grad}L(\boldsymbol{\theta}), \boldsymbol{E}_{i,k}^{\boldsymbol{R}} >_{\boldsymbol{R}} \cdot < \text{grad}L(\boldsymbol{\theta}), \boldsymbol{E}_{i_1,l}^{\boldsymbol{R}} >_{\boldsymbol{R}}^{\top}\} \\ =&\mathbb{E}\{\text{trace}(\text{grad}_i L(\boldsymbol{R})^{\top} \boldsymbol{R}_i \boldsymbol{E}_k) \cdot \text{trace}(\text{grad}_{i_1} L(\boldsymbol{R})^{\top} \boldsymbol{R}_{i_1} \boldsymbol{E}_l)^{\top}\} \end{aligned} \tag{56}$$

where $\boldsymbol{F}_{n_p+1+i,n_p+1+i_1}(k,l) \in \mathbb{R}$ means the (k,l)-th element of the (n_p+1+i, n_p+1+i_1)-th block of the FIM.

Based on (48) and (49), Lemma 1 and $\mathbb{E}\{g_{ij}\} = 0$, let $\widehat{g}_{ij,k} = (\boldsymbol{x}_i - \boldsymbol{x}_j)^\top \boldsymbol{R}_i \boldsymbol{E}_k \boldsymbol{y}_{ij}$, we have:

$$\begin{aligned}\boldsymbol{F}_{n_p+1+i,n_p+1+i_1}(k,l) = &\sum_{j\in V_i^+}\sum_{j_1\in V_{i_1}^+}\mathbb{E}\{\delta_{ij}^{-2}\delta_{i_1j_1}^{-2}\widehat{g}_{ij,k}\widehat{g}_{i_1j_1,l}^\top\}+\\ &\mathbb{E}\{\sum_{j\in V_i}\kappa_{ij}\text{trace}(\text{skew}(\boldsymbol{Z}_{ij})^\top\boldsymbol{E}_k)\cdot\\ &\sum_{j_1\in V_{i_1}}\kappa_{i_1j_1}\text{trace}(\text{skew}(\boldsymbol{Z}_{i_1j_1})^\top\boldsymbol{E}_l)\}\end{aligned} \tag{57}$$

Lemma 2. *If* $\boldsymbol{Z}_{ij} = \exp(\boldsymbol{\Phi}_{ij}^\wedge) \sim Lang(\boldsymbol{I}_{n\times n}, \kappa_{ij}), \boldsymbol{\Phi}_{ij} = [\Phi_1^{ij}, \Phi_2^{ij}, \Phi_3^{ij}]^\top$*, we have:* $\mathbb{E}\{{\Phi_k^{ij}}^2\} = \mathbb{E}\{{\Phi_l^{ij}}^2\} = \frac{1}{2}\mathbb{E}\{< skew(Z_{ij}), \boldsymbol{E}_k >_{\boldsymbol{R}}^2\} = \frac{\mathbb{E}\{\|skew(\boldsymbol{Z}_{ij})\|_F^2\}}{6}$.

Proof: See Appendix B.

a) : When $(i,j) = (i_1,j_1)$ and $k = l$, based on (57), we have:

$$\begin{aligned}\boldsymbol{F}_{n_p+1+i,n_p+1+i_1}(k,k) &= \sum_{j\in V_i}2\kappa_{ij}^2\mathbb{E}\{{\Phi_k^{ij}}^2\}\boldsymbol{I}_{3\times3}\\ &+\sum_{j\in V_i^+}\delta_{ij}^{-4}(\boldsymbol{x}_i-\boldsymbol{x}_j)^\top\boldsymbol{R}_i\boldsymbol{E}_k\mathbb{E}\{\boldsymbol{y}_{ij}\boldsymbol{y}_{ij}^\top\}\boldsymbol{E}_k^\top\boldsymbol{R}_i^\top(\boldsymbol{x}_i-\boldsymbol{x}_j)\\ &=\sum_{j\in V_i}\frac{\omega_{ij}}{3}\boldsymbol{I}_{3\times3}+\sum_{j\in V_i^+}\delta_{ij}^{-2}(\boldsymbol{x}_i-\boldsymbol{x}_j)^\top\boldsymbol{R}_i\boldsymbol{I}_{3\times3}^{k,k}\boldsymbol{R}_i^\top(\boldsymbol{x}_i-\boldsymbol{x}_j)\end{aligned} \tag{58}$$

where $\boldsymbol{I}_{3\times3}^{k,k} = \boldsymbol{E}_k\boldsymbol{E}_k^\top$,

$$\omega_{ij} = \frac{\kappa_{ij}^2\left(2I_0(2\kappa_{ij}) - I_1(2\kappa_{ij}) - 2I_2(2\kappa_{ij}) + I_3(2\kappa_{ij})\right)}{2I_0(2\kappa_{ij}) - 2I_1(2\kappa_{ij})} \tag{59}$$

where the specific derivation of this value is shown in Appendix C.

b) : When $(i,j) = (i_1,j_1)$ and $k \neq l$, based on (57), $\mathbb{E}\{\Phi_k^{ij}\} = \mathbb{E}\{\Phi_k^{ij}\}\mathbb{E}\{\Phi_l^{ij}\} = 0$, we have:

$$\begin{aligned}\boldsymbol{F}_{n_p+1+i,n_p+1+i_1}(k,l) &= \boldsymbol{0}\\ &+\sum_{j\in V_i^+}\delta_{ij}^{-4}(\boldsymbol{x}_i-\boldsymbol{x}_j)^\top\boldsymbol{R}_i\boldsymbol{E}_k\mathbb{E}\{\boldsymbol{y}_{ij}\boldsymbol{y}_{ij}^\top\}\boldsymbol{E}_l^\top\boldsymbol{R}_i^\top(\boldsymbol{x}_i-\boldsymbol{x}_j)\\ &=\sum_{j\in V_i^+}\delta_{ij}^{-2}(\boldsymbol{x}_i-\boldsymbol{x}_j)^\top\boldsymbol{R}_i\boldsymbol{I}_{3\times3}^{k,l}\boldsymbol{R}_i^\top(\boldsymbol{x}_i-\boldsymbol{x}_j)\end{aligned} \tag{60}$$

where $\boldsymbol{I}_{3\times3}^{k,l} = \boldsymbol{E}_k\boldsymbol{E}_l^\top$.

c) : When $(i,j) = (j_1,i_1) \in \mathcal{E}$ and $k = l$, based on (57), we have:

$$\begin{aligned}\boldsymbol{F}_{n_p+1+i,n_p+1+i_1}(k,l) &= \mathbb{E}\{\kappa_{ii_1}\text{trace}(\text{skew}(\boldsymbol{Z}_{ij})^\top\boldsymbol{E}_k)\cdot\\ &\left(-\kappa_{ii_1}\text{trace}(\text{skew}(\boldsymbol{Z}_{j_1i_1})^\top\boldsymbol{E}_k)\right)\}+\boldsymbol{0}\\ &=-2\kappa_{ij}^2\mathbb{E}\{{\Phi_k^{ij}}^2\}\boldsymbol{I}_{3\times3}\\ &=-\frac{\omega_{ii_1}}{3}\boldsymbol{I}_{3\times3}\end{aligned} \tag{61}$$

d) : When $(i_1,j_1) = (j,i) \in \mathcal{E}$ and $k = l$, based on (57), we have:

$$\begin{aligned}\boldsymbol{F}_{n_p+1+i,n_p+1+i_1}(k,l) &= \mathbb{E}\{\left(-\kappa_{i_1i}\text{trace}(\text{skew}(\boldsymbol{Z}_{ji})^\top\boldsymbol{E}_k)\right)\cdot\\ &\kappa_{i_1i}\text{trace}(\text{skew}(\boldsymbol{Z}_{i_1j_1})^\top\boldsymbol{E}_k)\}+\boldsymbol{0}\\ &=-2\kappa_{ij}^2\mathbb{E}\{{\Phi_k^{ij}}^2\}\boldsymbol{I}_{3\times3}\\ &=-\frac{\omega_{i_1i}}{3}\boldsymbol{I}_{3\times3}\end{aligned} \tag{62}$$

Finally, we have:

$$\begin{aligned}\boldsymbol{F}_{n_p+1+i,n_p+1+i_1} &= \begin{cases}\sum_{j\in V_i}\frac{\omega_{ij}}{3}\boldsymbol{I}_{3\times3}+\boldsymbol{\Psi}_i, & i=i_1\\ -\frac{\omega_{ii_1}}{3}\boldsymbol{I}_{3\times3}, & (i,i_1)\in\mathcal{E}\\ -\frac{\omega_{i_1i}}{3}\boldsymbol{I}_{3\times3}, & (i_1,i)\in\mathcal{E}\\ \boldsymbol{0}, & else\end{cases}\\ \boldsymbol{\Psi}_i &= \begin{bmatrix}\psi_i^{11} & \psi_i^{12} & \psi_i^{13}\\ \psi_i^{12} & \psi_i^{22} & \psi_i^{21}\\ \psi_i^{13} & \psi_i^{21} & \psi_i^{33}\end{bmatrix}\\ \psi_i^{kl} &= \sum_{j\in V_i^+}\delta_{ij}^{-2}(\boldsymbol{x}_i-\boldsymbol{x}_j)^\top\boldsymbol{R}_i\boldsymbol{I}_{3\times3}^{k,l}\boldsymbol{R}_i^\top(\boldsymbol{x}_i-\boldsymbol{x}_j)\end{aligned} \tag{63}$$

Except $\boldsymbol{\Psi}_i$, it is easy to find that the rest sub-matrix is a weighted Laplacian matrix after using the Kronecker product operation. This weighted Laplacian matrix can be written as $\boldsymbol{L}_w^{SO(3)}$.

5) The whole FIM: Combine (45), (54), (55) and (63), we can get the complete FIM:

$$\boldsymbol{\mathcal{I}}_{3D} = \left[\begin{array}{c|c}\boldsymbol{L}_w^{\mathbb{R}^3} & \triangle_w^{3D\top}\\ \hline \triangle_w^{3D} & \boldsymbol{L}_w^{SO(3)} + \text{diag}\{\boldsymbol{\Psi}_1,\cdots,\boldsymbol{\Psi}_{n_p}\}\end{array}\right] \tag{64}$$

Similar to the 2D pose-graph SLAM, for the 3D pose-graph SLAM, if the first node $\boldsymbol{x}_0$ is anchored, the corresponding rows and columns (node $\boldsymbol{x}_0$) of the FIM need to be deleted. So it is also related to the weighted reduced Laplacian matrix. The following CRLB and discussions on optimization metrics are focused on the pose-graph SLAM, so the FIM is defined based on the anchored situation.

VI. CRLB FOR POSE-GRAPH SLAM

Classical CRLB give a lower bound on the covariance matrix $\boldsymbol{C}$ of any unbiased estimator for an estimation problem in $\mathbb{R}^n$. In terms of the FIM $\boldsymbol{F} = \boldsymbol{\mathcal{I}}_{nD}$ of that problem, the classical result reads $\boldsymbol{C} \succeq \boldsymbol{F}^{-1}$. However, because our parameter space $\mathcal{P}$ is a manifold instead of a flat Euclidean space, the CRLB takes up the more general form $\boldsymbol{C} \succeq \boldsymbol{F}^{-1}$ + curvature terms.

Inspired by [23], we also show that when the signal-to-noise ratio (SNR) is large enough, the curvature terms will become negligible. For SLAM, the SNR snr is defined as:

$$snr = \frac{(n_p - |\mathcal{A}|)\mathbb{E}\{dist^2(\boldsymbol{Z}_{uni}^{\boldsymbol{R}}, \boldsymbol{I}_{n\times n}) + \|\boldsymbol{Z}_{uni}^{\boldsymbol{X}} - \boldsymbol{0}\|^2\}}{d^2\text{trace}(\boldsymbol{I}^{SO(3)^{-1}})} \tag{65}$$

where $\boldsymbol{I}^{SO(3)} = \boldsymbol{L}_w^{SO(3)} + \text{diag}\{\boldsymbol{\Psi}_1,\cdots,\boldsymbol{\Psi}_{n_p}\}$, $|\mathcal{A}|$ means the number of anchored node (in general $|\mathcal{A}| = 1$ for SLAM), $\boldsymbol{Z}_{uni}^{\boldsymbol{R}}$ and $\boldsymbol{Z}_{uni}^{\boldsymbol{X}}$ are respectively the expectations with a uniformly distributed over $SO(3)$ and $\mathbb{R}^3$, it is easy to find

that the SNR is inversely proportional to the trace function of the inverse of the information matrix. In other words, when the uncertainty of the rotation group is small (large-scale $\boldsymbol{I}^{SO(3)}$), the SNR snr will be large.

A. CRLB for 2D pose-graph SLAM

Before the discussion about the CRLB of the synchronization of the manifolds $\mathbb{R}^2 \times SO(2)$, we first introduce a lemma [36]:

Lemma 3. *Let $\mathcal{M} = \mathcal{M}_1 \times \mathcal{M}_2$ be the product of two Riemannian manifolds, and R be its curvature tensor, R_1, R_2 be curvature tensors for $\mathcal{M}_1$ and $\mathcal{M}_2$ respectively, then one can relate R, R_1 and R_2 by:*

$$R(X,Y) = R_1(X_1,Y_1) + R_2(X_2,Y_2) \tag{66}$$

where $X_i, Y_i \in \mathcal{T}(\mathcal{M}_i), i = 1,2$ and $X = X_1 + X_2, Y = Y_1 + Y_2 \in \mathcal{T}(\mathcal{M})$, $\mathcal{T}(\star)$ means the Tangent space.

Based on Lemma 3, we can get the curvature tensor of the parameter space $\mathcal{P}_1 = \{\mathbb{R}^2 \times \cdots \times \mathbb{R}^2\}_{n_p} \times \{SO(2) \times \cdots \times SO(2)\}_{n_p}$ is equal to the sum of the multiple curvature tensors of the group $\{\mathbb{R}^2 \times \cdots \times \mathbb{R}^2\}_{n_p}$ and the group $\{SO(2) \times \cdots \times SO(2)\}_{n_p}$. However, these two groups are both flat and their curvature tensors are $\mathbf{0}$. So we can get that the curvature tensor of the parameter space $\mathcal{P}_1$ is $\mathbf{0}$.

Following, the form of the CRLB for the 2D pose-graph SLAM on $\mathbb{R}^2 \times SO(2)$ is:

$$\boldsymbol{C} \succeq \boldsymbol{F}^{-1} + \text{curvature terms} = \boldsymbol{F}^{-1} + \mathbf{0} = \boldsymbol{F}^{-1} \tag{67}$$

So we can see that, for the parameter space $\mathcal{P}_1$, its CRLB form is the same as the classical CRLB result for the flat Euclidean space.

B. CRLB for 3D pose-graph SLAM

Based on the Theorem 4 in [25], CRLB follows: $\boldsymbol{C} \geq \boldsymbol{F}^{-1} - \frac{1}{3}(R_m(\boldsymbol{F}^{-1})\boldsymbol{F}^{-1} + \boldsymbol{F}^{-1}R_m(\boldsymbol{F}^{-1}))$, where the operator R_m: $\mathbb{R}_{6n_p \times 6n_p} \rightarrow \mathbb{R}_{6n_p \times 6n_p}$ involves the Riemannian curvature tensor of the parameter space. The operator $R_m(\boldsymbol{F}^{-1}) = \frac{1}{4}ddiag(\boldsymbol{I}^{SO(3)^{-1}})$ of the parameter space $\mathcal{P}_2 = \{SO(3) \times \cdots \times SO(3)\}$ for the anchored rotation synchronization situation has been shown by [23], ddiag sets all off-diagonal blocks of a matrix to zero.

Different from the $\mathbb{R}^2 \times SO(2)$ group, the manifolds $\mathbb{R}^3 \times SO(3)$ is not a flat space. So we need to compute the curvature terms. Based on Lemma 2, we known that the curvature tensor of the parameter space $\mathcal{P}_3 = \{\mathbb{R}^3 \times \cdots \times \mathbb{R}^3\}_{n_p} \times \{SO(3) \times \cdots \times SO(3)\}_{n_p}$ is equal to the sum of the curvature tensor of the manifold $\{\mathbb{R}^3 \times \cdots \times \mathbb{R}^3\}_{n_p}$ and the manifold $\mathcal{P}_2 = \{SO(3) \times \cdots \times SO(3)\}_{n_p}$. The Euclidean space $\mathbb{R}^3$ is flat with a $\mathbf{0}$ curvature tensor, so the curvature tensor of the parameter space $\mathcal{P}_3$ is determined by the curvature tensor of the manifold $\mathcal{P}_2 = \{SO(3) \times \cdots \times SO(3)\}_{n_p}$.

Following the same formulation, we can get the curvature tensor of the 3D pose-graph SLAM on the group $\mathbb{R}^3 \times SO(3)$ as:

$$\begin{aligned} \boldsymbol{C} &\succeq \boldsymbol{F}^{-1} - \frac{1}{12}(\text{ddiag}(\widetilde{\boldsymbol{L}})\boldsymbol{F}^{-1} + \boldsymbol{F}^{-1}\text{ddiag}(\widetilde{\boldsymbol{L}})) \\ \widetilde{\boldsymbol{L}} &= \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}^{SO(3)^{-1}} \end{bmatrix}_{6n_p \times 6n_p} \end{aligned} \tag{68}$$

It is easy to find that when snr is large enough, compared with $\boldsymbol{F}^{-1}$ the function $\text{ddiag}(\widetilde{\boldsymbol{L}})\boldsymbol{F}^{-1} + \boldsymbol{F}^{-1}\text{ddiag}(\widetilde{\boldsymbol{L}})$ is on a much smaller order of magnitude $O(\boldsymbol{F}^{-2})$. So it is negligible.

VII. Optimal experimental design metrics for pose-graph SLAM

It is known that the TOED [29], including T-optimality and D-optimality, for the FIM of the estimated state vector is a common way to evaluate the actions in terms of estimation accuracy. In this section, we will discuss and compare the optimal experimental design metrics of the 2D/3D pose-graph SLAM and show the bounds of these indexes, which are much easier to be computed than the accurate index for the FIM.

A. T-optimality design metric

1) T-optimality design metric for the synchronization on $\mathbb{R}^2 \times SO(2)$: Based on (42), we can get the T-optimality design metric of the FIM:

$$\begin{aligned} &\text{trace}(\boldsymbol{\mathcal{I}}_{2D}) \\ =&\text{trace}(\boldsymbol{L}_w^{\mathbb{R}^2}) + \text{trace}(\boldsymbol{L}_w^{SO(2)}) + \sum_{i=1}^{n_p}\sum_{j \in V_i^+} \delta_{ij}^{-2}\|\boldsymbol{x}_j - \boldsymbol{x}_i\|^2 \end{aligned} \tag{69}$$

2) T-optimality design metric for the synchronization on $\mathbb{R}^3 \times SO(3)$: Based on (64), we can get the T-optimality design metric of the FIM:

$$\begin{aligned} &\text{trace}(\boldsymbol{\mathcal{I}}_{3D}) \\ =&\text{trace}(\boldsymbol{L}_w^{\mathbb{R}^3}) + \text{trace}(\boldsymbol{L}_w^{SO(3)}) + \sum_{i=1}^{n_p}\text{trace}(\boldsymbol{\Psi}_i) \end{aligned} \tag{70}$$

We can find that, in these three parts, only one part $\sum_{i=1}^{n_p}\text{trace}(\boldsymbol{\Psi}_i)$ includes the state vector $\boldsymbol{x}_i$, $\boldsymbol{x}_j$, $\boldsymbol{R}_i$. In fact, because of the special structure about $\boldsymbol{\Psi}_i$, we have: $\sum_{i=1}^{n_p}\text{trace}(\boldsymbol{\Psi}_i) = \sum_{i=1}^{n_p}\sum_{j \in V_i^+}\text{trace}(\boldsymbol{\Psi}_{(i,j)}) = \sum_{i=1}^{n_p}\sum_{j \in V_i^+}\delta_{ij}^{-2}\|\boldsymbol{x}_j - \boldsymbol{x}_i\|^2$. The proof is shown in Appendix D.

3) Further analysis about the T-optimality design metric: We know that the measurement function is: $\boldsymbol{p}_{ij} = \boldsymbol{R}_i^\top(\boldsymbol{x}_j - \boldsymbol{x}_i) + \boldsymbol{y}_{ij}$. If the solution is the "optimal" or "near-optimal" solution with the small noise, we have:

$$\delta_{ij}^{-2}\|\boldsymbol{x}_j - \boldsymbol{x}_i\|^2 = \delta_{ij}^{-2}\|\boldsymbol{p}_{ij} - \boldsymbol{y}_{ij}\|^2 \approx \delta_{ij}^{-2}\|\boldsymbol{p}_{ij}\|^2 \tag{71}$$

Introduce (71) into (69) and (71), we have:

$$\begin{aligned} \text{trace}(\boldsymbol{\mathcal{I}}_{nD}) \approx &\text{trace}(\boldsymbol{L}_w^{\mathbb{R}^n}) + \text{trace}(\boldsymbol{L}_w^{SO(n)}) \\ &+ \sum_{i=1}^{n_p}\sum_{j \in V_i^+}\delta_{ij}^{-2}\|\boldsymbol{p}_{ij}\|^2,\ n = 2,3 \end{aligned} \tag{72}$$

Based on (72), we can see that the trace function of the FIM is weakly related to the state vector obtained by the estimated

result. So, in general, the T-optimality can be easily computed using above equation without considering the SLAM result. For many real word datasets, compared with the other parts, $\sum_{i=1}^{n_p}\sum_{j\in V_i^+}\delta_{ij}^{-2}\|\boldsymbol{p}_{ij}\|^2$ is relatively small in general. We have:

$$\begin{aligned}&\text{trace}(\boldsymbol{\mathcal{I}}_{nD})\\ \rightarrow&\text{trace}(\boldsymbol{L}_w^{\mathbb{R}^n})+\text{trace}(\boldsymbol{L}_w^{SO(n)})\\ =&\sum_{i=1}^{n_p}\left(\sum_{j\in V_i}\omega_{ij}+n\sum_{j_1\in V_i^+}\sigma_{ij_1}^{-2}+n\sum_{j_2\in V_i^-}\sigma_{j_2i}^{-2}\right)\\ =&\sum_{i=1}^{n_p}\sum_{j\in V_i}\left(\omega_{ij}+n\widehat{\sigma}_{ij}^{-2}\right)\end{aligned}\tag{73}$$

where $\widehat{\sigma}_{ij}$ means the covariance values corresponding to translation part, which are connected by the i-th and the j-th pose (in and out).

B. D-optimality design metric

D-optimality design is to use the log-determinant of the covariance matrix as an objective function. Due to the sparse structure of the FIM, we always compute $\log(\det(\boldsymbol{C}))$ via $\log(\det(\boldsymbol{F}))$: $\log(\det(\boldsymbol{C}))\approx\log(\det(\boldsymbol{F}^{-1}))=-\log(\det(\boldsymbol{F}))$. Some references show that the D-optimality metric can keep monotonicity during the exploration. Except this, the D-optimality metric and the entropy have an explicit relationship [30]. D-optimality is a useful metric for quantifying the uncertainty of the robot and the generated map in an active SLAM context. However, the determinant function of a matrix is really expensive. In this part, we will derive the bounds of the D-optimality metric which are easier to compute and can be used to approximate the original metric.

Some results about D-optimality design metric for the synchronization of the group $\mathbb{R}^2\times SO(2)$ based on the block-isotropic Gaussian noise has been discussed in [6]. Because the situation using the isotropic Langevin noise on $\mathbb{R}^2\times SO(2)$ is similar to the ones of [6], in this part, we only show the extended result in the group $\mathbb{R}^3\times SO(3)$.

Based on the Schur's determinant formula [6] and (64), because $\boldsymbol{L}_w^{\mathbb{R}^3}$ is invertible, we have:

$$\begin{aligned}&\log(\det\ (\boldsymbol{\mathcal{I}}_{3D}))\\ =&\log\left(\det\ (\boldsymbol{L}_w^{\mathbb{R}^3})\right)+\log(\ \det(\ \boldsymbol{L}_w^{SO(3)}+\\ &\text{diag}\{\boldsymbol{\Psi}_1,\cdots,\boldsymbol{\Psi}_{n_p}\}-\triangle_w^{3D}\boldsymbol{L}_w^{\mathbb{R}^3{}^{-1}}\triangle_w^{3D\top}\ )\ )\end{aligned}\tag{74}$$

Based on the proof of Theorem 3 in [6], $\triangle_w^{3D}\boldsymbol{L}_w^{\mathbb{R}^3{}^{-1}}\triangle_w^{3D\top}\succeq 0$ and $\text{diag}\{\boldsymbol{\Psi}_1,\cdots,\boldsymbol{\Psi}_{n_p}\}-\triangle_w^{3D}\boldsymbol{L}_w^{\mathbb{R}^3{}^{-1}}\triangle_w^{3D\top}\succeq 0$ are orthogonal projection matrices, then we have:

$$\begin{aligned}&\log(\det(\boldsymbol{L}_w^{\mathbb{R}^3}))+\log(\det(\boldsymbol{L}_w^{SO(3)}))=lb\leq\\ &\log\left(\det\ (\boldsymbol{\mathcal{I}}_{3D})\right)\leq\log(\det\ (\boldsymbol{L}_w^{\mathbb{R}^3}))+\log(\ \det(\ \boldsymbol{L}_w^{SO(3)}\\ &+\text{diag}\{\boldsymbol{\Psi}_1,\cdots,\boldsymbol{\Psi}_{n_p}\}\ )\ )<ub=\log(\det(\boldsymbol{L}_w^{\mathbb{R}^3}))+\\ &\log(\det\ (\boldsymbol{L}_w^{SO(3)}+\lambda_\infty\boldsymbol{I}_{3n_p\times 3n_p}))\\ &=\log(\det(\boldsymbol{L}_w^{\mathbb{R}^3}))+\sum_{i=1}^{3n_p}\log(\lambda_i(\boldsymbol{L}_w^{SO(3)})+\lambda_\infty)\end{aligned}\tag{75}$$

where $\lambda_\infty=\|\text{diag}\{\boldsymbol{\Psi}_1,\cdots,\boldsymbol{\Psi}_{n_p}\}\|_\infty$ and $\lambda_i(\boldsymbol{L}_w^{SO(3)})$ means the eigenvalue of $\boldsymbol{L}_w^{SO(3)}$. Some discussions about λ_∞ are shown as follows.

Based on the Rayleigh quotients [37] and Appendix D, we known that the variational description of the maximal eigenvalue of the real symmetric matrix $\boldsymbol{\Psi}_i,\ i=1,2,\cdots,n_p$ is:

$$\begin{aligned}&\lambda_\infty(\boldsymbol{\Psi}_i)=\max_{\boldsymbol{x}\neq 0}\frac{\boldsymbol{x}^\top\boldsymbol{\Psi}_i\boldsymbol{x}}{\boldsymbol{x}^\top\boldsymbol{x}}=\\ &\max_{\boldsymbol{x}\neq 0}\frac{\boldsymbol{x}^T\sum_{j\in V_i^+}\boldsymbol{\Psi}_{(i,j)}\boldsymbol{x}}{\boldsymbol{x}^\top\boldsymbol{x}}\leq\sum_{j\in V_i^+}\max_{\boldsymbol{x}\neq 0}\frac{\boldsymbol{x}^\top\boldsymbol{\Psi}_{(i,j)}\boldsymbol{x}}{\boldsymbol{x}^\top\boldsymbol{x}}\\ &=\sum_{j\in V_i^+}\lambda_\infty(\boldsymbol{\Psi}_{(i,j)})=\sum_{j\in V_i^+}\frac{1}{2}\delta_{ij}^{-2}\|\boldsymbol{x}_j-\boldsymbol{x}_i\|^2\end{aligned}\tag{76}$$

Because all eigenvalues of a block diagonal matrix are the eigenvalues of all block matrices on the diagonal [38], we have:

$$\begin{aligned}\lambda_\infty=&\max_{i=1,2,\cdots,n_p}\lambda_\infty(\boldsymbol{\Psi}_i)\\ \leq&\max_{i=1,2,\cdots,n_p}\frac{1}{2}\sum_{j\in V_i^+}\delta_{ij}^{-2}\|\boldsymbol{x}_j-\boldsymbol{x}_i\|^2\end{aligned}\tag{77}$$

C. Discussion and comparison

1) Efficiency of the metric:

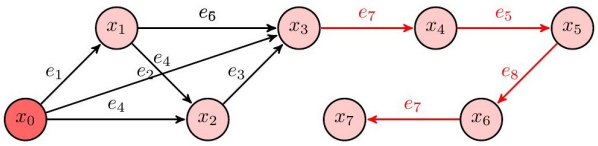
a) T-optimality metric: The parameters ω_{ij} and $\widehat{\sigma}_{ij}$ are constant. We can assume that these two parameters are the same in every measurement process. Set: $c^\omega=\omega_{ij}+3\widehat{\sigma}_{ij}^{-2}$, based on (73), we have:

$$\text{trace}(\boldsymbol{\mathcal{I}}_{nD})\approx\sum_{i=1}^{n_p}\sum_{j\in V_i}c^\omega=c^\omega\sum_{i=1}^{n_p}|V_i|,\ n=2,3\tag{78}$$

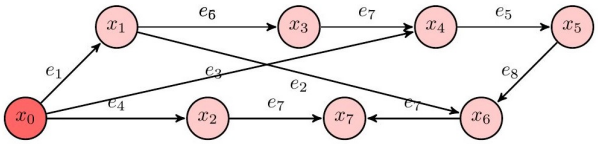
where $\sum_{i=1}^{n_p}|V_i|$ means the total node degree of the graph.

Based on (78), we can get a simple conclusion: For fixed number of pose, minimizing the estimated uncertainty is equivalent to gathering as many measurements as possible. Moreover, we can also see the limitation of the T-optimality metric: when two graphs have the same total node degree and same node number, the T-optimality metric can not be used to compute their uncertainty levels. Let's see following graph in Figure 1:

Because of the same nodes and total node degrees, they have the similar T-optimality objective function based on (75). However, it is easy to know that graph 2 is much better than graph 1 because of the loop closure. The uncertainty of poses $\boldsymbol{x}_4-\boldsymbol{x}_7$ in graph 1 is larger than the ones in graph 2. So in some situations, the T-optimality will lost its efficiency.



(a) G_1 (pose-graph 1).



(b) G_2 (pose-graph 2).

Fig. 1: Two examples of pose-graphs (with $\boldsymbol{x}_0$ as anchor)

b) D-optimality metric: Based on (75) and [6], we have the lower bound ($lb = \log(\det(\boldsymbol{L}_w^{\mathbb{R}^n})) + \log(\det(\boldsymbol{L}_w^{SO(n)}))$) and the upper bound ($ub = \log(\det(\boldsymbol{L}_w^{\mathbb{R}^n})) + \sum_{i=1}^{nn_p} \log(\lambda_i(\boldsymbol{L}_w^{SO(n)}) + \lambda_\infty)$). For most pose-graph problems, the relative translations $\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2$ are usually small, so we have $ub \to lb$. Finally, we can get: $\log(\det(\boldsymbol{\mathcal{I}}_{nD})) \to lb = \log(\det(\boldsymbol{L}_w^{\mathbb{R}^n})) + \log(\det(\boldsymbol{L}_w^{SO(n)}))$. We can the following lemma:

Lemma 4. *Consider the 2D/3D pose-graph SLAM problem in Section III-B,* $\xi \triangleq \max_{i=1,2,\cdots,n_p} \sum_{j\in V_i^+} \delta_{ij}^{-2} \|\boldsymbol{x}_j - \boldsymbol{x}_i\|^2$ *(2D) or* $\xi \triangleq \max_{i=1,2,\cdots,n_p} \frac{1}{2} \sum_{j\in V_i^+} \delta_{ij}^{-2} \|\boldsymbol{x}_j - \boldsymbol{x}_i\|^2$ *(3D) and* $\lambda_{min}(\boldsymbol{L}_w^{SO(n)})$*, define* $\varepsilon = \log(\det(\boldsymbol{\mathcal{I}}_{\boldsymbol{nD}})) - \log(\det(\boldsymbol{L}_w^{\mathbb{R}^n})) - \log(\det(\boldsymbol{L}_w^{SO(n)}))$*. Then we have,*

$$0 \leq \varepsilon \leq dn_p \log(1 + \xi/\lambda_{min}(\boldsymbol{L}_w^{SO(n)})) \tag{79}$$

Proof: See Appendix E.

Based on the Kirchhoff's matrix tree theorem [39], it is easy to know that $\det(\boldsymbol{L}_w^{\mathbb{R}^3})$ and $\det(\boldsymbol{L}_w^{SO(3)})$ are equivalent to the weighted number of the spanning trees of the translation graph and the rotation graph. So, for a weighted graph $\mathcal{G}$ in pose-graph SLAM, the D-optimality design of the weighted Laplacian matrix is almost equal to maximize the weighted number of spanning trees of the graph $\mathcal{G}$ (also named weighted tree connectivity).

The 2D SLAM results with the Gaussian noise about the relationship between the D-optimality metric with the weighted tree connectivity have been discussed in [6]. In this paper, we extend it into the 3D pose-graph SLAM situation. Thus, results and all the further algorithms derived from these results (including $k-ESP^+$ problem [6]) can be extended to 3D case.

2) Computational Complexity:

a) T-optimality metric: After constructing the FIM, the computational complexity of the trace function of the FIM is $O(n_p)$. In some special scenarios and applications, such as the active SLAM, some parts and some functions of the FIM can be reused and form the incremental method. The complexity can be further reduced to $O(L_p)$, where L_p is the predicted horizon. It is easy to find that the T-optimality metric is also a cost-effective metric.

b) D-optimality metric: We can see that the bounds lb and ub of the D-optimality metric of the FIM are almost independent of the values of the pose $\boldsymbol{x}_i$ and $\boldsymbol{R}_i$, which leads to robust performance. Beyond that the update and operations on these bounds are easier than the real D-optimality metric of the FIM. In this part, we talk about the computational complexity of these bounds in 3D case (2D case is similar):

For the lower bound, we have two parts: $\log(\det(\boldsymbol{L}_w^{\mathbb{R}^3}))$ and $\log(\det(\boldsymbol{L}_w^{SO(3)}))$. Using the sparse Cholesky decomposition with a good fill-reducing permutation (Algorithm 1 in [6]), they can be computed much faster than the log-determinant function of the dense matrix $O(n_p^3)$; However, because $\boldsymbol{L}_w^{\mathbb{R}^3}$ and $\boldsymbol{L}_w^{SO(3)}$ have the same sparse structure, we can simply modify the original Algorithm 1 in [6] by the more efficient new algorithm (Algorithm 1):

In Algorithm 1, $\widehat{\boldsymbol{L}_w^{SO(3)}}$ and $\widehat{\boldsymbol{L}_w^{\mathbb{R}^3}}$ are the weighted Laplacian matrix corresponding to the rotation graph and the translation graph without using Kronecker product, i. e. $\boldsymbol{L}_w^{SO(3)} = \widehat{\boldsymbol{L}_w^{SO(3)}} \otimes \boldsymbol{I}_3$ and $\boldsymbol{L}_w^{\mathbb{R}^3} = \widehat{\boldsymbol{L}_w^{\mathbb{R}^3}} \otimes \boldsymbol{I}_3$. $(\star)_{i,i}$ means the i-th diagonal element of the matrix $\star$. At the same time, the same order l can also be used to compute the upper bound.

The above algorithm is similar to the Algorithm 1 in [6]. The only difference is to re-use the result of the column approximate minimum degree reordering algorithm because of the same sparse structure.

For the upper bound ub, there are three parts: $\log(\det(\boldsymbol{L}_w^{\mathbb{R}^3}))$, $\lambda_i(\boldsymbol{L}_w^{SO(3)})$ and λ_∞. For the part $\lambda_i(\boldsymbol{L}_w^{SO(3)})$, we can use the Lanczos algorithm [40] and Fast Multi-pole Method [41] for the sparse Hermitian matrix (Information matrix meets). The Lanczos algorithm can help to generate a tridiagonal real symmetric matrix from the matrix $\boldsymbol{L}_w^{SO(3)}$ in complexity $O(d_n m_p n_p)$ and $O(d_n n_p^2)$ if $(m_p = n_p)$, where d_n is the average number of nonzero elements in a row of the matrix $\boldsymbol{L}_w^{SO(3)}$, m_p is the number of iterations in the Lanczos algorithm (as default, let $m_p = n_p$). For tridiagonal matrices, the Fast Multi-pole Method compute all eigenvalues in just $O(n_p \log n_p)$ operations. So the computing complexity of $\lambda_i(\boldsymbol{L}_w^{SO(3)})$ is totally $O(dn_p^2) + O(n_p \log n_p)$; For the part λ_∞, we have shown the analytical results $\lambda_\infty \approx \max_{i=1,2,\cdots,n_p} \frac{1}{2} \sum_{j\in V_i^+} \delta_{ij}^{-2} \|\boldsymbol{p}_{ij}\|^2$.

The above discussion is based on the operations on the whole FIM. Similar to the T-optimality metric, in some special scenarios and applications, such as the active SLAM, the incremental computation using the matrix determinant and re-use of calculation method (rAMDL) has been introduced into the computation of the D-optimality metric of the FIM [33]. Without the loop-closure, its computational complexity is reduced to $O(L_p^3)$, where L_p is the look-ahead step which is usually constant and independent of the pose number n_p. In future, we will try to apply the rAMDL method in the computation of these bounds to further improve the running ability. This point is one of our future research directions.

VIII. Applications

A. Active SLAM

The active SLAM is a decision making problem where the

Algorithm 1 Lower bound computation for D-optimality metric (3D case)

procedure LB($\widehat{\boldsymbol{L}_w^{SO(3)}}, \widehat{\boldsymbol{L}_w^{\mathbb{R}^3}}$)
  $\boldsymbol{l} \leftarrow COLAMD(\widehat{\boldsymbol{L}_w^{SO(3)}})$ ▷ Column approximate minimum degree permutation
  $\boldsymbol{L}_1 \leftarrow SparseCholesky(\widehat{\boldsymbol{L}_w^{SO(3)}}(\boldsymbol{l},\boldsymbol{l}))$ ▷ Sparse Cholesky factor based on $\boldsymbol{l}$ for $\widehat{\boldsymbol{L}_w^{SO(3)}}$
  $\boldsymbol{L}_2 \leftarrow SparseCholesky(\widehat{\boldsymbol{L}_w^{\mathbb{R}^3}}(\boldsymbol{l},\boldsymbol{l}))$ ▷ Sparse Cholesky factor based on the same $\boldsymbol{l}$ for $\widehat{\boldsymbol{L}_w^{\mathbb{R}^3}}$
  return $2 \cdot 3 \cdot \sum_i (\log(\boldsymbol{L}_1)_{i,i} + \log(\boldsymbol{L}_2)_{i,i})$

robot's trajectory is chosen both to improve SLAM results, and meanwhile, to perform other tasks such as coverage or exploration [42]. The optimality design metrics of the FIM and the covariance matrix are the most common objective function in the active SLAM [43] and other belief space planning problems [33]. The computation complexity of the objective function is one of the key questions in the active SLAM. This paper presents the new metrics, including weighted node degrees and weighted tree connectivity, in 3D situation for the active SLAM, which helps to improve the real-time ability of the algorithm [42]. In future, we will apply these metrics to perform the on-line active SLAM in some plat-forms, like quad-rotor UAV and other robots.

B. Measurement selection

Because of the limitation of the weight, size, power budgets, and on-board hardware capability of the robot, for the scalable and long-term autonomy, the measurement selection is one of several mechanisms through which the SLAM systems can achieve resource adaptation. In the work [6], the edge selection problem, which is proved to be the sub-modular optimization problem, is presented based on the 2D result and solved by the greedy algorithm and the convex relaxation method. This edge selection problem shows its enormous potential applications in the front/back-end computational reduction [6] and communication resource assignment in cooperative SLAM [44].

IX. Experimental results

In this section, we validate the correctness of our conclusions and evaluate the performance of the bounds on a variety of synchronization problems. These experiments are based on a variety of 2D/3D rigid body motion synchronization problems drawn from the pose-graph SLAM. The datasets and the estimation results in these simulation results are obtained based on the open source of the literature [3].

All of the following experiments are performed on a HP EliteDesk 800 G2 desktop with an Intel Core i5-6500 3.20 GHz processor and 8 GB of RAM running Windows 10 Enterprise. Our experimental implementations are written in MATLAB R2016a.

A. Efficency of the metrics

In this experiment, we verify the results in Section VII-A3 and VII-C1a, which show that the T-optimality metric has a strong and direct relationship with the total node degree of the pose-graph and the D-optimality metric has a better performance than the T-optimality metric. We use a small example with 8 nodes to simulate their relation. A regular octagon is constructed based on these eight poses. The variances of the translation δ_{ij}^2 and the concentrations κ_{ij} of all measurements are 1×10^{-4} and 5×10^3. The initial pose graph only includes the odometer without the additional measurements, as shown in Figure 2. Then, we begin to add the measurements gradually from one edge to 21 edges without repetition. It is noted that because the first pose is the anchor, the additional edges are not allowed to connected with the first pose. For example in Figure 3, we add 3 additional measurements. Then, for every graph, we can estimate the trace function and the log-determinant function of the FIM after using the SE-sync [3]. The results (Black points) and their corresponding total node degrees (Red Line) are shown in Figure 4 and 5.

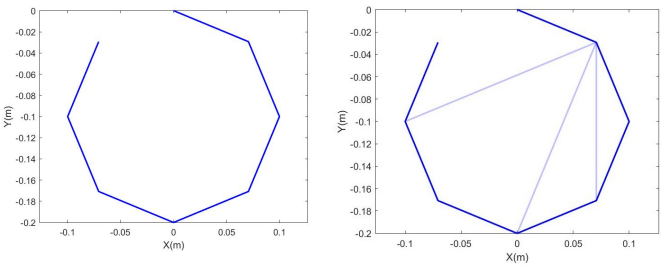


Fig. 2: Inital pose-graph

Fig. 3: Pose-graph added 3 measurements

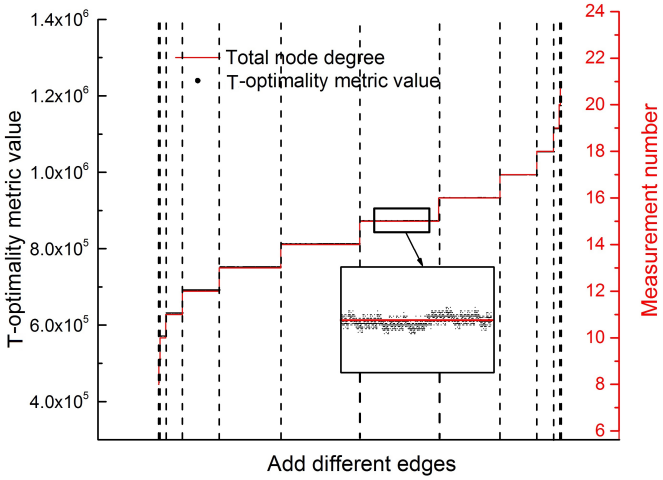


Fig. 4: Direct relationship between T-optimality metric with total node degree

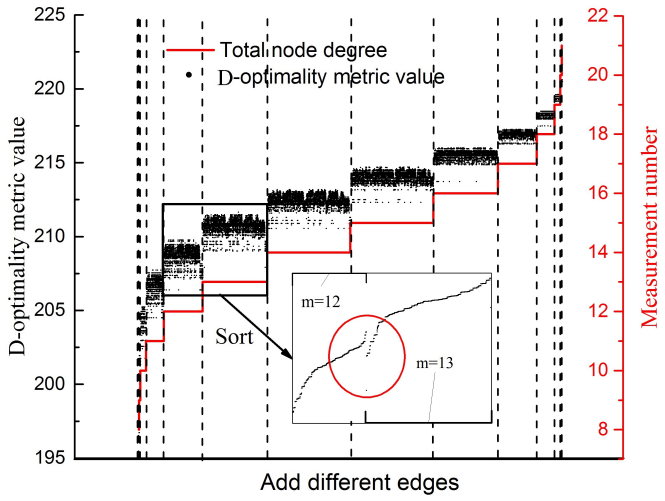


Fig. 5: Part relationship between D-optimality metric with total node degree

In Figure 4, because the term $\sum_{i=1}^{n_p}\sum_{j\in V_i^+}\delta_{ij}^{-2}\|\boldsymbol{p}_{ij}\|^2$ is relatively small, we can find that the T-optimality metric is almost proportional to the total node degree of the whole pose graph. The T-optimality metric values of the graphs, which have the same node degrees, are similar. It may lead the obvious local minimal problem in its application. In Figure 5, we can learn that the D-optimality metric has a weaker connection with the total node degree compared with the T-optimality metric. We sort the D-optimality metric values of the graphs with $m = 12$ and $m = 13$ measurements. Pay attention to the red circle part. We find that the D-optimality metric values of some graphs with more measurements are worse than the ones of the graphs with less measurements. This situation is common for many real systems.

B. Bound efficiency on D-optimality metric

In this set of experiments, we evaluate the bound efficiency and computational advantage on a variety of classical heterogeneous 3D pose-graph SLAM benchmarks. These datasets include: the synthetic datasets (3D: sphere, torus, and tiny/small/normal grid datasets) and the real-world datasets (2D: CSAIL, Intel Research Lab, manhattan (M3500), KITTI, city10000, and ais2klinik datasets; 3D: garage, cubicle, and rim datasets).

For the lower bound, the proposed Algorithm 1 is used to compute it. The largest eigenvalue of the weighted Laplacian matrix is obtained using the QR decomposition with the same ordering shown in Algorithm 1. Results for these experiments are shown in Table I (2D) and Table II (3D). We demonstrate the number of the poses and the measurements, the log-determinant functions of the FIM, its upper and lower bounds, and their computational time.

On each of these examples, the log-determinant function of the FIM is bounded within lb and ub successfully. We can also find that, the benefit from the great dimensionality reduction, the computation of the log-determinant function of the FIM is much more expensive than the ones of the upper and lower bounds. Because the dimension of the weighted Laplacian matrix is one-fourth (2D)/one-sixth (3D) of that of the full FIM, this great computational gap is sensible.

It is easy to find that, except the 'garage' dataset, the log-determinant function of the whole FIM gets closed to its lower bound (in 3%), because of the small $\sum_{i=1}^{n_p}\sum_{j\in V_i^+}\delta_{ij}^{-2}\|\boldsymbol{p}_{ij}\|^2$. So in the real application of the D-optimality metric, such as the active pose-graph SLAM [45], we suggest using the tight lower bound to replace the original objective function.

Besides the computational time of obtaining these metrics, we also show the computational time of constructing the whole FIM, the weighted Laplacian matrix for $\mathbb{R}^2$ and $\mathbb{R}^3$ parts and the weighted Laplacian matrix for $SO(2)$ and $SO(3)$ part. They are used to compute the metrics. Table III and Table IV show the computation time. The results show that the sparser weighted Laplacian matrices for the lower bound are easier to be constructed than the whole FIM.

TABLE III: Matrix constructing time for the 2D SLAM datasets

	Time $[s]$	
Dataset	Whole FIM	Weighted Laplacian matrices
CSAIL	8.8581×10^{-2}	1.5723×10^{-2}
Intel	1.3518×10^{-1}	1.9283×10^{-2}
manhattan	2.3982×10^{-1}	2.6794×10^{-2}
KITTI	2.2071×10^{-1}	2.5118×10^{-2}
city10000	7.3709×10^{-1}	6.3875×10^{-2}
ais2klinik	6.3033×10^{-1}	5.4222×10^{-2}

TABLE IV: Matrix constructing time for the 3D SLAM datasets

	Time $[s]$	
Dataset	Whole FIM	Weighted Laplacian matrices
tiny-grid	7.1865×10^{-2}	2.0968×10^{-2}
small-grid	1.1677×10^{-1}	1.9557×10^{-2}
garage	9.6897×10^{-1}	5.3639×10^{-2}
sphere	8.3076×10^{-1}	4.8680×10^{-2}
torus	1.3884×10^{0}	7.3766×10^{-2}
cubicle	2.3300×10^{0}	1.1033×10^{-1}
grid	3.0566×10^{0}	1.1413×10^{-1}
rim	4.0647×10^{0}	1.7824×10^{-1}

C. Efficiency of CRLB

In this section, our main purpose is to validate the CRLB is reachable using the SE-sycn pose-graph SLAM method (which belongs to maximum likelihood estimator MLE) [3]. These experiments are based on 'tiny-3Dgrid' and 'CSAIL' problems drawn from the pose-graph SLAM.

In order to compute the covariance by the statistical way, we need a ground truth and then sample noises to it. The ground truth is obtained by the optimization results of original datasets present in [3] using SE-sync. Then, we set these estimated poses as the ground truth. After obtaining the ground truth, the random noises obeying the isotropic Gaussian distribution and the isotropic Lagevin distribution are generated. We

TABLE I: Metric Results for the 2D SLAM datasets

Dataset	# Poses	# Measurements	Original FIM		Upper bound		Lower bound	
			Metric value	Time $[s]$	Value	Time $[s]$	Value	Time $[s]$
CSAIL	1045	1172	1.9860×10^{4}	2.3284×10^{-3}	1.9876×10^{4}	9.2391×10^{-4}	1.9858×10^{4}	3.6712×10^{-3}
Intel	1728	2512	3.0194×10^{4}	6.0744×10^{-3}	3.0352×10^{4}	4.0055×10^{-3}	3.0155×10^{4}	2.0743×10^{-3}
manhattan	3500	5453	7.1124×10^{4}	3.3530×10^{-1}	7.1193×10^{4}	7.3602×10^{-3}	7.1104×10^{4}	4.5105×10^{-3}
KITTI	4541	4677	1.1697×10^{5}	8.9039×10^{-3}	1.1703×10^{5}	4.6468×10^{-3}	1.1695×10^{5}	3.1092×10^{-3}
city10000	10000	20687	1.6667×10^{5}	9.0052×10^{-2}	1.6816×10^{5}	2.4292×10^{-2}	1.6520×10^{5}	2.0664×10^{-2}
ais2klinik	15115	16727	2.3642×10^{5}	3.6301×10^{-2}	2.3975×10^{5}	1.4693×10^{-2}	2.3567×10^{5}	1.3116×10^{-2}

TABLE II: Metric Results for the 3D SLAM datasets

Dataset	# Poses	# Measurements	Original FIM		Upper bound		Lower bound	
			Metric value	Time $[s]$	Value	Time $[s]$	Value	Time $[s]$
tiny-grid	9	11	2.4562×10^{2}	2.1558×10^{-4}	2.7561×10^{2}	2.9866×10^{-2}	2.4163×10^{2}	1.4115×10^{-4}
small-grid	125	297	4.4452×10^{3}	3.6373×10^{-3}	4.8488×10^{3}	3.1836×10^{-2}	4.3815×10^{3}	4.3020×10^{-4}
garage	1661	6275	2.0618×10^{4}	3.9511×10^{-2}	3.6618×10^{4}	3.6117×10^{-2}	1.5845×10^{4}	2.2989×10^{-3}
sphere	2500	4949	9.9105×10^{4}	7.4856×10^{-2}	1.0607×10^{5}	7.4023×10^{-2}	9.9054×10^{4}	6.7317×10^{-3}
torus	5000	9048	2.4986×10^{5}	1.9488×10^{-1}	2.6526×10^{5}	1.0156×10^{-1}	2.4985×10^{5}	1.6390×10^{-2}
cubicle	5750	16869	2.3729×10^{5}	2.1579×10^{-1}	3.1839×10^{5}	5.2156×10^{-2}	2.3685×10^{5}	1.4141×10^{-2}
grid	8000	22236	4.2613×10^{5}	3.2318×10^{0}	4.4902×10^{5}	1.5065×10^{-1}	4.2610×10^{5}	9.0280×10^{-2}
rim	10195	29743	4.7257×10^{5}	4.1626×10^{-1}	5.8622×10^{5}	6.5298×10^{-2}	4.7209×10^{5}	4.1260×10^{-2}

use 'normrnd' MATLAB function to generate the isotropic Gaussian distribution. For the isotropic Lagevin distribution, the noises are generated by the Acceptance-Rejection Method (ARM) [46]. Then we can add these noises into our relative measurements by the edge data using following equations:

$$\begin{aligned}\boldsymbol{H}_{ij} =&\boldsymbol{Z}_{ij}\boldsymbol{R}_j\boldsymbol{R}_i^{\top}, \boldsymbol{Z}_{ij}\sim Lang(\boldsymbol{I}_{n\times n}, k_c\cdot\kappa_{ij})\\ \boldsymbol{p}_{ij} =&\boldsymbol{R}_i^{\top}(\boldsymbol{x}_j-\boldsymbol{x}_i)+\boldsymbol{y}_{ij}, \boldsymbol{y}_{ij}\sim\mathcal{N}(\boldsymbol{0}, {\delta_{ij}}^2\boldsymbol{I}_{n\times n})\end{aligned} \tag{80}$$

where k_c is the coefficient to determine the uncertainty level.

For an estimated result, we can compute the trace function of covariance by:

$$\text{trace}(\boldsymbol{C}) = \mathbb{E}\{\sum_{i=1}^{n_p}\left(dist(\boldsymbol{R}_i,\bar{\boldsymbol{R}}_i)^2+\|\boldsymbol{x}_i-\bar{\boldsymbol{x}}_i\|_2^2\right)\} \tag{81}$$

where $\bar{\star}$ means the ground truth of $\star$. Every measurement dataset can generate a value using (81), then repeat many times to obtain the mathematical expectation $\mathbb{E}\{\text{trace}(\boldsymbol{C})\}$ by average. Finally, we can compute and compare the CRLB and the average mean squared error (MSE) for every pose.

For the 2D situation, we use the 'CSAIL' dataset to obtain the average MSE and CRLB. The initial κ_{ij} and δ_{ij}^{-2} are respectively set as 150 and 140. The coefficient k_c changes from 1 to 20. The simulations are repeated 100 times to get the covariance matrix. The simulation results are shown in Fig. 6.

For the 3D situation, we use the 'tiny-3Dgrid' dataset to obtain the covariance, CRLB with curvature terms and CRLB without curvature terms. κ_{ij}=12.5. Then the coefficient k_c changes from 1 to 20. Based on the Monte Carlo simulation, the processes are repeated 50 times. The simulation results are shown in Fig.7.

The above two simulations show that the CRLB is reachable when the SNR is large enough.

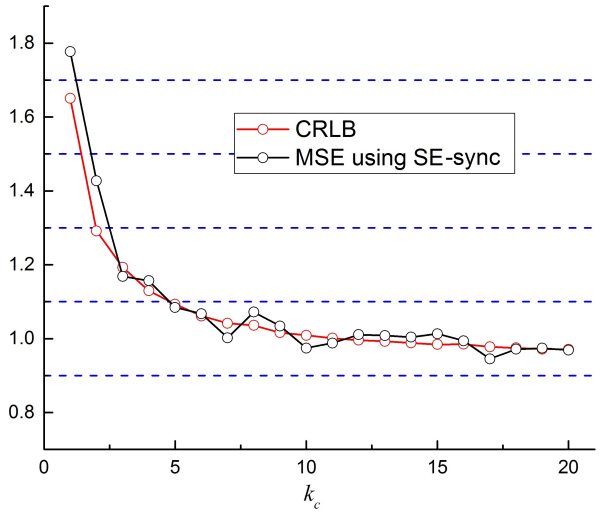


Fig. 6: Comparison of the CRLB with the mean squared error ($MSE = \frac{\text{trace}(\boldsymbol{C})}{n_p}$) of known estimators for the synchronization of $n_p = 1045$ poses with a complete measurement graph and one anchor. The SE-sync method appears to reach the CRLB.

X. Conclusion and Future Work

Based on the assumption of the isotropic Langevin noise for rotation and the block-isotropic Gaussian noise for translation, the FIM and CRLB of the 2D/3D pose-graph SLAM, which are formulated as the synchronization on $\mathbb{R}^n\times SO(n)$, $n=2,3$, are derived and shown to be closely related to the weighted Laplacian matrix of the pose-graph SLAM. Then, the TOED, including T-optimality and D-optimality are discussed. It shows that the T-optimality metric is almost directly determined by the total node degrees, and the D-optimality design is

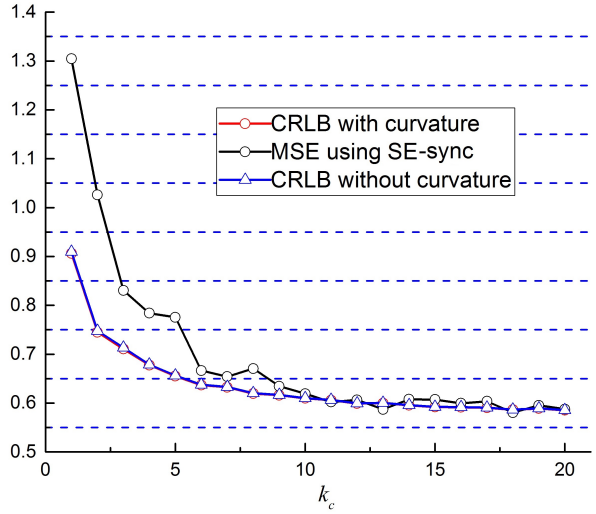


Fig. 7: Comparison of the CRLB with the mean squared error ($MSE = \frac{\text{trace}(\boldsymbol{C})}{n_p}$) of known estimators for the synchronization of $n_p = 9$ poses with a complete measurement graph and one anchor. The SE-sync method appears to reach the CRLB.

equivalent to maximize the weighted number of the spanning trees of the pose-graph. We find that the T-optimality metric is cheaper than the D-optimality metric, but the D-optimality metric has a better performance. Furthermore, the lower and upper bounds of the D-optimality metric are presented, which are cheaper and can be used to replace the original metric. The experiment results, which verify our standpoints, are performed based on both the synthetic dataset and the real-world dataset.

Finally, the results presented in this paper are only the first steps towards the use of graph topology in areas such as the active SLAM. Similar to our other previous work [42], we are trying to use these proposed technologies into the planning and decision making pipelines.

APPENDIX A
DERIVATION OF $\text{grad}_i\ \log\ f(\boldsymbol{H}\boldsymbol{R}_i)$

Let f is a function $SO(n) \to \mathbb{R}$, $\boldsymbol{H}$ and $\boldsymbol{R}_i \in SO(n)$, then this part is used to derivate the gradient of this function $\log\ f(\boldsymbol{H}\boldsymbol{R}_i)$ on $SO(n)$.

We firstly define the directional derivative of $\log(f(\boldsymbol{H}\boldsymbol{R}_i))$ at $\boldsymbol{R}_i$ along $\boldsymbol{U}$, where $\boldsymbol{U}$ is a $n \times n$ matrix:

$$\mathrm{D}_i \log\ f(\boldsymbol{H}\boldsymbol{R}_i)[\boldsymbol{U}] =< \bigtriangledown_{\boldsymbol{R}_i} \log(f(\boldsymbol{H}\boldsymbol{R}_i)), \boldsymbol{U} >_{\boldsymbol{R}} \quad (82)$$

Then we have a chain rule:

$$\mathrm{D}_i \log f(\boldsymbol{H}\boldsymbol{R}_i)[\boldsymbol{U}] = (f(\boldsymbol{H}\boldsymbol{R}_i))^{-1} \mathrm{D}_{\boldsymbol{H}\boldsymbol{R}_i} f(\boldsymbol{H}\boldsymbol{R}_i)[\boldsymbol{H}\boldsymbol{U}] \quad (83)$$

where $\mathrm{D}_{\boldsymbol{H}\boldsymbol{R}_i} f(\boldsymbol{H}\boldsymbol{R}_i)[\boldsymbol{H}\boldsymbol{U}] =< \bigtriangledown_{\boldsymbol{H}\boldsymbol{R}_i} f(\boldsymbol{H}\boldsymbol{R}_i), \boldsymbol{H}\boldsymbol{U} >_{\boldsymbol{R}}$. We can remove $\boldsymbol{H}$ into the other side, we have $\mathrm{D}_{\boldsymbol{H}\boldsymbol{R}_i} f(\boldsymbol{H}\boldsymbol{R}_i)[\boldsymbol{H}\boldsymbol{U}] =< \boldsymbol{H}^\top \bigtriangledown_{\boldsymbol{H}\boldsymbol{R}_i} f(\boldsymbol{H}\boldsymbol{R}_i), \boldsymbol{U} >_{\boldsymbol{R}}$. We can get:

$$\begin{aligned} &\mathrm{D}_i \log f(\boldsymbol{H}\boldsymbol{R}_i)[\boldsymbol{U}] \\ &=< (f(\boldsymbol{H}\boldsymbol{R}_i))^{-1} \boldsymbol{H}^\top \bigtriangledown_{\boldsymbol{H}\boldsymbol{R}_i} f(\boldsymbol{H}\boldsymbol{R}_i), \boldsymbol{U} >_{\boldsymbol{R}} \end{aligned} \quad (84)$$

Compare (82) and (84), we can get:

$$\bigtriangledown_{\boldsymbol{R}_i} \log(f(\boldsymbol{H}\boldsymbol{R}_i)) = (f(\boldsymbol{H}\boldsymbol{R}_i))^{-1} \boldsymbol{H}^\top \bigtriangledown_{\boldsymbol{H}\boldsymbol{R}_i} f(\boldsymbol{H}\boldsymbol{R}_i) \quad (85)$$

Based on (9), the gradient of $\log(f(\boldsymbol{H}\boldsymbol{R}_i))$ can be computed by:

$$\begin{aligned} &\text{grad}_i\ \log f(\boldsymbol{H}\boldsymbol{R}_i) \\ =&\boldsymbol{R}_i \text{skew}(\boldsymbol{R}_i^\top \bigtriangledown_{\boldsymbol{R}_i} \log(f(\boldsymbol{H}\boldsymbol{R}_i))) \\ =&\boldsymbol{R}_i \text{skew}(\boldsymbol{R}_i^\top (f(\boldsymbol{H}\boldsymbol{R}_i))^{-1} \boldsymbol{H}^\top \bigtriangledown_{\boldsymbol{H}\boldsymbol{R}_i} f(\boldsymbol{H}\boldsymbol{R}_i)) \end{aligned} \quad (86)$$

APPENDIX B
PROOF OF LEMMA 2

If $\boldsymbol{\Phi}_{ij} \to \boldsymbol{0}$, $\boldsymbol{Z}_{ij} = \exp(\boldsymbol{\Phi}_{ij}^\wedge) \approx \begin{bmatrix} 1 & -\Phi_3^{ij} & \Phi_2^{ij} \\ \Phi_3^{ij} & 1 & -\Phi_1^{ij} \\ -\Phi_2^{ij} & \Phi_1^{ij} & 1 \end{bmatrix}$ [35], we have:

$$\begin{aligned} \text{skew}(\boldsymbol{Z}_{ij}) &= \begin{bmatrix} 0 & -\Phi_3^{ij} & \Phi_2^{ij} \\ \Phi_3^{ij} & 0 & -\Phi_1^{ij} \\ -\Phi_2^{ij} & \Phi_1^{ij} & 0 \end{bmatrix} \\ &= \sqrt{2}(\Phi_1^{ij}\boldsymbol{E}_1 + \Phi_2^{ij}\boldsymbol{E}_2 + \Phi_3^{ij}\boldsymbol{E}_3) \end{aligned} \quad (87)$$

We also have:

$$\Phi_k^{ij} = \frac{1}{\sqrt{2}} < \text{skew}(\boldsymbol{Z}_{ij}), \boldsymbol{E}_k >_{\boldsymbol{R}} \quad (88)$$

Substitute (88) into (87), we have:

$$\text{skew}(\boldsymbol{Z}_{ij}) = \sum_{k=1}^{3} < \text{skew}(\boldsymbol{Z}_{ij}), \boldsymbol{E}_k >_{\boldsymbol{R}} \boldsymbol{E}_k \quad (89)$$

Let's analysis the isotropic Lagevin distribution first: For every $\boldsymbol{Z} \in SO(3)$, there exists some $\boldsymbol{U} \in O(3)$ such that [3]:

$$\boldsymbol{U}^\top \boldsymbol{Z} \boldsymbol{U} = \begin{bmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (90)$$

Based on (90) and (89), the trace appearing in the isotropic Langevin density can be equivalently expressed as:

$$\text{trace}(\boldsymbol{Z}) = \text{trace}(\boldsymbol{U}^\top \boldsymbol{Z} \boldsymbol{U}) = 1 + 2\cos(\theta) \quad (91)$$

It is easy to know that θ is independent of the choose of basis $\boldsymbol{E}_k$ and $\boldsymbol{E}_l$. θ is depend on the magnitude of the rotation. So it is isotropic based on the basis. We have:

$$\begin{aligned} \mathbb{E}\{< \text{skew}(\boldsymbol{Z}_{ij}), \boldsymbol{E}_k >_{\boldsymbol{R}}^2\} &= \mathbb{E}\{< \text{skew}(\boldsymbol{Z}_{ij}), \boldsymbol{E}_l >_{\boldsymbol{R}}^2\} \\ \Longleftrightarrow \mathbb{E}\{{\Phi_k^{ij}}^2\} &= \mathbb{E}\{{\Phi_l^{ij}}^2\} \end{aligned} \quad (92)$$

Based on (89) and (92), we have:

$$\begin{aligned} &\mathbb{E}\{\|\text{skew}(\boldsymbol{Z}_{ij})\|_F^2\} \\ =&\mathbb{E}\{\| \sum_{k=1}^{3} < \text{skew}(\boldsymbol{Z}_{ij}), \boldsymbol{E}_k >_{\boldsymbol{R}} \boldsymbol{E}_k \|_F^2\} \\ =&\mathbb{E}\{\sum_{k=1}^{3} 2 \cdot \frac{1}{2} \left(< \text{skew}(\boldsymbol{Z}_{ij}), \boldsymbol{E}_k >_{\boldsymbol{R}}^\top \cdot < \boldsymbol{E}_k, \text{skew}(\boldsymbol{Z}_{ij}) >_{\boldsymbol{R}}\right) \\ =&6\mathbb{E}\{{\Phi_k^{ij}}^2\} \end{aligned} \quad (93)$$

So we have: $\mathbb{E}\{{\Phi_k^{ij}}^2\} = \mathbb{E}\{{\Phi_l^{ij}}^2\} = \frac{1}{2}\mathbb{E}\{< \text{skew}(\boldsymbol{Z}_{ij}), \boldsymbol{E}_k >_{\boldsymbol{R}}^2\} = \frac{1}{6}\mathbb{E}\{\|\text{skew}(\boldsymbol{Z}_{ij})\|_F^2\}$

APPENDIX C
THE COMPUTATION OF THE WEIGHT ω_{ij}

Based on [23], the PDF of the isotropic Langevin distribution on $SO(n)$ are class functions, which can meet the following property, which is translated by the extended bi-invariance property:

$$\begin{aligned} &\forall\ \boldsymbol{U}_1, \boldsymbol{U}_2 \in O(n)\ s.t.\ \det(\boldsymbol{U}_1\boldsymbol{U}_2) = 1, \\ &\int_{SO(n)} f(\boldsymbol{U}_1\boldsymbol{Z}\boldsymbol{U}_2)d\mu(\boldsymbol{Z}) = \int_{SO(n)} f(\boldsymbol{Z})d\mu(\boldsymbol{Z}). \end{aligned} \tag{94}$$

So we can use the Weyl integration formula specialized to $SO(3)$ [47]. Let $\boldsymbol{Z} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}$,We have:

$$\begin{aligned} \omega_{ij} =& \kappa_{ij}^2 \mathbb{E}\{\|\text{skew}(\boldsymbol{Z}_{ij})\|_F^2\} \\ =& \frac{\kappa_{ij}^2}{4} \int_{SO(3)} \|\boldsymbol{Z} - \boldsymbol{Z}^T\|_F^2 \widehat{f}_{ij}(\boldsymbol{Z}) d\mu(\boldsymbol{Z}) \\ =& \frac{\kappa_{ij}^2}{4} \frac{1}{2\pi} \int_{-\pi}^{-\pi} 8\sin^2(\theta) \widehat{f}_{ij}(\boldsymbol{Z}) (1-\cos(\theta)) d\theta \\ =& \frac{2\kappa_{ij}^2 \int_{-\pi}^{-\pi} \sin^2(\theta)(1-\cos(\theta)) \exp(2\kappa_{ij}\cos(\theta)) d\theta}{2\pi(I_0(2\kappa_{ij}) - I_1(2\kappa_{ij}))} \\ =& \frac{2\kappa_{ij}^2}{I_0(2\kappa_{ij}) - I_1(2\kappa_{ij})} \frac{1}{2\pi} \int_{-\pi}^{-\pi} \rho(\theta) \exp(2\kappa_{ij}\cos(\theta)) d\theta \end{aligned} \tag{95}$$

where $\rho(\theta) = (1 - \cos(\theta) - \cos^2(\theta) + \cos^3(\theta))$.

We know that:

$$\begin{aligned} \cos(0\theta) &= 1 \\ \cos(1\theta) &= \cos(\theta) \\ \cos(2\theta) &= 2\cos^2(\theta) - 1 \\ \cos(3\theta) &= 4\cos^3(\theta) - 3\cos(\theta) \end{aligned} \tag{96}$$

Based on some trigonometric operations, we can get:

$$\begin{aligned} &\omega_{ij} \\ =& \frac{2\kappa_{ij}^2 \left(\frac{1}{2} I_0(2\kappa_{ij}) - \frac{1}{4} I_1(2\kappa_{ij}) - \frac{I_2(2\kappa_{ij})}{2} + \frac{I_3(2\kappa_{ij})}{4} \right)}{I_0(2\kappa_{ij}) - I_1(2\kappa_{ij})} \\ =& \frac{\kappa_{ij}^2 \left(2I_0(2\kappa_{ij}) - I_1(2\kappa_{ij}) - 2I_2(2\kappa_{ij}) + I_3(2\kappa_{ij}) \right)}{2I_0(2\kappa_{ij}) - 2I_1(2\kappa_{ij})} \end{aligned} \tag{97}$$

APPENDIX D
PROOF OF THE EIGENVALUES OF $\boldsymbol{\Psi}_i$

$\boldsymbol{\Psi}_i$ can be written as:

$$\begin{aligned} \boldsymbol{\Psi}_i &= \sum_{j \in V_i^+} \boldsymbol{\Psi}_{(i,j)} \\ \boldsymbol{\Psi}_{(i,j)} &= \begin{bmatrix} \psi_{(i,j)}^{11} & \psi_{(i,j)}^{12} & \psi_{(i,j)}^{13} \\ \psi_{(i,j)}^{12} & \psi_{(i,j)}^{22} & \psi_{(i,j)}^{21} \\ \psi_{(i,j)}^{13} & \psi_{(i,j)}^{21} & \psi_{(i,j)}^{33} \end{bmatrix} \\ \psi_{(i,j)}^{kl} &= \delta_{ij}^{-2} (\boldsymbol{x}_i - \boldsymbol{x}_j)^\top \boldsymbol{R}_i \boldsymbol{I}_{3\times 3}^{k,l} \boldsymbol{R}_i^\top (\boldsymbol{x}_i - \boldsymbol{x}_j) \end{aligned} \tag{98}$$

Let $\frac{1}{\sqrt{2}} \boldsymbol{R}_i^\top (\boldsymbol{x}_i - \boldsymbol{x}_j) = (p_1,\ p_2,\ p_3)^\top$, we can get:

$$\boldsymbol{\Psi}_{(i,j)} = \delta_{ij}^{-2} \begin{bmatrix} p_2^2 + p_3^2 & -p_1 p_2 & -p_1 p_3 \\ -p_1 p_2 & p_1^2 + p_3^2 & -p_2 p_3 \\ -p_1 p_3 & -p_2 p_3 & p_1^2 + p_2^2 \end{bmatrix} \tag{99}$$

For a 3×3 maxtrix $\boldsymbol{B} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$, $\det(\lambda \boldsymbol{I} - \boldsymbol{B}) = 0$ can be written as:

$$\begin{aligned} &\lambda^3 + k_1\lambda^2 + k_2\lambda + k_3 = 0 \\ &k_1 = -b_{11} - b_{22} - b_{33} \\ &k_2 = b_{11}b_{22} + b_{11}b_{33} + b_{22}b_{33} - b_{12}^2 - b_{23}^2 - b_{33}^2 \\ &k_3 = -\det(\boldsymbol{B}) \end{aligned} \tag{100}$$

We can find that $\det(\boldsymbol{\Psi}_{(i,j)}) = 0$ easily. So $k_3 = \det(\boldsymbol{\Psi}_{(i,j)}) = 0$. Substitute (99) into (100), we have:

$$\begin{aligned} &\lambda^2 + k_1\lambda + k_2 = 0 \\ &k_1 = -2(p_1^2 + p_2^2 + p_3^2) \\ &k_2 = (p_1^2 + p_2^2 + p_3^2)^2 \\ \Longrightarrow &\left(\lambda - (p_1^2 + p_2^2 + p_3^2)\right)^2 = 0 \end{aligned} \tag{101}$$

Finally, based on $p_1^2 + p_2^2 + p_3^2 = (\boldsymbol{x}_i - \boldsymbol{x}_j)^\top \boldsymbol{R}_i \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \boldsymbol{R}_i^\top (\boldsymbol{x}_i - \boldsymbol{x}_j) = \frac{1}{2}\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2$, we can known that the eigenvalues of $\boldsymbol{\Psi}_{(i,j)}$ are $\lambda_1^{ij} = \lambda_2^{ij} = \frac{1}{2}\delta_{ij}^{-2}\|\boldsymbol{x}_j - \boldsymbol{x}_i\|^2$ and $\lambda_3^{ij} = 0$.

APPENDIX E
PROOF OF LEMMA 4

Based on Eq. (75) and Eq. (77), we have:

$$\begin{aligned} \varepsilon &= \log(\det(\boldsymbol{\mathcal{I}}_{\boldsymbol{nD}})) - \log(\det(\boldsymbol{L}_w^{\mathbb{R}^n})) - \log(\det(\boldsymbol{L}_w^{SO(n)})) \\ &\le \log(\det(\boldsymbol{L}_w^{SO(n)} + \xi \boldsymbol{I}_{dn_p \times dn_p})) - \log(\det(\boldsymbol{L}_w^{SO(n)})) \\ &= \log \prod_{i=1}^{dn_p} \frac{\lambda_i(\boldsymbol{L}_w^{SO(n)}) + \xi}{\lambda_i(\boldsymbol{L}_w^{SO(n)})} \\ &= \log \prod_{i=1}^{dn_p} \left(1 + \frac{\xi}{\lambda_i(\boldsymbol{L}_w^{SO(n)})}\right) \\ &\le \log \left(1 + \frac{\xi}{\lambda_{min}(\boldsymbol{L}_w^{SO(n)})}\right)^{dn_p} \\ &= dn_p \log\left(1 + \xi/\lambda_{min}(\boldsymbol{L}_w^{SO(n)})\right) \end{aligned} \tag{102}$$

where $\lambda_i(\boldsymbol{L}_w^{SO(n)})$ are the eigenvalues of $\boldsymbol{L}_w^{SO(n)}$, $\lambda_{min}(\boldsymbol{L}_w^{SO(n)})$ is the minimal eigenvalue of $\boldsymbol{L}_w^{SO(n)}$.

ACKNOWLEDGMENT

We would like to thank Dr. N. Boumal. He has answered many questions about his research [1] and [23], which helps us significantly in this research. We also want to thank David M. Rosen. He offers us the code to generate the random noise for our simulation.

REFERENCES

[1] N. Boumal, "On intrisic Cramér-Rao bounds for Riemannian submanifolds and quotient manifolds," *IEEE Transactions Signal Process*, vol. 61, no. 7, pp. 1809-1821, 2013.

[2] 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, 2016, pp. 1309-1332.

[3] D. M. Rosen, L. Carlone, A. S. Bandeira, and J. J. Leonard, "SE-Sync: A Certifiably Correct Algorithm for Synchronization over the Special Euclidean Group," arXiv Prepr. arXiv1612.07386, pp. 149, 2016.

[4] J. Briales, J. Gonzalez-Jimenez, "Cartan-Sync: Fast and Global SE (d)-Synchronization," *IEEE Robotics and Automation Letters*, vol. 2, no. 4, pp. 2127-2134, 2017.

[5] K. Khosoussi, S. Huang, and G. Dissanayak, "Tree-Connectivity: Evaluating the Graphical Structure of SLAM," in *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, 2016, pp. 1316-1322.

[6] K. Khosoussi, M. Giamou, G. S. Sukhatme, S. Huang, G. Dissanayake, and J. P. How, Reliable Graph Topologies for SLAM, *The International Journal of Robotics Research*, 2017. http://www.mit.edu/ mr-robot/filez/ijrr17.pdf

[7] K. Khosoussi, S. Huang, and G. Dissanayake, "Novel insights into the impact of graph structure on SLAM," in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2014, pp. 2707-2714.

[8] H. Cramr, (1946). Mathematical Methods of Statistics. Princeton, NJ: Princeton Univ. Press.

[9] S. Wenhardt, B. Deutsch, E. Angelopoulou, and H. Niemann, "Active visual object reconstruction using D-, E-, and T-optimal next best views," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2007, pp. 1-7.

[10] X. Shen, and P.K. Varshney, "Sensor selection based on generalized information gain for target tracking in large sensor networks," *IEEE Transactions on Signal Processing*, vol. 62, no. 2, 2014, pp. 363-375.

[11] S. Liu, S.P. Chepuri, M. Fardad, E. Maazade, G. Leus, and P.K. Varshney, "Sensor selection for estimation with correlated measurement noise," *IEEE Transactions on Signal Processing*, vol. 64, no. 13, 2016, pp. 3509-3522.

[12] R. A. Bailey and P. J. Cameron, "Combinatorics of optimal designs," *Surveys in Combinatorics*, vol. 365, 2009, pp. 19-73.

[13] R. Kmmerle, G. Grisetti, H. Strasdat, K. Konolige, and W. Burgard. "g2o: A general framework for graph optimization," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2011, pp. 3607-3613.

[14] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J.J. Leonard, and F. Dellaert. "iSAM2: Incremental smoothing and mapping using the Bayes tree," *The International Journal of Robotics Research*, vol. 31, no. 2, pp. 216-235, 2012.

[15] V. Ila, L. Polok, M. Solony, and P. Svoboda. "SLAM++-A highly efficient and temporally scalable incremental SLAM framework," *The International Journal of Robotics Research*, vol. 36, no. 2, pp. 210-230, 2017.

[16] S. Agarwal and K. Mierle and Others, Ceres Solver, http://ceres-solver.org.

[17] L. Carlone, G.C. Calafiore, C. Tommolillo, and F. Dellaert. "Planar pose graph optimization: Duality, optimal solutions, and verification," *IEEE Transactions on Robotics*, vol. 32, no. 3, pp. 545-565, 2016.

[18] O. Ozyesil, N. Sharon, and A. Singer. "Synchronization over Cartan motion groups via contraction," *SIAM Journal on Applied Algebra and Geometry*, vol. 2, no. 2, pp. 207-241, 2018.

[19] L. Carlone, G. C. Calafiore. "Convex Relaxations for Pose Graph Optimization With Outliers," *IEEE Robotics and Automation Letters*, Vol. 3, no. 2, pp. 1160-1167, 2018.

[20] J. Briales, and J. Gonzalez-Jimenez, "Convex global 3D registration with Lagrangian duality," in *International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[21] O. Ozyesil, and A. Singer, "Robust camera location estimation by convex programming," *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 2674-2683.

[22] S.M. Ross, (2014). Introduction to probability models. Academic press.

[23] N. Boumal, A. Singer, P. A. Absil, and V. D. Blondel, "CramérRao bounds for synchronization of rotations," *Information and Inference: A Journal of the IMA*, vol. 3, no. 1, pp. 1-39, 2013.

[24] R. C. Rao, "Information and accuracy attainable in the estimation of statistical parameters, *Breakthroughs in statistics*, vol. 37, pp. 8191, 1945.

[25] S. Smith, "Covariance, subspace, and intrinsic Cramér-Rao bounds, *IEEE Transactions Signal Process*, vol. 53, no. 5, pp. 1610-1630, 2005.

[26] J. Xavier and V. Barroso, "Intrinsic Variance Lower Bound (IVLB): an extension of the Cramr-Rao bound to Riemannian manifolds," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2005, pp. 1033-1036.

[27] P. Vakili, H. Mirzaei, S. Zarbafian, I.C. Paschalidis, D. Kozakov, and S. Vajda, "Optimization on the space of rigid and flexible motions: an alternative manifold optimization approach," in *IEEE Annual Conference on Decision and Control (CDC)*, 2014, pp. 5825-5830.

[28] O. Brüls, M. Arnold and A. Cardona, "Two Lie group formulations for dynamic multibody systems with large rotations," in *ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 2011, pp. 85-94.

[29] H. Carrillo, Y. Latif, M. L. Rodr´guez, J. Neira, and J. A. Castellanos, "On the Monotonicity of Optimality Criteria during Exploration in Active SLAM," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 1476-1483.

[30] H. Carrillo, I. Reid, and J. A. Castellanos, "On the Comparison of Uncertainty Criteria for Active SLAM," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2012, pp. 2080-2087.

[31] M. L. Rodríguez-Arévalo, J. Néira, and J.A. Castellanos. "On the Importance of Uncertainty Representation in Active SLAM," *IEEE Transactions on Robotics*, 2018.

[32] T. D. Barfoot, and P.T. Furgale. "Associating uncertainty with three-dimensional poses for use in estimation problems," *IEEE Transactions on Robotics*, vol. 30, no.3, pp. 679-693, 2014.

[33] D. Kopitkov, V. Indelman, "No belief propagation required: Belief space planning in high-dimensional state spaces via factor graphs, the matrix determinant lemma, and re-use of calculation," *The International Journal of Robotics Research*, vol. 36, no. 10, pp. 1088-1130, 2017.

[34] Wolfram (2001) Modified Bessel function of the first kind: Integral representations. http://functions.wolfram.com/03.02.07.0007.01.

[35] E. Eade (2011). Lie groups for 2d and 3d transformations, http://ethaneade.com/lie.pdf, revised May. 2017.

[36] M. P. Do Carmo (1992). Riemannian geometry. Birkhauser.

[37] J. R. Magnus and H. Neudecker. Matrix differential calculus with applications in statistics and econometrics. Wiley series in probability and mathematical statistics, 1988.

[38] D. A. Harville. Matrix algebra from a statistician's perspective, Springer, 1997.

[39] S. Chaiken, and D. J. Kleitman. "Matrix tree theorems," *Journal of combinatorial theory*, Series A, vol. 24, no. 3, pp. 377-381, 1978.

[40] C. Lanczos,"An iteration method for the solution of the eigenvalue problem of linear differential and integral operators," *Journal of Research of the National Bureau of Standards*, vol. 45, no. 4, 1950.

[41] E. S. Coakley, and V. Rokhlin. "A fast divide-and-conquer algorithm for computing the spectra of real symmetric tridiagonal matrices," *Applied and Computational Harmonic Analysis*, vol. 34, no. 3, pp. 379-414, 2013.

[42] Y. Chen, S. Huang, R. Fitch, and J. Yu, "Efficient Active SLAM based on Submap Joining, Graph Topology and Convex Optimization," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2018. https://github.com/cyb1212/Conference-paper.git

[43] Huang, S., Kwok, N.M., Dissanayake, G., Ha, Q.P. and Fang, G., "Multi-step look-ahead trajectory planning in SLAM: Possibility and necessity," in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 1091-1096, 2005.

[44] M. Giamou, K. Khosoussi, and J.P. How, "Talk Resource-Efficiently to Me: Optimal Communication Planning for Distributed SLAM Front-Ends," arXiv preprint arXiv:1709.06675, 2017.

[45] R. Valencia, M. Morta, J. Andrade-Cetto, and J.M. Porta. "Planning reliable paths with pose SLAM," *IEEE Transactions on Robotics*, vol. 29, no. 4, pp.1050-1059, 2013.

[46] B. D. Flury. "Acceptance-rejection sampling made easy," *SIAM Review*, vol. 32, no.3, 1990, pp. 474-476.

[47] D. Bump (2004) Lie Groups. Graduate Texts in Mathematics, vol. 225. Berlin: Springer.