

# Cramér-Rao bounds and optimal design metrics for pose-graph SLAM: Technical report

Yongbo Chen<sup>†</sup>, Shoudong Huang<sup>†</sup>, Liang Zhao<sup>†</sup>, and Gamini Dissanayake<sup>†</sup>

<sup>†</sup> Centre for Autonomous Systems (CAS), University of Technology Sydney, Australia

Yongbo.Chen@student.uts.edu.au, Shoudong.Huang@uts.edu.au,  
Liang.Zhao@uts.edu.au, Gamini.Dissanayake@uts.edu.au

January 22, 2020

This document provides more detailed material to the paper [1]. It can be considered as a self-contained document and have everything in [1]. It can be cited as:

*“Y. Chen, S. Huang, L. Zhao, and G. Dissanayake, “Cramér-Rao bounds and optimal design metrics for pose-graph SLAM: Technical report,” 2020. <https://github.com/cyb1212/A-submittedpaper>”*

Throughout this document, notations used are the same as those in [1].

## 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 report discusses about the Fisher information matrix (FIM), Cramér-Rao bounds (CRBs) and optimal design metrics (T-optimality and D-optimality) of synchronization of  $\mathbb{R}^n \times SO(n)$ , which is a standard 2/3D pose-graph SLAM problem. The FIM tells us about the “quality” of the SLAM result (i.e., how accurate it is compared to the ground-truth). We start from the synchronization problem on  $\mathbb{R}^1$  and  $\mathbb{R}^2$  (linear cases), then talk about the synchronization on  $SO(2)$ ,  $SO(3)$ ,  $\mathbb{R}^2 \times SO(2)$ , and  $\mathbb{R}^3 \times SO(3)$ , respectively. The FIM is shown to be closely related to the graph structure, in particular, the weighted Laplacian matrix. The CRBs is a classical tool in estimation method that provide a lower bound on the variance of any unbiased estimator for an estimation problem. Because of the non-flatness of the manifold  $\mathbb{R}^3 \times SO(3)$ , a new form  $C \succeq F^{-1} + curvature\ terms$  is used to compute the CRLBs. We show that the curvature terms can be negligible at large signal-to-noise ratios (SNR). We also prove that total node degree and weighted number of spanning trees, as two graph connectivity metrics, are respectively closely related to the trace and determinant of the FIM. The discussions show that, compared with the D-optimality, the T-optimality metric is more easily computed but less effective. 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 results are verified with several wellknown datasets, such as Intel, KITTI, sphere and so on.

## Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
1.1	Contributions of this report . . . . .	4
1.2	Outline . . . . .	5
1.3	Notations . . . . .	5
<b>2</b>	<b>Related work</b>	<b>5</b>
<b>3</b>	<b>Synchronization problem (pose-graph SLAM)</b>	<b>7</b>
3.1	Graph preliminaries for pose-graph SLAM . . . . .	7
3.2	The general synchronization problem and definition of FIM . . . . .	8
3.3	Geometry of the parameter spaces . . . . .	8
3.3.1	Tangent space on $SO(n)$ [38] . . . . .	8
3.3.2	Inner product on $\mathbb{R}^n$ and $SO(n)$ [13] . . . . .	8
3.3.3	Gradient on $SO(n)$ [40] . . . . .	9

<b>4 Synchronization on <math>\mathbb{R}^1</math></b>	<b>9</b>
4.1 Synchronization on $\mathbb{R}^1$	9
4.2 FIM for $\mathbb{R}^1$	10
4.3 Discussion about the directed graph and undirected graph	13
4.4 FIM using Definition in [39] under Gaussian distribution	13
<b>5 Synchronization on <math>\mathbb{R}^2</math></b>	<b>14</b>
5.1 Synchronization on $\mathbb{R}^2$	14
5.2 FIM for $\mathbb{R}^2$	15
<b>6 Synchronization on <math>SO(2)</math></b>	<b>17</b>
6.1 Discussion about isotropic Langevin distribution	17
6.2 Synchronization on $SO(2)$	17
6.3 FIM for $SO(2)$	18
<b>7 Synchronization on <math>SO(3)</math></b>	<b>21</b>
7.1 Synchronization on $SO(3)$	21
7.2 FIM for $SO(3)$	22
<b>8 Synchronization on <math>\mathbb{R}^n \times SO(n)</math></b>	<b>28</b>
<b>9 2D pose-graph SLAM</b>	<b>29</b>
9.1 Synchronization on $\mathbb{R}^2 \times SO(2)$	29
9.2 Geometry of the parameter space	29
9.3 FIM for 2D pose-graph SLAM	30
9.3.1 $\mathbb{R}^2$ sub-matrix	30
9.3.2 $SO(2)$ by $\mathbb{R}^2$ coupling sub-matrix	32
9.3.3 $\mathbb{R}^2$ by $SO(2)$ coupling sub-matrix	35
9.3.4 $SO(2)$ sub-matrix	35
9.3.5 The whole FIM matrix for 2D pose graph	36
<b>10 3D pose-graph SLAM</b>	<b>37</b>
10.1 Synchronization on $\mathbb{R}^3 \times SO(3)$	37
10.2 Geometry of the parameter space	37
10.2.1 Tangent space	37
10.2.2 Orthonormal basis	38
10.2.3 Gradient	38
10.3 FIM for 3D pose-graph SLAM	38
10.3.1 $\mathbb{R}^3$ sub-matrix	39
10.3.2 $SO(3)$ by $\mathbb{R}^3$ coupling sub-matrix	39
10.3.3 $\mathbb{R}^3$ by $SO(3)$ coupling sub-matrix	42
10.3.4 $SO(3)$ sub-matrix	43
10.3.5 The whole FIM matrix for 3D pose graph	45
<b>11 CRLB for pose-graph SLAM</b>	<b>45</b>
11.1 CRLB for 2D pose-graph SLAM	46
11.2 CRLB for 3D pose-graph SLAM	46
<b>12 Optimal Experimental Design</b>	<b>47</b>
12.1 T-optimality design metric	47
12.1.1 T-optimality design metric for the synchronization on $\mathbb{R}^2 \times SO(2)$	47
12.1.2 T-optimality design metric for the synchronization on $\mathbb{R}^3 \times SO(3)$	47
12.1.3 Further analysis	48
12.2 D-optimality design metric	48
12.3 Discussion and comparison	49
12.3.1 Efficiency of the metric	49
12.3.2 Computational complexity	51

<b>13 Code and simulation for D-optimality bound</b>	<b>51</b>
13.1 Code . . . . .	51
13.2 Load dataset . . . . .	52
13.3 Construct FIM . . . . .	52
13.4 Compute log-determinant function and Obtain bounds . . . . .	52
13.5 Obtain ground truth . . . . .	53
13.6 Sample noise . . . . .	53
13.7 Obtain noisy measurement . . . . .	53
13.8 Obtain results using SE(3)-sycn . . . . .	53
13.9 Obtain covariance based on statistical way . . . . .	53
13.10 Compute CRLB based on FIM . . . . .	53
13.11 Simulation setting . . . . .	54
13.12 Relation between T-/D-optimality metrics with graph topology . . . . .	54
13.13 T-/D-optimality metrics in active SLAM application . . . . .	56
13.14 Bound efficiency on D-optimality metric . . . . .	56
13.15 Efficiency of CRLB . . . . .	58

# 1 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 [2]. Multiple estimation problems, including pose-graph SLAM, fall into this category [3]. These synchronization problems in general give rise to a weighted graph representation. In essence, there is 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, one question to ask is how *reliable* that solution will be (i.e., accuracy compared to ground-truth and the uncertainty of the estimated solution). Both the covariance matrix and the FIM can be used to assess the uncertainty of the estimated parameters. However, compared with the FIM, updating/storing the dense full covariance matrix is prohibitively expensive with the growth of the dimension of the state vector. Thus, the FIM is the top-priority choice to evaluate the uncertainty of the SLAM solution in the maximum likelihood (ML) estimate. Under the assumption that the rotation noises obeying the isotropic Langevin distribution on  $SO(2)$  and  $SO(3)$  and using the Frobenius norm based distance, the state-of-the-art algorithms of SLAM, including SE-Sync [4] and Cartan-Sync [5], show outstanding computational efficiency (more than an order of magnitude faster) compared with the Gauss-Newton (GN) based approach. Nevertheless, the corresponding FIM is not presented in the literature.

From the graphical point of view, we know that adding relative measurements among 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 proved to be closely related to the graph structure of the measurements network, in particular, the weighted Laplacian matrix [6] [7] assuming Gaussian noise on the relative pose orientation and ignoring the wraparound issue. In this paper, as the first contribution, we derive the FIM based on the assumption of zero-mean isotropic Langevin noise for orientation and block-isotropic Gaussian noise for translation in 2D and 3D pose-graph SLAM.

It is known that, in a flat Euclidean space, the classical CRLB result for any unbiased estimator provides us with a simple but strong relation between the covariance matrix  $\mathbf{C}$  and the FIM  $\mathbf{F}$ :  $\mathbf{C} \succeq \mathbf{F}^{-1}$  [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 need to be considered to derive the rigorous CRLB. As the second contribution, we derive the CRLB for 2D and 3D pose-graph SLAM based on isotropic Langevin noise for orientation and block-isotropic Gaussian noise for translation.

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 [3]. For example, in [9], the D-optimality, E-optimality, and T-optimality metrics are used in 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 an expression related to the weighted number of spanning trees [7]. In this paper, we extend the results in [7] 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.

## 1.1 Contributions of this report

The main contributions of this paper are listed below.

- Extension of the formula derivation of the FIM and the CRLB of the synchronization problem in [13] from the rotation group  $SO(n)$  to the group of rigid body motions,  $\mathbb{R}^2 \times SO(2)$  and  $\mathbb{R}^3 \times SO(3)$ , based on the assumption of isotropic Langevin noise (for rotation) and block-isotropic Gaussian noise (for translation);
- Derivation of the relationship between the FIM of 3D pose-graph SLAM and the weighted Laplacian matrix;
- Extension of the analysis results of the D-optimality metric in [6] and [7] from 2D pose-graph SLAM into 3D case with  $\mathbb{R}^3 \times SO(3)$  relative-pose measurements;

- Comparison of the D-optimality and T-optimality metrics of 2D and 3D pose-graph SLAM from the graphical perspective.

In order to make the report more self-contained and easy to understand, we have included the simple linear cases ( $\mathbb{R}^1$  and  $\mathbb{R}^2$ ) and the Lie group ( $SO(2)$  and  $SO(3)$ ) as well.

## 1.2 Outline

In Section 2, we review the related works about the pose-graph SLAM, the FIM, the CRLB and the TOED. In Section 3, the preliminaries of the synchronization problem are presented and defined. In Section 4 and Section 5, we present the mathematical formulations of the synchronization on the Euclidean space,  $\mathbb{R}^n$ ,  $n = 1, 2$  and their FIMs. In Section 6.1, we provide some discussions about isotropic Langevin distribution. The mathematical formulations of the synchronization on the Lie group,  $SO(n)$ ,  $n = 2, 3$  and their FIMs are shown in Section 6 and 7. In Section 9 and 10, the group of rigid body motions in 2D plane and 3D space,  $\mathbb{R}^n \times SO(n)$ ,  $n = 2, 3$  (pose-graph SLAM), their corresponding FIMs are presented to show its strong relationship with the weighted Laplacian matrix of the measurement graph.

in Section V. In Section VI, the CRLB of 2D and 3D pose-graph SLAM are presented. The TOED metrics, including D-optimality and T-optimality, of the FIM of 3D pose-graph SLAM are discussed focusing on the graph topology in Section VII. Results about the optimality metrics and the CRLB using different publicly available datasets are given in Section VIII. Conclusions, potential applications and future work are presented in Section IX. To improve the readability of the paper, some details about the derivations and proofs are provided in the appendices.

## 1.3 Notations

Throughout this paper, unless otherwise noted, bold lowercase and bold uppercase letters are reserved for vectors and matrices, respectively. Sets are shown by uppercase letters.  $SO(n)$  (special orthogonal group) and  $\bar{O}(n)$  (orthogonal group) are respectively defined as:  $SO(n) \triangleq \{\mathbf{R} \in \mathbb{R}^{n \times n} : \mathbf{R}^\top \mathbf{R} = \mathbf{I}_{n \times n}, \det(\mathbf{R}) = 1\}$  and  $\bar{O}(n) \triangleq \{\mathbf{G} \in \mathbb{R}^{n \times n} : \mathbf{G}^\top \mathbf{G} = \mathbf{I}_{n \times n}\}$ .  $\mathbf{S}_1 \succeq \mathbf{S}_2$  means matrix  $\mathbf{S}_1 - \mathbf{S}_2$  is positive semidefinite. The Kronecker product is denoted by  $\otimes$ .  $\text{trace}(\star)$  and  $\det(\star)$  represent the trace and determinant of the matrix  $\star$ . We denote by  $\text{diag}(\mathbf{M}_1, \dots, \mathbf{M}_k)$  the block-diagonal matrix with matrices  $\mathbf{M}_1, \dots, \mathbf{M}_k$  as blocks on its main diagonal. The squared Frobenius norm is  $\|\star\|_F^2 = \text{trace}(\star^\top \cdot \star)$ .  $\|\star\|_{eig}$  means the biggest eigenvalue of the matrix  $\star$ ; For a symmetric positive definite matrix  $\star$ ,  $\|\star\|_{eig} = \|\star\|_2$ .  $\text{dist}(\star, \bullet) = \|\log(\star^\top \cdot \bullet)\|_F$  is the geodesic distance between  $\star$  and  $\bullet$  in  $SO(n)$ . The squared vector norm is  $\|\star\|_2^2 = \star^\top \cdot \star$ , for a vector  $\star$ .  $|\star|$  means the cardinality of the set  $\star$ .  $\mathbb{E}\{\star\}$  means the mathematical expectation of  $\star$ .  $\nabla_\star \bullet$  means the partial derivative of a function  $\bullet$  with respect to parameter  $\star$ .  $\star \ltimes \bullet$  and  $\star \times \bullet$  respectively mean the semi-product group and the direct product group of the group  $\star$  and the group  $\bullet$ .  $\text{ddiag}(\star)$  sets all off-diagonal entries of a square matrix  $\star$  to zero.

## 2 Related work

Pose-graph SLAM leads to a non-convex optimization problem, whose (globally optimal) solution is the ML estimate for the unknown poses. It is well-known that SLAM back-end methods roughly fall into two categories. The first one is to use high-efficiency iterative nonlinear optimization methods based on GN method, Levenberg-Marquardt (LM) method and Powell's dogleg method, such as: g2o [14], iSAM2 [15], SLAM++ [16] and ceres [17], to obtain locally optimal solutions. Because of its non-convex property, starting from a poor initial guess, the iterative techniques may be trapped into a local minimum, which corresponds to a wrong estimate. The other one is to compute globally optimal solutions via convex relaxations [4] [18] [19]. Some works have shown that the duality gap of the general pose-graph SLAM problem in practical applications is close to zero [20], which implies that it can be solved exactly via convex relaxations. Beside pose-graph SLAM, the similar technologies have been widely used in other robotics and computer vision applications, such as: 3D registration [21] and structure from motion (SfM) [22].

Although many efficient optimization algorithms have been developed, the achievable estimated uncertainty is not well studied. Boumal [2] proposes the FIM for the estimation problems when the actual parameter space is a Riemannian sub-manifolds or a Riemannian quotient manifold. In [2], Boumal also shows two simple examples based on isotropic Gaussian noise: synchronization on the group of translation  $\mathbb{R}^n$  and synchronization on the group of rotation  $SO(3)$ . In his later work [13], the conclusions are extended to the general rotation group  $SO(n)$  based on several kinds of Gaussian-like, but

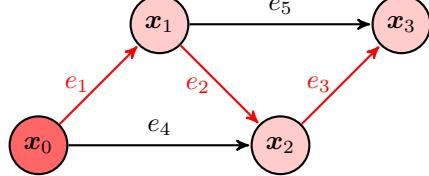
non-Gaussian, noises, especially for the isotropic Langevin noise, which has attracted some robotics researchers' attention. As the state-of-the-art back-end algorithms, the convex relaxation based SE-Sync [4] and Cartan-Sync methods [5] are built based on the assumption of isotropic Gaussian noise (for translation) and isotropic Langevin noise (for rotation). To the best of our knowledge, the FIM for pose-graph SLAM based on these noises, whose parameter space is the product manifold  $\mathbb{R}^n \times SO(n)$ ,  $n = 2, 3$ , has not 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 [13]. The traditional CRLB is defined in a flat Euclidean space. Smith [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 the anchor-free case. Based on these new extended tools, Boumal presents the CRLB for the synchronization of rotations  $SO(n)$  in both the anchored and the anchor-free cases. 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 of [13], we will present its CRLB in this paper.

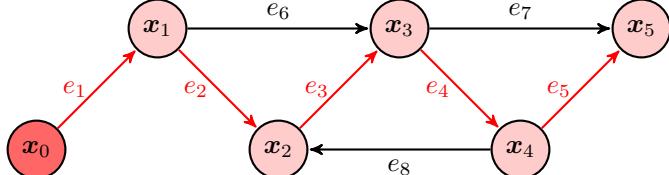
In fact, there are two different Lie group representations corresponding to 2D/3D pose-graph SLAM:  $\mathbb{R}^n \times SO(n)$  and  $SE(n) = \mathbb{R}^n \times SO(n)$ . The direct product of  $SO(n)$  and  $\mathbb{R}^n$  manifolds  $\mathbb{R}^n \times SO(n)$  can be represented as a  $(2n+1) \times (2n+1)$  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 (semi-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  $SE(3)$ , the direct product group  $\mathbb{R}^n \times SO(n)$  can keep the bi-invariance property by the simple combined Riemannian metric, which is defined as the sum of the metrics of  $\mathbb{R}^n$  and  $SO(n)$ . However, the manifold  $SE(3)$  does not have any bi-invariant metric, which results in more complicated formulations for FIM, CRLB and curvature terms. Meanwhile, many recent popular SLAM pose-graph optimization methods, such as SE-sync method and initialization techniques [29], separate the SLAM problem into two parts: the rotation estimation and the linear least squares estimation for translation. Their objective functions are built based on the direct product group  $\mathbb{R}^n \times SO(n)$ . Hence, in this paper, we only consider the pose-graph SLAM problem defined 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 [30] [31] [32]. In recent work, the authors show that the monotonicity of all optimality criteria, A-optimality, D-optimality, and E-optimality, and Shannon's entropy is greatly affected by the uncertainty representation [33]. As a common representation, the uncertainty representations on the Lie group  $SE(3)$  based on the Gaussian noise have also been considered [34]. In these metrics, the D-optimality metric and T-optimality metric do not need to perform the inverse operation for the FIM, which leads to a lower computational complexity. The D-optimality metric is the most popular metric with a good performance [30]. However, from a computational complexity perspective, the T-optimality metric still has its obvious advantage compared with the D-optimality metric, even though the state-of-the-art incremental technology could be used for computing the D-optimality metric [35]. So both metrics have great significance to the different estimation and planning requirements. We will investigate both two metrics based on the graphical structure of the SLAM.

In the previous work [6] [7] and [36], we analyze the impact of the graphical structure on some of the desirable attributes of some estimation problems: linear sensor network (SN), compass-SLAM and 2D pose-graph SLAM with block-isotropic Gaussian noise. In linear-SN and 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 structure of the measurement graph. For 2D pose-graph SLAM with the block-isotropic Gaussian noises, it is stated that the D-optimality metric of the FIM can be bounded by an expression related to the weighted number of spanning trees of the measurement graph (weighted tree-connectivity). Based on the lower bounds, a new near- $t$ -optimal graph synthesis framework is put forward for the measurement selection, pose-graph pruning problems and D-optimality-aware SLAM front-end. In this paper, we extend our analysis and conclusions [6, 7, 36] into 2D and 3D pose-SLAM with  $\mathbb{R}^n \times SO(n)$ ,  $n = 2, 3$  relative-pose measurements.



(a)  $\mathcal{G}_1$  (pose-graph 1).



(b)  $\mathcal{G}_2$  (pose-graph 2).

Figure 1: Two examples of pose-graphs (with  $x_0$  as anchor).

### 3 Synchronization problem (pose-graph SLAM)

#### 3.1 Graph preliminaries for pose-graph SLAM

A directed graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , which is weakly connected, is used to represent pose-graph SLAM problem naturally. In this paper,  $\mathcal{V} = \{1, 2, \dots, n_p\} \cup \{0\}$ ,  $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$  and  $|\mathcal{E}| = m$ . Each node corresponds to a robot pose, and each edge  $(i, j) \in \mathcal{E}$  corresponds to a relative measurement from robot pose  $i$  to robot pose  $j$ . A new undirected rotation graph  $\mathcal{G}_1 = (\mathcal{V}_1, \mathcal{F})$ , whose nodes only represent the rotations of the poses and edges mean the relative rotation measurements, is created. It is similar to the pose graph  $\mathcal{G}$ . The rotation graph  $\mathcal{G}_1$  is un-directed, because the relative rotation measurements  $(i, j)$  introduce the same information for  $i$ -th node and  $j$ -th node.

For the  $i$ -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  $V_i = V_i^- \cup V_i^+$ , so we have  $|V_i^+| + |V_i^-| = |V_i| = d_i$ , where  $d_i$  is the  $i$ -th node degree. Without loss of generality, the first pose (corresponding to 0-th node) is assumed to be the origin of our global coordinates system.

The incidence matrix of  $\mathcal{G}$  is denoted by  $\mathbf{A}_0 \in \{-1, 0, 1\}^{(n_p+1) \times m}$ .  $a_{ik} = -1$  and  $a_{jk} = 1$  (the  $(i, k)$ -th and  $(j, k)$ -th element of  $\mathbf{A}_0$ ,  $k$  means the  $k$ -th measurement (edge)) are non-zero, if the  $k$ -th edge is  $e_k = (i, j) \in \mathcal{E}$ . The incidence matrix after anchoring to the origin,  $\mathbf{A} \in \{-1, 0, 1\}^{n_p \times m}$ , is obtained simply by removing the row corresponding to the first node in  $\mathbf{A}_0$ . The Laplacian matrix and the reduced Laplacian matrix of  $\mathcal{G}$  are respectively defined as  $\mathbf{L}_0 \triangleq \mathbf{A}_0 \mathbf{A}_0^\top$  and  $\mathbf{L} \triangleq \mathbf{A} \mathbf{A}^\top$ . It can be shown that  $\mathbf{L}_0$  and  $\mathbf{L}$  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  $\mathbf{L}_0 = \mathbf{D}_0 - \mathbf{W}_0$  and  $\mathbf{L} = \mathbf{D} - \mathbf{W}$ , where  $\mathbf{D}_0 \triangleq \text{diag}(d_0, d_1, \dots, d_n)$ ,  $\mathbf{D} \triangleq \text{diag}(d_1, \dots, d_n)$ ,  $\mathbf{W}_0$  and  $\mathbf{W}$  are respectively the original adjacency matrix and the adjacency matrix after removing the row and column corresponding to node 0. The weighted Laplacian matrix and the weighted reduced Laplacian matrix are defined as  $\mathbf{L}_\omega^0 \triangleq \mathbf{A}_0 \Sigma_0 \mathbf{A}_0^\top$  and  $\mathbf{L}_\omega \triangleq \mathbf{A} \Sigma \mathbf{A}^\top$ , where  $\Sigma_0$  and  $\Sigma$  are diagonal matrices whose diagonal elements are the weight values of the graph edges.

The weighted Laplacian matrix and reduced weighted Laplacian matrix for the two graphs in Figure 1 are:

$$\mathbf{L}_{w0}^{\mathcal{G}_1} \triangleq \mathbf{A}_0^{\mathcal{G}_1} \Sigma_0^1 \mathbf{A}_0^{\mathcal{G}_1 \top}, \quad \mathbf{L}_{w0}^{\mathcal{G}_2} \triangleq \mathbf{A}_0^{\mathcal{G}_2} \Sigma_0^2 \mathbf{A}_0^{\mathcal{G}_2 \top}, \quad (1)$$

$$\mathbf{L}^{\mathcal{G}_1} \triangleq \mathbf{A}_{\mathcal{G}_1} \Sigma^1 \mathbf{A}_{\mathcal{G}_1}^\top, \quad \mathbf{L}^{\mathcal{G}_2} \triangleq \mathbf{A}_{\mathcal{G}_2} \Sigma^2 \mathbf{A}_{\mathcal{G}_2}^\top, \quad (2)$$

where  $\Sigma_0^1 = \text{diag}\{\omega_0^1, \omega_1^1, \dots, \omega_5^1\}$ ,  $\Sigma^1 = \text{diag}\{\omega_1^1, \dots, \omega_5^1\}$ ,  $\Sigma_0^2 = \text{diag}\{\omega_0^2, \omega_1^2, \dots, \omega_8^2\}$ ,  $\Sigma^2 = \text{diag}\{\omega_1^2, \dots, \omega_8^2\}$ ;  $\omega_i$  is the weight value for  $i$ -th edge;  $\mathbf{A}_0^{\mathcal{G}_1}$  and  $\mathbf{A}_{\mathcal{G}_1}$  are the incident matrix and the reduced incident matrix of the graph  $\mathcal{G}_1$ ;  $\mathbf{A}_0^{\mathcal{G}_2}$  and  $\mathbf{A}_{\mathcal{G}_2}$  are the incident matrix and the reduced incident matrix of the graph  $\mathcal{G}_2$ .

## 3.2 The general synchronization problem and definition of FIM

Now we introduce the general synchronization problem and the definition of FIM. In general, synchronization on space  $\mathcal{P}$  is the problem of estimating a set of variables  $X_0, X_1, X_2, \dots, X_{n_p} \in \mathcal{P}$  from noisy measurements  $H_{ij}$  of relative information between  $X_i$  and  $X_j$  (edges in the graph,  $0 \leq i, j \leq n_p$ ).

The noisy measurements  $H_{ij}$  can be expressed by

$$H_{ij} = Z_{ij} \oplus (X_j \ominus X_i) \quad (3)$$

where  $Z_{ij}$  is a random variable distributed over  $\mathcal{P}'$  following a PDF  $f_{ij} : \mathcal{P}' \rightarrow \mathbb{R}^+$ . The PDF  $f_{ij}$  depends on the space  $\mathcal{P}'$ , for  $\mathcal{P}' = \mathbb{R}^1$  or  $\mathcal{P}' = \mathbb{R}^2$ , we assume isotropic Gaussian with zero-mean; for  $\mathcal{P}' = SO(3)$ , we assume the isotropic Langevin distribution.  $X_j \ominus X_i$  means the relative information between  $X_i$  and  $X_j$ , the exact meanings of  $\ominus$  and  $\oplus$  also depend on  $\mathcal{P}'$ .

**Definition:** The log-likelihood of the estimator  $\boldsymbol{\theta} = (X_0, \dots, X_{n_p})$ , given the measurements  $H_{ij}$ , is given by:

$$L(\mathbf{y}; \boldsymbol{\theta}) = \sum_{(i,j) \in \mathcal{E}} \log f_{ij}(H_{ij} \ominus (X_j \ominus X_i)). \quad (4)$$

The problem of synchronization on space  $\mathcal{P} = \{\mathcal{P}' \times \dots \times \mathcal{P}'\}_{n_p}$  is to find  $\boldsymbol{\theta}$  such that the log-likelihood  $L(\mathbf{y}; \boldsymbol{\theta})$  is maximized.

**Definition [13]:**  $\boldsymbol{\theta} \in \mathcal{P}'$  be unknown parameter and  $f(\mathbf{y}; \boldsymbol{\theta})$  be the PDF of the measurement  $\mathbf{y}$  conditioned by  $\boldsymbol{\theta}$ . Based on the log-likelihood function  $L(\mathbf{y}; \boldsymbol{\theta}) = \log f(\mathbf{y}; \boldsymbol{\theta})$  shown in (4) and the orthonormal basis, the  $(i, j)$ -th element of the FIM is defined as:

$$\mathbf{F}_{(i,j)} = \mathbb{E} \{ \langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), e_i \rangle \cdot \langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), e_j \rangle \}, \quad (5)$$

where  $e_i$  and  $e_j$  are the  $i$ -th and  $j$ -th bases of the tangent space of the parameters. Expectations are taken w.r.t. the measurement  $\mathbf{y}$ . They will be defined based on the parameter space [2]. It is noted that the FIM is directly decided by the bases. For a parameter space, there may exist more than one kind of bases. The inner product  $\langle \bullet, \star \rangle = \langle \bullet, \star \rangle_{\mathcal{P}'}$  is defined based on the parameter space  $\mathcal{P}'$ .

It is noted that, for the anchored situation, we only need to set the anchored nodes as the constant values. The formulations are similar.

Now we will discuss the synchronization on different space  $\mathcal{P}'$  and the corresponding FIM.

## 3.3 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.

### 3.3.1 Tangent space on $SO(n)$ [38]

As a Lie group, the dimension of its minimal representation is  $d = d(n) = \frac{n(n-1)}{2}$  ( $d = 1$  and  $3$ , for 2D and 3D case). We can admit a tangent space  $\mathcal{T}_Q SO(n)$ ,  $n = 2, 3$  for each rotation:

$$\mathcal{T}_Q SO(n) = Q so(n) \triangleq \{Q\Omega : \Omega \in \mathbb{R}^{n \times n}, \Omega^\top + \Omega = 0\}, \quad (6)$$

where  $Q \in SO(n)$ ,  $so(n)$  is the Lie algebra corresponding to  $SO(n)$ .

### 3.3.2 Inner product on $\mathbb{R}^n$ and $SO(n)$ [13]

Based on 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:

$$\begin{cases} \langle \rho_1, \rho_2 \rangle_x = \rho_1^\top \rho_2 & \rho_1, \rho_2 \in \mathbb{R}^n \\ \langle \Omega_1, \Omega_2 \rangle_R = \text{trace}(\Omega_1^\top \Omega_2) & \Omega_1, \Omega_2 \in \mathcal{T}_Q SO(n). \end{cases} \quad (7)$$

Because  $(Q\Omega_1)^\top Q\Omega_2 = \Omega_1^\top \Omega_2$ , we omit the tangent subscripts  $Q$  in above equation and all related inner product equations of this paper, for better readability. The Riemannian metric of  $\mathbb{R}^n \times SO(n)$  is the sum of the Riemannian metrics of  $\mathbb{R}^n$  and  $SO(n)$ , which helps to maintain the bi-invariance property.  $\forall (\rho, Q) \in \mathbb{R}^n \times SO(n)$ ,  $\forall (\rho_i, \Omega_i) \in \mathcal{T}_{(\rho, Q)}(\mathbb{R}^n \times SO(n))$ ,  $i = 1, 2$ , the specific expression of the Riemannian metric of  $\mathbb{R}^n \times SO(n)$  is defined as:

$$\langle (\rho_1, \Omega_1), (\rho_2, \Omega_2) \rangle := \langle \rho_1, \rho_2 \rangle_x + \langle \Omega_1, \Omega_2 \rangle_R. \quad (8)$$

### 3.3.3 Gradient on $SO(n)$ [40]

Let  $h : SO(n) \rightarrow \mathbb{R}$  be a differentiable function, we can define the gradient of  $h$  by:

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

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

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

$$\langle \text{grad } h(\mathbf{Q}), \mathbf{Q}\Omega \rangle_{\mathbf{R}}, \quad (10)$$

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

## 4 Synchronization on $\mathbb{R}^1$

### 4.1 Synchronization on $\mathbb{R}^1$

When  $\mathcal{P}' = \mathbb{R}^1$ , the relative information between  $x_i$  and  $x_j$  is simply  $x_j \ominus x_i = x_j - x_i$  and the  $\oplus$  operation is simply the addition in  $\mathbb{R}^1$ . So the measurement  $p_{ij}$  is given by

$$p_{ij} = y_{ij} + x_j - x_i. \quad (11)$$

where  $y_{ij}$  is a zero-mean Gaussian noise in  $\mathbb{R}^1$  with variance  $\delta_{ij}^2$ . That is,

$$f_{ij}(x) = \frac{1}{\sqrt{2\pi}\delta_{ij}} e^{-\frac{x^2}{2\delta_{ij}^2}} \quad (12)$$

Hence

$$L(\mathbf{y}; \boldsymbol{\theta}) = \sum_{(i,j) \in \mathcal{E}} \log f_{ij}(p_{ij} \ominus (x_j \ominus x_i)) = \sum_{(i,j) \in \mathcal{E}} \log \left( \frac{1}{\sqrt{2\pi}\delta_{ij}} e^{-\frac{(p_{ij}-x_j+x_i)^2}{2\delta_{ij}^2}} \right) \quad (13)$$

For a given  $i$ , there are two sets connecting with  $i$ :  $V_i^+$  and  $V_i^-$ .  $V_i^+$  means the set whose directed edges come out from  $i$  and  $V_i^-$  means the set whose directed edges come into  $i$ . We have:

$$\sum_{(i,j) \in \mathcal{E}} \log f_{ij}(H_{ij} \ominus (X_j \ominus X_i)) = \sum_{j_1 \in V_i^+} \log f_{ij_1}(H_{ij_1} \ominus (X_{j_1} \ominus X_i)) + \sum_{j_2 \in V_i^-} \log f_{j_2 i}(H_{j_2 i} \ominus (X_i \ominus X_{j_2})) \quad (14)$$

Note that  $L(\mathbf{y}; \boldsymbol{\theta})$  could be simplified as

$$\begin{aligned} L(\mathbf{y}; \boldsymbol{\theta}) &= \sum_{(i,j) \in \mathcal{E}} \left( -\log(\sqrt{2\pi}\delta_{ij}) - \frac{(p_{ij}-x_j+x_i)^2}{2\delta_{ij}^2} \right) \\ &= -\sum_{(i,j) \in \mathcal{E}} \log(\sqrt{2\pi}\delta_{ij}) - \sum_{(i,j) \in \mathcal{E}} \frac{(p_{ij}-x_j+x_i)^2}{\delta_{ij}^2} \\ &= -\sum_{(i,j) \in \mathcal{E}} \log(\sqrt{2\pi}\delta_{ij}) - L_0(\mathbf{y}; \boldsymbol{\theta}) \end{aligned} \quad (15)$$

where

$$L_0(\mathbf{y}; \boldsymbol{\theta}) = \sum_{(i,j) \in \mathcal{E}} \frac{(p_{ij} - x_j + x_i)^2}{\delta_{ij}^2}. \quad (16)$$

Since  $\sum_{(i,j) \in \mathcal{E}} \log(\sqrt{2\pi}\delta_{ij})$  is a constant, it is clear that maximizing  $L(\mathbf{y}; \boldsymbol{\theta})$  is equivalent to minimizing  $L_0(\mathbf{y}; \boldsymbol{\theta})$ . The later is a standard weighted linear least squares problem.

There are two ways to compute the FIM. One is to compute FIM based on the general FIM formula Eq.(131). Another way is to use linear least squares formula.

**Example 1.** Let's consider the pose-graph 1 in Fig. 1 as a example. Its measurement will be:

$$\begin{aligned} p_{01} &= y_{01} + x_1 - x_0, \\ p_{12} &= y_{12} + x_2 - x_1, \\ p_{23} &= y_{23} + x_2 - x_0. \\ p_{02} &= y_{02} + x_2 - x_0. \\ p_{13} &= y_{13} + x_3 - x_1. \end{aligned} \quad (17)$$

Its edge set  $\mathcal{E}$  will be:

$$\mathcal{E} = \{(0, 1), (1, 2), (2, 3), (0, 2), (1, 3)\} \quad (18)$$

Its log-likelihood function will be:

$$L(\mathbf{y}; \boldsymbol{\theta}) = \log f_{01}(p_{01} - (x_1 - x_0)) + \log f_{12}(p_{12} - (x_2 - x_1)) + \log f_{23}(p_{23} - (x_3 - x_2)) + \log f_{02}(p_{02} - (x_2 - x_0)) + \log f_{13}(p_{13} - (x_3 - x_1)) \quad (19)$$

For the 2-nd node,  $V_2^-$  includes  $\{1, 0\}$  and  $V_2^+$  includes  $\{3\}$ .

## 4.2 FIM for $\mathbb{R}^1$

For the definition of FIM, when  $\mathcal{P}' = \mathbb{R}^1$ , the parameter space  $\mathcal{P}$  is  $\mathbb{R}^1 \times \cdots \times \mathbb{R}^1$ .  $\boldsymbol{\theta} = \mathbf{x} = (x_1, \dots, x_{n_p})$ ,  $y = \{p_{ij}\}_{(i,j) \in \mathcal{E}}$ . The tangent space  $T_\theta \mathcal{P}$  is still  $\mathbb{R}^1 \times \cdots \times \mathbb{R}^1$ , and the  $i$ -th element of the orthonormal basis  $\mathbf{e}_i$  is 1. The basis  $\mathbf{e}_i$  is:

$$\mathbf{E}_i^X = \left[ 0, \dots, 0, \underbrace{1}_{i-th}, 0, \dots, 0 \right]^\top, \quad (20)$$

The inner product is defined by  $\langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_i^X \rangle_{\mathbf{x}} = \text{grad}L(\mathbf{y}; \boldsymbol{\theta})^\top \cdot \mathbf{E}_i^X$

Based on the definition of FIM, we have: The  $(i, i_1)$ -th element of the information matrix is

$$\begin{aligned} F_{ii_1} &= \mathbb{E}\{\langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_i^X \rangle_{\theta} \cdot \langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_{i_1}^X \rangle_{\theta}^\top\} \\ &= \mathbb{E}\{\text{grad}_{x_i} L(\mathbf{y}; \boldsymbol{\theta}) \cdot \text{grad}_{x_{i_1}} L(\mathbf{y}; \boldsymbol{\theta})^\top\} \end{aligned} \quad (21)$$

Let's first compute the directional derivative  $\text{grad}_i L(\boldsymbol{\theta})$ , in  $\mathbb{R}$ :

$$\begin{aligned} \text{grad}_{x_i} L(\mathbf{y}; \boldsymbol{\theta}) &= \nabla_{y_{ij}} L(\mathbf{y}; \boldsymbol{\theta}) \nabla_{x_i} y_{ij} \\ &= \nabla_{y_{ij}} \sum_{(i,j) \in \mathcal{E}} \log \left( \frac{1}{\sqrt{2\pi} \delta_{ij}} e^{-\frac{(y_{ij})^2}{2\delta_{ij}^2}} \right) \nabla_{x_i} (p_{ij} - x_j + x_i) \\ &= \sum_{(i,j) \in \mathcal{E}} -2 \frac{y_{ij}}{2\delta_{ij}^2} \nabla_{x_i} y_{ij} \\ &= \sum_{(i,j) \in \mathcal{E}} -\frac{y_{ij}}{\delta_{ij}^2} \nabla_{x_i} y_{ij} \end{aligned} \quad (22)$$

Based on  $V_i^+$  and  $V_i^-$ , we have:

$$\begin{aligned} \text{grad}_{x_i} L(\mathbf{y}; \boldsymbol{\theta}) &= \sum_{j_1 \in V_i^+} -\frac{y_{ij_1}}{\delta_{ij_1}^2} \nabla_{x_i} y_{ij_1} + \sum_{j_2 \in V_i^-} -\frac{y_{j_2 i}}{\delta_{j_2 i}^2} \nabla_{x_i} y_{j_2 i} \\ &= \sum_{j_1 \in V_i^+} P_{j_1}(x_i) + \sum_{j_2 \in V_i^-} P_{j_2}(x_i) \end{aligned} \quad (23)$$

where  $P_{j_1}(x_i) = -\frac{y_{ij_1}}{\delta_{ij_1}^2} \nabla_{x_i} y_{ij_1}$ ,  $P_{j_2}(x_i) = -\frac{y_{j_2 i}}{\delta_{j_2 i}^2} \nabla_{x_i} y_{j_2 i}$ .

Introduce Eq.(23) into Eq.(21), we have:

$$\begin{aligned}
F_{ii_1} &= \mathbb{E} \left\{ \text{grad}_{x_i} L(\mathbf{y}; \boldsymbol{\theta}) \cdot \text{grad}_{x_{i_1}} L(\mathbf{y}; \boldsymbol{\theta})^\top \right\} \\
&= \mathbb{E} \left\{ \left( \sum_{j_1 \in V_i^+} P_{j_1}(x_i) + \sum_{j_2 \in V_i^-} P_{j_2}(x_i) \right) \left( \sum_{j_3 \in V_{i_1}^+} P_{j_3}(x_{i_1}) + \sum_{j_4 \in V_{i_1}^-} P_{j_4}(x_{i_1}) \right) \right\} \\
&= \mathbb{E} \left\{ \sum_{j_1 \in V_i^+} P_{j_1}(x_i) \sum_{j_3 \in V_{i_1}^+} P_{j_3}(x_{i_1}) + \sum_{j_1 \in V_i^+} P_{j_1}(x_i) \sum_{j_4 \in V_{i_1}^-} P_{j_4}(x_{i_1}) \right. \\
&\quad \left. + \sum_{j_2 \in V_i^-} P_{j_2}(x_i) \sum_{j_3 \in V_{i_1}^+} P_{j_3}(x_{i_1}) + \sum_{j_2 \in V_i^-} P_{j_2}(x_i) \sum_{j_4 \in V_{i_1}^-} P_{j_4}(x_{i_1}) \right\} \\
&= \sum_{j_1 \in V_i^+} \sum_{j_3 \in V_{i_1}^+} \mathbb{E}\{P_{j_1}(x_i)P_{j_3}(x_{i_1})\} + \sum_{j_1 \in V_i^+} \sum_{j_4 \in V_{i_1}^-} \mathbb{E}\{P_{j_1}(x_i)P_{j_4}(x_{i_1})\} \\
&\quad + \sum_{j_2 \in V_i^-} \sum_{j_3 \in V_{i_1}^+} \mathbb{E}\{P_{j_2}(x_i)P_{j_3}(x_{i_1})\} + \sum_{j_2 \in V_i^-} \sum_{j_4 \in V_{i_1}^-} \mathbb{E}\{P_{j_2}(x_i)P_{j_4}(x_{i_1})\}
\end{aligned} \tag{24}$$

Introduce (143) into (24), we have:

$$\begin{aligned}
&\sum_{j_1 \in V_i^+} \sum_{j_3 \in V_{i_1}^+} \mathbb{E}\{P_{j_1}(x_i)P_{j_3}(x_{i_1})\} \\
&= \mathbb{E} \left\{ \sum_{j_1 \in V_i^+} \sum_{j_3 \in V_{i_1}^+} \frac{y_{ij_1}}{\delta_{ij_1}^2} \nabla_{x_i} y_{ij_1} \nabla_{x_{i_1}} y_{i_1 j_3} \frac{y_{i_1 j_3}}{\delta_{i_1 j_3}^2} \right\} \\
&= \sum_{j_1 \in V_i^+} \sum_{j_3 \in V_{i_1}^+} 1 \cdot 1 \cdot \mathbb{E} \left\{ \frac{y_{ij_1}}{\delta_{ij_1}^2} \frac{y_{i_1 j_3}}{\delta_{i_1 j_3}^2} \right\}
\end{aligned} \tag{25}$$

$$\begin{aligned}
&\sum_{j_2 \in V_i^-} \sum_{j_4 \in V_{i_1}^-} \mathbb{E}\{P_{j_2}(x_i)P_{j_4}(x_{i_1})\} \\
&= \mathbb{E} \left\{ \sum_{j_2 \in V_i^-} \sum_{j_4 \in V_{i_1}^-} \frac{y_{j_2 i}}{\delta_{j_2 i}^2} \nabla_{x_i} y_{j_2 i} \nabla_{x_{i_1}} y_{j_4 i_1} \frac{y_{j_4 i_1}}{\delta_{j_4 i_1}^2} \right\} \\
&= \sum_{j_2 \in V_i^-} \sum_{j_4 \in V_{i_1}^-} -1 \cdot -1 \cdot \mathbb{E} \left\{ \frac{y_{j_2 i}}{\delta_{j_2 i}^2} \frac{y_{j_4 i_1}}{\delta_{j_4 i_1}^2} \right\}
\end{aligned} \tag{26}$$

$$\begin{aligned}
&\sum_{j_1 \in V_i^+} \sum_{j_4 \in V_{i_1}^-} \mathbb{E}\{P_{j_1}(x_i)P_{j_4}(x_{i_1})\} \\
&= \mathbb{E} \left\{ \sum_{j_1 \in V_i^+} \sum_{j_4 \in V_{i_1}^-} \frac{y_{ij_1}}{\delta_{ij_1}^2} \nabla_{x_i} y_{ij_1} \nabla_{x_{i_1}} y_{j_4 i_1} \frac{y_{j_4 i_1}}{\delta_{j_4 i_1}^2} \right\} \\
&= \sum_{j_1 \in V_i^+} \sum_{j_4 \in V_{i_1}^-} 1 \cdot -1 \cdot \mathbb{E} \left\{ \frac{y_{ij_1}}{\delta_{ij_1}^2} \frac{y_{j_4 i_1}}{\delta_{j_4 i_1}^2} \right\}
\end{aligned} \tag{27}$$

$$\begin{aligned}
& \sum_{j_2 \in V_i^-} \sum_{j_3 \in V_{i_1}^+} \mathbb{E}\{y_{j_2} y_{j_3}\} \\
&= \mathbb{E} \left\{ \sum_{j_2 \in V_i^-} \sum_{j_3 \in V_{i_1}^+} \frac{y_{j_2 i}}{\delta_{j_2 i}^2} \nabla_{x_i} y_{j_2 i} \nabla_{x_{i_1}} y_{i_1 j_3} \frac{y_{i_1 j_3}}{\delta_{i_1 j_3}^2} \right\} \\
&= \sum_{j_2 \in V_i^-} \sum_{j_3 \in V_{i_1}^+} -1 \cdot 1 \cdot \mathbb{E} \left\{ \frac{y_{j_2 i}}{\delta_{j_2 i}^2} \frac{y_{i_1 j_3}}{\delta_{i_1 j_3}^2} \right\}
\end{aligned} \tag{28}$$

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

Because we know that  $\mathbb{E}\{y_{ij}\} = 0$ . Based on Lemma.1 and when  $i = i_1$ , we have:

$$\mathbb{E} \left\{ \frac{y_{ij_1}}{\delta_{ij_1}^2} \frac{y_{i_1 j_3}}{\delta_{i_1 j_3}^2} \right\} = \mathbb{E} \left\{ \frac{y_{ij_1}}{\delta_{ij_1}^2} \frac{y_{ij_3}}{\delta_{ij_3}^2} \right\} = \begin{cases} \delta_{ij_1}^{-2}, & j_1 = j_3 \\ \mathbb{E} \left\{ \frac{y_{ij_1}}{\delta_{ij_1}^2} \right\} \mathbb{E} \left\{ \frac{y_{i_1 j_3}}{\delta_{i_1 j_3}^2} \right\} = 0, & j_1 \neq j_3 \end{cases} \tag{29}$$

$$\mathbb{E} \left\{ \frac{y_{j_2 i}}{\delta_{j_2 i}^2} \frac{y_{j_4 i_1}}{\delta_{j_4 i_1}^2} \right\} = \mathbb{E} \left\{ \frac{y_{j_2 i}}{\delta_{j_2 i}^2} \frac{y_{j_4 i}}{\delta_{j_4 i}^2} \right\} = \begin{cases} \delta_{j_2 i}^{-2}, & j_2 = j_4 \\ \mathbb{E} \left\{ \frac{y_{j_2 i}}{\delta_{j_2 i}^2} \right\} \mathbb{E} \left\{ \frac{y_{j_4 i}}{\delta_{j_4 i}^2} \right\} = 0, & j_2 \neq j_4 \end{cases} \tag{30}$$

$$\mathbb{E} \left\{ \frac{y_{ij_1}}{\delta_{ij_1}^2} \frac{y_{j_4 i_1}}{\delta_{j_4 i_1}^2} \right\} = \mathbb{E} \left\{ \frac{y_{ij_1}}{\delta_{ij_1}^2} \frac{y_{j_4 i}}{\delta_{j_4 i}^2} \right\} = 0 \tag{31}$$

$$\mathbb{E} \left\{ \frac{y_{j_2 i}}{\delta_{j_2 i}^2} \frac{y_{i_1 j_3}}{\delta_{i_1 j_3}^2} \right\} = \mathbb{E} \left\{ \frac{y_{j_2 i}}{\delta_{j_2 i}^2} \frac{y_{ij_3}}{\delta_{ij_3}^2} \right\} = 0 \tag{32}$$

Then, we have, if  $i = i_1$ :

$$\begin{aligned}
F_{ii_1} &= \sum_{j_1 \in V_i^+} \sum_{j_3 \in V_{i_1}^+} \mathbb{E}\{y_{j_1} y_{j_3}\} + \sum_{j_1 \in V_i^+} \sum_{j_4 \in V_{i_1}^-} \mathbb{E}\{y_{j_1} y_{j_4}\} + \sum_{j_2 \in V_i^-} \sum_{j_3 \in V_{i_1}^+} \mathbb{E}\{y_{j_2} y_{j_3}\} + \sum_{j_2 \in V_i^-} \sum_{j_4 \in V_{i_1}^-} \mathbb{E}\{y_{j_2} y_{j_4}\} \\
&= \sum_{j_1 \in V_i^+} \delta_{ij_1}^{-2} + \sum_{j_1 \in V_i^+} \sum_{j_4 \in V_{i_1}^-} 0 + \sum_{j_2 \in V_i^-} \sum_{j_3 \in V_{i_1}^+} 0 + \sum_{j_2 \in V_i^-} \delta_{j_2 i}^{-2} \\
&= \sum_{j_1 \in V_i^+} \delta_{ij_1}^{-2} + \sum_{j_2 \in V_i^-} \delta_{j_2 i}^{-2}
\end{aligned} \tag{33}$$

When  $i \neq i_1$ , we have:

$$\mathbb{E} \left\{ \frac{y_{ij_1}}{\delta_{ij_1}^2} \frac{y_{i_1 j_3}}{\delta_{i_1 j_3}^2} \right\} = 0 \tag{34}$$

$$\mathbb{E} \left\{ \frac{y_{j_2 i}}{\delta_{j_2 i}^2} \frac{y_{j_4 i_1}}{\delta_{j_4 i_1}^2} \right\} = 0 \tag{35}$$

$$\mathbb{E} \left\{ \frac{y_{ij_1}}{\delta_{ij_1}^2} \frac{y_{j_4 i_1}}{\delta_{j_4 i_1}^2} \right\} = \begin{cases} \delta_{ij_1}^{-2}, & i = j_4 \text{ and } j_1 = i_1, (i, i_1) \in \mathcal{E} \\ \mathbb{E} \left\{ \frac{y_{ij_1}}{\delta_{ij_1}^2} \right\} \mathbb{E} \left\{ \frac{y_{j_4 i_1}}{\delta_{j_4 i_1}^2} \right\} = 0, & \text{else} \end{cases} \tag{36}$$

$$\mathbb{E} \left\{ \frac{y_{j_2 i}}{\delta_{j_2 i}^2} \frac{y_{i_1 j_3}}{\delta_{i_1 j_3}^2} \right\} = \begin{cases} \delta_{j_2 i}^{-2}, & j_2 = i_1 \text{ and } i = j_3, (i_1, i) \in \mathcal{E} \\ \mathbb{E} \left\{ \frac{y_{j_2 i}}{\delta_{j_2 i}^2} \right\} \mathbb{E} \left\{ \frac{y_{i_1 j_3}}{\delta_{i_1 j_3}^2} \right\} = 0, & j_2 \neq j_3 \end{cases} \tag{37}$$

We known that for one edge  $(i, i_1) \in \mathcal{E}$ , (36) is nonzero and (37) is zero. And the same time, if its corresponding  $(i_1, i)$  belongs to  $\mathcal{E}$ , (37) will be nonzero and (36) will be zero.

Then, combine (36) and (37), and  $i \neq i_1$ , we have:

$$F_{ii_1} = \begin{cases} -\delta_{i_1}^{-2}, & (i, i_1) \in \mathcal{E} \\ -\delta_{i_1 i}^{-2}, & (i_1, i) \in \mathcal{E} \\ 0, & \text{else} \end{cases} \quad (38)$$

Combine Eq.(38) and Eq.(33), we have:

$$F_{ii_1} = \begin{cases} \sum_{j_1 \in V_i^+} \delta_{ij_1}^{-2} + \sum_{j_2 \in V_i^-} \delta_{j_2 i}^{-2}, & i = i_1 \\ -\delta_{i_1 i}^{-2}, & (i, i_1) \in \mathcal{E} \\ -\delta_{i_1 i}^{-2}, & (i_1, i) \in \mathcal{E} \\ 0, & \text{else} \end{cases} \quad (39)$$

It is easy to find that it equals to the weighted Laplacian matrix. In order to unify the symbol, we can re-write the above equation by the new symbol  $(\mathbf{L}_{w_{\mathbb{R}}})_{i,i_1}$ , of which  $(i, i_1)$ -th block is:

$$(\mathbf{L}_{w_{\mathbb{R}}})_{i,i_1} = \begin{cases} \sum_{j \in V_i} w_{ij}^{\mathbb{R}}, & i = i_1 \\ -w_{i_1 i}^{\mathbb{R}}, & (i, i_1) \in \mathcal{E} \\ -w_{i_1 i}^{\mathbb{R}}, & (i_1, i) \in \mathcal{E} \\ 0, & \text{else,} \end{cases} \quad (40)$$

where  $w_{ij}^{\mathbb{R}} = \delta_{ij}^{-2}$ ,  $\mathbf{L}_w^{\mathbb{R}^1} = \mathbf{L}_{w_{\mathbb{R}}} \otimes \mathbf{I}_{1 \times 1}$ . So the FIM corresponding to the space  $\mathbb{R}^1$  is  $\mathbf{F} = \mathbf{L}_w^{\mathbb{R}^1}$ .

### 4.3 Discussion about the directed graph and undirected graph

The FIM of  $\mathbb{R}^1$  is deduced by a directed graph. Its log-likelihood function does not need to add  $\frac{1}{2}$ . We get this result. If you see it as an undirected graph and add  $\frac{1}{2}$  to its log-likelihood function. You can get the same result.

### 4.4 FIM using Definition in [39] under Gaussian distribution

We can re-write all measurement equations based on Eq.(11) as:

$$\mathbf{p} = \mathbf{A}_0^{G^\top} \mathbf{x} + \mathbf{y} \quad (41)$$

where  $\mathbf{p} = (\dots, p_{ij}, \dots)^\top$ ,  $\mathbf{y} = (\dots, y_{ij}, \dots)^\top$ ,  $(i, j) \in \mathcal{E}$ , It is easy to compute the FIM by:

$$\begin{aligned} \mathbf{F} &= J(\mathbf{x})^\top \Sigma^{-1} J(\mathbf{x}) \\ &= \mathbf{A}_0^G \Sigma_0^{-1} \mathbf{A}_0^{G^\top} \\ &= \mathbf{L}_{w_{\mathbb{R}}} \end{aligned} \quad (42)$$

where  $J(\mathbf{x})$  means the Jacobian matrix corresponding to the state vector  $\mathbf{x}$ . It is easy to find that it equal to the weighted Laplacian matrix. For the anchored situation (anchored by  $x_0$ ), we have:

$$\begin{aligned} \mathbf{p} &= \mathbf{A}^G \bar{\mathbf{x}} + \mathbf{y} \\ \mathbf{x} &= [x_0^\top, \bar{\mathbf{x}}^\top]^\top \\ \mathbf{F} &= J(\bar{\mathbf{x}})^\top \Sigma^{-1} J(\bar{\mathbf{x}}) \\ &= \mathbf{A}^G \Sigma^{-1} \mathbf{A}^{G^\top} \end{aligned} \quad (43)$$

where  $J(\bar{\mathbf{x}})$  means the Jacobian matrix corresponding to the state vector  $\bar{\mathbf{x}}$ .

**Example 2.** Let's consider the pose-graph 1 as a example. Its log-likelihood function is shown in (139). Let's discuss  $F_{22}$  and  $F_{23}$ .

For  $F_{22}$ , where  $i = 2$  and  $i_1 = 2$ , we have:

$$\begin{aligned}
\text{grad}_{x_2} L(\mathbf{y}; \boldsymbol{\theta}) &= \nabla_{x_2} (\log f_{21}(p_{21} - (x_2 - x_1)) + \log f_{32}(p_{32} - (x_3 - x_2)) + \log f_{20}(p_{20} - (x_2 - x_0))) \\
&= -\frac{y_{21}}{\delta_{21}^2} \nabla_{x_2} y_{21} - \frac{y_{32}}{\delta_{32}^2} \nabla_{x_2} y_{32} - \frac{y_{20}}{\delta_{20}^2} \nabla_{x_2} y_{20} \\
&= -\frac{y_{21}}{\delta_{21}^2} \cdot (-1) - \frac{y_{32}}{\delta_{32}^2} \cdot 1 - \frac{y_{20}}{\delta_{20}^2} \cdot (-1) \\
&= \frac{y_{21}}{\delta_{21}^2} - \frac{y_{32}}{\delta_{32}^2} + \frac{y_{20}}{\delta_{20}^2}
\end{aligned} \tag{44}$$

Based on (44), we can get:

$$\begin{aligned}
F_{22} &= \mathbb{E} \{ \text{grad}_{x_2} L(\mathbf{y}; \boldsymbol{\theta}) \text{grad}_{x_2} L(\mathbf{y}; \boldsymbol{\theta})^\top \} \\
&= \mathbb{E} \left\{ \left( \frac{y_{21}}{\delta_{21}^2} - \frac{y_{32}}{\delta_{32}^2} + \frac{y_{20}}{\delta_{20}^2} \right) \left( \frac{y_{21}}{\delta_{21}^2} - \frac{y_{32}}{\delta_{32}^2} + \frac{y_{20}}{\delta_{20}^2} \right)^\top \right\} \\
&= \mathbb{E} \left\{ \frac{y_{21}}{\delta_{21}^2} \frac{y_{21}}{\delta_{21}^2} + \frac{y_{32}}{\delta_{32}^2} \frac{y_{32}}{\delta_{32}^2} + \frac{y_{20}}{\delta_{20}^2} \frac{y_{20}}{\delta_{20}^2} \right\} \\
&= \delta_{21}^{-2} + \delta_{32}^{-2} + \delta_{20}^{-2}
\end{aligned} \tag{45}$$

For  $F_{23}$ , where  $i = 2$  and  $i_1 = 3$ , we have:

$$\begin{aligned}
\text{grad}_{x_3} L(\mathbf{y}; \boldsymbol{\theta}) &= \nabla_{x_3} (\log f_{32}(p_{32} - (x_3 - x_2)) + \log f_{31}(p_{31} - (x_3 - x_1))) \\
&= -\frac{y_{32}}{\delta_{32}^2} \nabla_{x_3} y_{32} - \frac{y_{31}}{\delta_{31}^2} \nabla_{x_3} y_{31} \\
&= -\frac{y_{32}}{\delta_{32}^2} \cdot (-1) - \frac{y_{31}}{\delta_{31}^2} \cdot (-1) \\
&= \frac{y_{32}}{\delta_{32}^2} + \frac{y_{31}}{\delta_{31}^2}
\end{aligned} \tag{46}$$

Based on (44) and (46), we have:

$$\begin{aligned}
F_{32} = F_{23} &= \mathbb{E} \{ \text{grad}_{x_2} L(\mathbf{y}; \boldsymbol{\theta}) \text{grad}_{x_3} L(\mathbf{y}; \boldsymbol{\theta})^\top \} \\
&= \mathbb{E} \left\{ \left( \frac{y_{21}}{\delta_{21}^2} - \frac{y_{32}}{\delta_{32}^2} + \frac{y_{20}}{\delta_{20}^2} \right) \left( \frac{y_{32}}{\delta_{32}^2} + \frac{y_{31}}{\delta_{31}^2} \right)^\top \right\} \\
&= \mathbb{E} \left\{ -\frac{y_{32}}{\delta_{32}^2} \frac{y_{32}}{\delta_{32}^2} \right\} \\
&= -\delta_{32}^{-2}
\end{aligned} \tag{47}$$

## 5 Synchronization on $\mathbb{R}^2$

### 5.1 Synchronization on $\mathbb{R}^2$

Synchronization on  $\mathbb{R}^2$  is the problem of estimating a set of vectors  $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{n_p} \in \mathbb{R}^2$  from noise measurements of some relative vectors  $\mathbf{p}_{ij}$ . The operations  $\oplus$  and  $\ominus$  are defined as:  $\mathbf{X}_i \oplus \mathbf{X}_j = \mathbf{x}_i + \mathbf{x}_j$  and  $\mathbf{X}_j \ominus \mathbf{X}_i = \mathbf{x}_j - \mathbf{x}_i$ .

So the measurement  $\mathbf{p}_{ij}$  is given by

$$\mathbf{p}_{ij} = \mathbf{y}_{ij} + \mathbf{x}_j - \mathbf{x}_i. \tag{48}$$

where  $\mathbf{y}_{ij}$  is a isotropic zero-mean Gaussian noise in  $\mathbb{R}^2$  with covariance  $\delta_{ij}^2 \mathbf{I}_{2 \times 2}$ . Its distributed function  $f_{ij} : \mathbb{R}^2 \rightarrow \mathbb{R}^+$ :

$$\begin{aligned}
f_{ij}(\mathbf{y}_{ij}) &= \frac{1}{2\pi \det(\boldsymbol{\Sigma}_{ij})^{1/2}} \exp(-\frac{1}{2}(\mathbf{y}_{ij}^\top \boldsymbol{\Sigma}_{ij}^{-1} \mathbf{y}_{ij})) \\
\boldsymbol{\Sigma}_{ij} &= \delta_{ij}^2 \mathbf{I}_{2 \times 2}
\end{aligned} \tag{49}$$

Based on a undirected graph  $\mathcal{F}$ , which has the same nodes with the graph  $\mathcal{E}$ , hence

$$\begin{aligned}
L(\mathbf{y}; \boldsymbol{\theta}) &= \frac{1}{2} \sum_{(i,j) \in \mathcal{F}} \log f_{ij}(\mathbf{y}_{ij} + \mathbf{x}_j - \mathbf{x}_i) \\
&= \frac{1}{2} \sum_{(i,j) \in \mathcal{F}} \log \frac{1}{2\pi \det(\boldsymbol{\Sigma}_{ij})^{1/2}} \exp(-\frac{1}{2}((\mathbf{y}_{ij} - \mathbf{x}_j + \mathbf{x}_i)^\top \boldsymbol{\Sigma}_{ij}^{-1} (\mathbf{y}_{ij} - \mathbf{x}_j + \mathbf{x}_i)))
\end{aligned} \tag{50}$$

## 5.2 FIM for $\mathbb{R}^2$

The synchronization on space  $\mathcal{P}' = \mathbb{R}^2$  is the problem of estimating a set of variables (nodes in the graph except the anchor  $\mathbf{x}_0$ ),  $\boldsymbol{\theta} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n_p}\} \in \mathcal{P} = \{\mathcal{P}' \times \mathcal{P}' \times \dots \times \mathcal{P}'\}$  from noisy measurements  $\mathbf{p}_{ij}$  of relative information between  $\mathbf{x}_i$  and  $\mathbf{x}_j$  (edges in the graph,  $0 \leq i, j \leq n_p$ ). The orthonormal basis  $e = (\mathbf{E}_{1,1}^X, \mathbf{E}_{1,2}^X, \dots, \mathbf{E}_{i,1}^X, \mathbf{E}_{i,2}^X, \dots, \mathbf{E}_{n_p,1}^X, \mathbf{E}_{n_p,2}^X)$  of the tangent space  $\mathcal{T}_{\boldsymbol{\theta}}\mathcal{P}'$ :

$$\mathbf{E}_{i,k}^X = (\mathbf{0}_{1 \times 2}, \dots, \underbrace{\mathbf{0}_{1 \times 2}, 0, \underbrace{\mathbf{0}_{1 \times 2}, \dots, \mathbf{0}_{1 \times 2}}_{k-\text{th}}, \dots, \underbrace{\mathbf{0}_{1 \times 2}, \dots, \mathbf{0}_{1 \times 2}}_{i-\text{th}})_{1 \times 2(n_p+1)}, \quad k = 1, 2 \quad (51)$$

We deal with the inner product corresponding to the bases of the coordinate of the  $i$ -th pose as a  $\mathbb{R}^{2 \times 1}$  vector meeting:

$$\langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_i^X \rangle_{\mathbf{X}}^* = (\langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_{i,1}^X \rangle_{\mathbf{X}}, \langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_{i,2}^X \rangle_{\mathbf{X}})^\top, \quad (52)$$

Then, based on the definition (5), we have:

$$\begin{aligned} \mathbf{F}_{i,i_1} &= \mathbb{E}\{\langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_i^X \rangle_{\mathbf{X}}^* \cdot \langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_{i_1}^X \rangle_{\mathbf{X}}^\top\} \\ &= \mathbb{E}\{\text{grad}_{\mathbf{x}_i} L(\mathbf{y}; \boldsymbol{\theta})^\top \cdot \text{grad}_{\mathbf{x}_{i_1}} L(\mathbf{y}; \boldsymbol{\theta})\}, \end{aligned} \quad (53)$$

where  $\mathbf{F}_{i,i_1} \in \mathbb{R}^{2 \times 2}$ ,  $\text{grad}_{\mathbf{x}_i} L(\mathbf{y}; \boldsymbol{\theta})$  means the gradient of  $L(\mathbf{y}; \boldsymbol{\theta})$  with respect to parameter  $\mathbf{x}_i$ .

Based on the chain rule, we have:

$$\begin{aligned} &\text{grad}_{\mathbf{x}_i} L(\mathbf{y}; \boldsymbol{\theta}) \\ &= \nabla_i \frac{1}{2} \sum_{(i,j) \in \mathcal{F}} \log f_{ij}(\mathbf{p}_{ij} - \mathbf{x}_j + \mathbf{x}_i) \\ &= \nabla_i \frac{1}{2} \sum_{(i,j) \in \mathcal{F}} \log \frac{1}{2\pi \det(\boldsymbol{\Sigma}_{ij})^{1/2}} \exp(-\frac{1}{2}((\mathbf{y}_{ij} - \mathbf{x}_j + \mathbf{x}_i)^\top \boldsymbol{\Sigma}_{ij}^{-1} (\mathbf{y}_{ij} - \mathbf{x}_j + \mathbf{x}_i))) \\ &= \nabla_i \frac{1}{2} \sum_{(i,j) \in \mathcal{F}} \log \frac{1}{2\pi \det(\boldsymbol{\Sigma}_{ij})^{1/2}} \exp(-\frac{1}{2}((\mathbf{y}_{ij} - \mathbf{x}_j + \mathbf{x}_i)^\top \boldsymbol{\Sigma}_{ij}^{-1} (\mathbf{y}_{ij} - \mathbf{x}_j + \mathbf{x}_i))) \\ &= \nabla_i \sum_{j \in V_i} \log \frac{1}{2\pi \det(\boldsymbol{\Sigma}_{ij})^{1/2}} \exp(-\frac{1}{2}((\mathbf{y}_{ij} - \mathbf{x}_j + \mathbf{x}_i)^\top \boldsymbol{\Sigma}_{ij}^{-1} (\mathbf{y}_{ij} - \mathbf{x}_j + \mathbf{x}_i))) \\ &= \nabla_i \sum_{j \in V_i} \log \frac{1}{2\pi \det(\boldsymbol{\Sigma}_{ij})^{1/2}} + \log \exp(-\frac{1}{2}((\mathbf{y}_{ij} - \mathbf{x}_j + \mathbf{x}_i)^\top \boldsymbol{\Sigma}_{ij}^{-1} (\mathbf{y}_{ij} - \mathbf{x}_j + \mathbf{x}_i))) \\ &= \nabla_i \sum_{j \in V_i} \log \exp(-\frac{1}{2}((\mathbf{y}_{ij} - \mathbf{x}_j + \mathbf{x}_i)^\top \boldsymbol{\Sigma}_{ij}^{-1} (\mathbf{y}_{ij} - \mathbf{x}_j + \mathbf{x}_i))) \\ &= \nabla_i \sum_{j \in V_i} -\frac{1}{2}((\mathbf{y}_{ij} - \mathbf{x}_j + \mathbf{x}_i)^\top \boldsymbol{\Sigma}_{ij}^{-1} (\mathbf{y}_{ij} - \mathbf{x}_j + \mathbf{x}_i)) \\ &= \nabla_i \sum_{j \in V_i} -\frac{1}{2}(\mathbf{y}_{ij}^\top \boldsymbol{\Sigma}_{ij}^{-1} \mathbf{y}_{ij}) \\ &= \nabla_{\mathbf{y}_{ij}} \sum_{j \in V_i} -\frac{1}{2}(\mathbf{y}_{ij}^\top \boldsymbol{\Sigma}_{ij}^{-1} \mathbf{y}_{ij}) \nabla_{\mathbf{x}_i} \mathbf{y}_{ij} \\ &= \sum_{j \in V_i} -\frac{1}{2} \mathbf{y}_{ij}^\top (\boldsymbol{\Sigma}_{ij}^{-1} + \boldsymbol{\Sigma}_{ij}^{-T}) \nabla_{\mathbf{x}_i} \mathbf{y}_{ij} \\ &= \sum_{j \in V_i} -\delta_{ij}^{-2} \mathbf{y}_{ij}^\top \mathbf{I}_{2 \times 2} \nabla_{\mathbf{x}_i} \mathbf{y}_{ij} \\ &= \sum_{j \in V_i} -\delta_{ij}^{-2} \mathbf{y}_{ij}^\top \nabla_{\mathbf{x}_i} \mathbf{y}_{ij} \end{aligned} \quad (54)$$

Introduce (54) into (53), we have:

$$\begin{aligned}
\mathbf{F}_{ii_1} &= \mathbb{E}\left\{\sum_{j \in V_i} \delta_{ij}^{-2} (\mathbf{y}_{ij}^\top \nabla_{\mathbf{x}_i} \mathbf{y}_{ij})^\top \cdot \sum_{j_1 \in V_{i_1}} \delta_{i_1 j_1}^{-2} \mathbf{y}_{i_1 j_1}^\top \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1}\right\} \\
&= \mathbb{E}\left\{\sum_{j \in V_i} \sum_{j \in V_{i_1}} \delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} (\nabla_{\mathbf{x}_i} \mathbf{y}_{ij})^\top \mathbf{y}_{ij} \cdot \mathbf{y}_{i_1 j_1}^\top \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1}\right\} \\
&= \sum_{j \in V_i} \sum_{j \in V_{i_1}} \delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \mathbb{E}\{(\nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{ij})^\top \mathbf{y}_{ij} \cdot \mathbf{y}_{i_1 j_1}^\top \nabla_{\mathbf{x}_i} \mathbf{y}_{i_1 j_1}\}
\end{aligned} \tag{55}$$

When  $i = i_1$ , we have:

$$\begin{aligned}
\mathbf{F}_{ii_1} &= \sum_{j \in V_i} \sum_{j \in V_{i_1}} \delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \mathbb{E}\{(\nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{ij})^\top \mathbf{y}_{ij} \cdot \mathbf{y}_{i_1 j_1}^\top \nabla_{\mathbf{x}_i} \mathbf{y}_{i_1 j_1}\} \\
&= \sum_{j \in V_i} \delta_{ii_1}^{-4} \mathbb{E}\{\mathbf{y}_{ij} \mathbf{y}_{i_1 j_1}^\top\} \\
&= \sum_{j \in V_i} \delta_{ii_1}^{-2} \mathbf{I}_{2 \times 2}
\end{aligned} \tag{56}$$

When  $i \neq i_1$  and  $(i, i_1) \in \varepsilon$ , we have:

$$\mathbf{F}_{ii_1} = \sum_{j \in V_i} \sum_{j \in V_{i_1}} \delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \mathbb{E}\{(\nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{ij})^\top \mathbf{y}_{ij} \cdot \mathbf{y}_{i_1 j_1}^\top \nabla_{\mathbf{x}_i} \mathbf{y}_{i_1 j_1}\} \tag{57}$$

Because  $\mathbf{y}_{ii_1} = \mathbf{p}_{ii_1} - \mathbf{x}_{i_1} + \mathbf{x}_i$ , we have:

$$\begin{aligned}
\nabla_{\mathbf{x}_i} \mathbf{y}_{ii_1} &= \mathbf{I}_{2 \times 2} \\
\nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{ii_1} &= -\mathbf{I}_{2 \times 2}
\end{aligned} \tag{58}$$

Using (55) and (56), we can get:

$$\begin{aligned}
\mathbf{F}_{ii_1} &= -\delta_{ii_1}^{-4} \mathbb{E}\{\mathbf{y}_{ii_1} \mathbf{y}_{ii_1}^\top\} \\
&= -\delta_{ii_1}^{-2} \mathbf{I}_{2 \times 2}
\end{aligned} \tag{59}$$

Combine Eq.(59) and Eq.(56), we have:

$$\mathbf{F}_{ii_1} = \begin{cases} \sum_{j \in V_i} \delta_{ij}^{-2} \mathbf{I}_{2 \times 2}, & i = i_1 \\ -\delta_{ii_1}^{-2} \mathbf{I}_{2 \times 2}, & \{i, i_1\} \in \varepsilon \\ -\delta_{ii_1}^{-2} \mathbf{I}_{2 \times 2}, & \{i_1, i\} \in \varepsilon \\ \mathbf{0}_{2 \times 2}, & \text{else} \end{cases} \tag{60}$$

It can be seen that  $\mathbf{L}_w^{\mathbb{R}^2} = \mathbf{L}_{w_{\mathbb{R}}} \otimes \mathbf{I}_{2 \times 2}$ , where:

$$(\mathbf{L}_{w_{\mathbb{R}}})_{i,i_1} = \begin{cases} \sum_{j \in V_i} w_{ij}^{\mathbb{R}} & i = i_1 \\ -w_{ii_1}^{\mathbb{R}} & (i, i_1) \in \mathcal{E} \\ -w_{i_1 i}^{\mathbb{R}} & (i_1, i) \in \mathcal{E} \\ 0 & \text{else,} \end{cases} \tag{61}$$

where  $w_{ij}^{\mathbb{R}} = \delta_{ij}^{-2}$ . Thus  $\mathbf{L}_{w_{\mathbb{R}}}$  is a weighted Laplacian matrix.

## 6 Synchronization on $SO(2)$

### 6.1 Discussion about isotropic Langevin distribution

The isotropic Langevin distribution on  $SO(n), n = 2, 3$  with mean  $\mathbf{Q} \in SO(n)$  and concentration  $\kappa \geq 0$  has PDF [2]:

$$\begin{aligned}\widehat{f}(\mathbf{Z}) &= \frac{1}{c_n(\kappa)} \exp(\kappa \operatorname{trace}(\mathbf{Q}^\top \mathbf{Z})) \\ c_2(\kappa) &= I_0(2\kappa) \\ c_3(\kappa) &= \exp(\kappa)(I_0(2\kappa) - I_1(2\kappa)) \\ I_0(2\kappa) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(2\kappa \cos(\theta)) d\theta \\ I_1(2\kappa) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(2\kappa \cos(\theta)) \cos(\theta) d\theta\end{aligned}\tag{62}$$

where  $c_n(\kappa)$  is a normalization constant such that  $f$  has unit mass. We write  $\mathbf{Z} \sim Lang(\mathbf{Q}, \kappa)$  to mean that  $\mathbf{Z}$  is a random rotation (in  $SO(n)$ ) with PDF (62).

Let's first discuss  $SO(2)$  situation compared with the Gaussian distribution  $\mathcal{N}(0, \delta^2)$ .  $\mathbf{Q}$  is set as an identity matrix.  $\kappa$  changes from 0 to 10.  $\eta$  represents the orientation angle of  $\mathbf{Z} \in SO(2)$  changes from  $-\pi$  to  $\pi$  at 0.01 interval. The PDF needs firstly be normalized by  $2\pi$  and then meets following equation:

$$\delta = \frac{1}{\sqrt{2\kappa}}\tag{63}$$

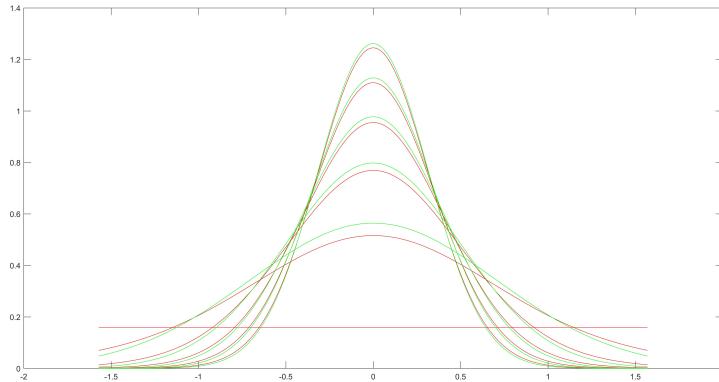


Figure 2: Green is the Gaussian distribution and Red is the normalized isotropic Langevin distribution

Then, let's discuss the isotropic Langevin distribution on  $SO(3)$ . In fact, the isotropic Langevin distribution on  $SO(3)$  is obtained by the combination of the isotropic Langevin distribution on  $SO(2)$  and the uniform distribution. Just like Fig.3, the spherical surface is the surface for  $SO(3)$ . We assume that some random points (corresponding to some rotations in  $SO(3)$ ), following the isotropic Langevin distribution on  $SO(3)$  with a mean rotation  $\mathbf{Q}$  (corresponding to a center point on surface), localizing at the surface of the sphere. The way to obtain these random points are shown as follows: (1) We first randomly generate a circle surface  $\mathcal{A}$  passes the sphere center and the center point by the uniform distribution. (2) Based on this circle surface  $\mathcal{A}$ , we can generate the random points using the isotropic Langevin distribution on  $SO(2)$ .

### 6.2 Synchronization on $SO(2)$

Synchronization on  $SO(2)$  is the problem of estimating a set of rotations  $\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_{n_p} \in SO(2)$  from noise measurements of some relative rotations  $\mathbf{R}_i \mathbf{R}_j^\top$ . The operations  $\oplus$  and  $\ominus$  are defined as:  $\mathbf{R}_i \oplus \mathbf{R}_j = \mathbf{R}_i \mathbf{R}_j$  and  $\mathbf{R}_i \ominus \mathbf{R}_j = \mathbf{R}_i \mathbf{R}_j^\top$ .

In our estimation problem, the target quantities (the parameters) are the rotation matrices  $\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_{n_p} \in SO(2)$ . The natural parameter space is thus:

$$\mathcal{P} = \{SO(2) \times \dots \times SO(2)\}_{n_p}\tag{64}$$

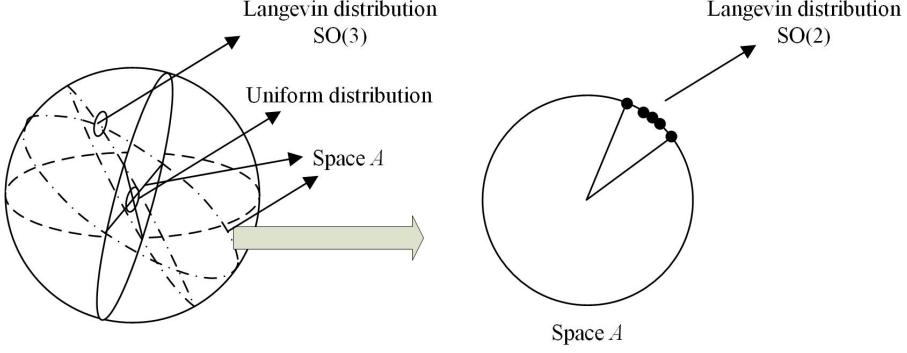


Figure 3: Green is the Gaussian distribution and Red is the normalized isotropic Langevin distribution

We can get a measurement un-directed graph  $\mathcal{G} = ([n_p], \mathcal{F})$ , we have one measurement  $\mathbf{H}_{ij} \in SO(2)$  of a form:

$$\mathbf{H}_{ij} = \mathbf{Z}_{ij} \mathbf{R}_j \mathbf{R}_i^\top \quad (65)$$

where  $\mathbf{Z}_{ij}$  is a random value whose distributed function  $\hat{f}_{ij} : SO(2) \rightarrow \mathbb{R}^+$  meets an isotropic Langevin distribution:

$$\begin{aligned} \hat{f}_{ij}(\mathbf{Z}_{ij}) &= \frac{1}{c_2(k_{ij})} \exp(\kappa_{ij} \operatorname{trace}(\mathbf{Z}_{ij})) \\ c_2(\kappa_{ij}) &= I_0(2\kappa_{ij}) \\ I_0(2\kappa_{ij}) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(2\kappa_{ij} \cos(\theta)) d\theta \end{aligned} \quad (66)$$

Hence

$$L(\mathbf{y}; \boldsymbol{\theta}) = \frac{1}{2} \sum_{(i,j) \in \mathcal{F}} \log \hat{f}_{ij}(\mathbf{H}_{ij} \mathbf{R}_i \mathbf{R}_j^\top) = \frac{1}{2} \sum_{(i,j) \in \mathcal{F}} \log \left( \frac{1}{c_2(\kappa_{ij})} \exp(\kappa_{ij} \operatorname{trace}(\mathbf{H}_{ij} \mathbf{R}_i \mathbf{R}_j^\top)) \right) \quad (67)$$

### 6.3 FIM for $SO(2)$

The synchronization on  $SO(2)$  is the problem of estimating a set of variables (nodes in the graph except the anchor  $\mathbf{R}_0$ ),  $\theta = \{\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_{n_p}\} \in \mathcal{P}$  from noisy measurements  $\mathbf{H}_{ij}$  of relative information between  $\mathbf{R}_i$  and  $\mathbf{R}_j$  (edges in the rotation graph). The orthonormal basis  $e = (e_1, \dots, e_i, \dots, e_{n_p})$  of the tangent space  $\mathcal{T}_{\mathbf{R}} SO(2)$  as:  $\mathbf{R}_1 \mathbf{E}, \dots, \mathbf{R}_{n_p} \mathbf{E}$ , where  $\mathbf{E} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$  is:

$$\mathbf{E}_i^R = (\mathbf{0}, \dots, \mathbf{0}, \mathbf{R}_i \mathbf{E}, \mathbf{0}, \dots, \mathbf{0})_{2 \times 2n_p}^\top \quad (68)$$

The FIM of the estimation problem w.r.t. the basis  $e$  is defined as:

$$\begin{aligned} \mathbf{F}_{ii} &= \mathbb{E}\{\langle \operatorname{grad} L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_i^R \rangle_R \cdot \langle \operatorname{grad} L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_{i_1}^R \rangle_R\} \\ &= \mathbb{E}\{\langle \operatorname{grad}_{\mathbf{R}_i} L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{R}_i \mathbf{E} \rangle_R \cdot \langle \operatorname{grad}_{\mathbf{R}_{i_1}} L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{R}_{i_1} \mathbf{E} \rangle_R\} \end{aligned} \quad (69)$$

Based on the PDF (67) for the measurement error  $\mathbf{Z}_{ij}$ , solving (129) for  $\mathbf{Z}_{ij}$  and substituting into the PDF (67), we have:

$$\log \hat{f}_{ij}(\mathbf{Z}_{ij}) = -\log(c_n(\kappa_{ij})) + \kappa_{ij} \operatorname{trace}(\mathbf{H}_{ij} \mathbf{R}_i \mathbf{R}_j^\top). \quad (70)$$

Since the inner product on  $\mathbb{R}^{n \times n}$  is  $\langle \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2 \rangle = \operatorname{trace}(\boldsymbol{\Omega}_1^\top \boldsymbol{\Omega}_2)$ , the gradient (99) implies that the Euclidean gradients of the log-likelihood function are:

$$\nabla_{\mathbf{R}_i} \log \hat{f}_{ij}(\mathbf{Z}_{ij}) = \kappa_{ij} \mathbf{H}_{ij}^\top \mathbf{R}_j. \quad (71)$$

Finally, based on the gradient on  $SO(n)$  (1), we have the Riemannian gradient of the log-likelihood function is:

$$\begin{aligned} \operatorname{grad}_{\mathbf{R}_i} \sum_{j \in V_i} \log \hat{f}_{ij}(\mathbf{Z}_{ij}) &= \sum_{j \in V_i} \kappa_{ij} \mathbf{R}_i \operatorname{skew}(\mathbf{R}_i^\top \mathbf{H}_{ij}^\top \mathbf{R}_j) \\ &\quad \sum_{j \in V_i} \kappa_{ij} \mathbf{R}_i \operatorname{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j). \end{aligned} \quad (72)$$

For 2D case, because of  $\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j = \mathbf{Z}_{ij}^\top$ , so we have:

$$\text{grad}_{\mathbf{R}_i} \sum_{j \in V_i} \log \hat{f}_{ij}(\mathbf{Z}_{ij}) = \sum_{j \in V_i} -\kappa_{ij} \mathbf{R}_i \text{skew}(\mathbf{Z}_{ij}). \quad (73)$$

So the FIM will be:

$$\mathbf{F}_{ii_1} = \mathbb{E}\left\{ \left\langle \sum_{j \in V_i} -\kappa_{ij} \mathbf{R}_i \text{skew}(\mathbf{Z}_{ij}), \mathbf{R}_i \mathbf{E} \right\rangle_{\mathbf{R}} \cdot \left\langle \sum_{j \in V_{i_1}} -\kappa_{ij} \mathbf{R}_{i_1} \text{skew}(\mathbf{Z}_{i_1 j}), \mathbf{R}_{i_1} \mathbf{E} \right\rangle_{\mathbf{R}} \right\} \quad (74)$$

For SO(2), we have:

$$\begin{aligned} & \left\langle \sum_{j \in V_i} -\kappa_{ij} \mathbf{R}_i \text{skew}(\mathbf{Z}_{ij}), \mathbf{R}_i \mathbf{E} \right\rangle_{\mathbf{R}} \\ &= \sum_{j \in V_i} \text{trace}(-\kappa_{ij} \text{skew}(\mathbf{Z}_{ij})^\top \mathbf{R}_i^\top \mathbf{R}_i \mathbf{E}) \\ &= \sum_{j \in V_i} \text{trace}(-\kappa_{ij} \text{skew}(\mathbf{Z}_{ij})^\top \mathbf{E}) \end{aligned} \quad (75)$$

Set  $\mathbf{Z}_{ij} = \begin{bmatrix} c_{ij} & -s_{ij} \\ s_{ij} & c_{ij} \end{bmatrix}$ , where  $c_{ij} = \cos(\alpha_{ij})$  and  $s_{ij} = \sin(\alpha_{ij})$ ,  $\alpha_{ij}$  is the minimal representation of rotation matrix [41], we have:

$$\text{trace}(\text{skew}(\mathbf{Z}_{ij}) \mathbf{E}) = -2\kappa_{ij} s_{ij}. \quad (76)$$

We can get (75) will be:

$$\begin{aligned} & \left\langle \sum_{j \in V_i} -\kappa_{ij} \mathbf{R}_i \text{skew}(\mathbf{Z}_{ij}), \mathbf{R}_i \mathbf{E} \right\rangle_{\mathbf{R}} \\ &= \sum_{j \in V_i} \kappa_{ij} \text{trace}\left(-\begin{pmatrix} 0 & -s_{ij} \\ s_{ij} & 0 \end{pmatrix}^\top \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\right) \\ &= \sum_{j \in V_i} \kappa_{ij} \text{trace}\left(\begin{pmatrix} 0 & -s_{ij} \\ s_{ij} & 0 \end{pmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\right) \\ &= \sum_{j \in V_i} \kappa_{ij} (-2s_{ij}) \\ &= \sum_{j \in V_i} -2\kappa_{ij} s_{ij} \end{aligned} \quad (77)$$

Introduce (77) into (186), we have:

$$\begin{aligned} \mathbf{F}_{ii_1} &= \mathbb{E}\left\{ \sum_{j \in V_i} -2\kappa_{ij} s_{ij} \sum_{j \in V_{i_1}} -2\kappa_{i_1 j_1} s_{i_1 j_1} \right\} \\ &= \sum_{j \in V_i} \sum_{j \in V_{i_1}} 4\kappa_{ij} \kappa_{i_1 j_1} \mathbb{E}\{s_{ij} s_{i_1 j_1}\} \end{aligned} \quad (78)$$

Because when the minimal representation  $\alpha_{ij}$  is small, it meets  $\alpha_{ij} \approx s_{ij}$ . Because the isotropic Langevin distribution is similar to Gaussian distribution, we can regard it as:  $s_{ij} \sim N(0, \frac{1}{2\kappa_{ij}})$ . In fact, the accurate value for  $\mathbb{E}\{s_{ij} s_{i_1 j_1}\}$  is  $\frac{I_1(2\kappa_{ij})}{2\kappa_{ij} I_0(2\kappa_{ij})}$ . When  $\kappa_{ij}$  becomes large, we will have:  $\frac{I_1(2\kappa_{ij})}{2\kappa_{ij} I_0(2\kappa_{ij})} \rightarrow \frac{1}{2\kappa_{ij}}$ . The specific derivation process is shown as follows:

We have:

$$\begin{aligned} \text{skew}(\mathbf{Z}_{ij}) &= \frac{1}{2} \left( \begin{bmatrix} 0 & -s_{ij} \\ s_{ij} & 0 \end{bmatrix} - \begin{bmatrix} 0 & -s_{ij} \\ s_{ij} & 0 \end{bmatrix}^\top \right) \\ &= \frac{1}{2} \begin{bmatrix} 0 & -2s_{ij} \\ 2s_{ij} & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & -s_{ij} \\ s_{ij} & 0 \end{bmatrix} \\ &= s_{ij} \mathbf{E} \end{aligned} \quad (79)$$

We also have:

$$s_{ij} = \frac{1}{2} \langle \text{skew}(\mathbf{Z}_{ij}), \mathbf{E} \rangle_{\mathbf{R}} \quad (80)$$

Because of  $\text{skew}(\mathbf{Z}_{ij}) = s_{ij} \mathbf{E}$  and  $s_{ij} = \frac{1}{2} \langle \text{skew}(\mathbf{Z}_{ij}), \mathbf{E} \rangle_{\mathbf{R}}$ , we have:

$$\text{skew}(\mathbf{Z}_{ij}) = \frac{1}{2} \langle \text{skew}(\mathbf{Z}_{ij}), \mathbf{E} \rangle_{\mathbf{R}} \mathbf{E}. \quad (81)$$

Combine above equations, we have:

$$\begin{aligned} & \mathbb{E}\{\|\text{skew}(\mathbf{Z}_{ij})\|_F^2\} \\ &= \mathbb{E}\left\{\left\|\frac{1}{2} \langle \text{skew}(\mathbf{Z}_{ij}), \mathbf{E} \rangle_{\mathbf{R}} \mathbf{E}\right\|_F^2\right\} \\ &= \mathbb{E}\left\{\frac{1}{4} \cdot 2 \left( \langle \text{skew}(\mathbf{Z}_{ij}), \mathbf{E} \rangle_{\mathbf{R}}^T \langle \text{skew}(\mathbf{Z}_{ij}), \mathbf{E} \rangle_{\mathbf{R}} \right)\right\} \\ &= \mathbb{E}\{2s_{ij}s_{ij}\}. \end{aligned} \quad (82)$$

So we have:  $\mathbb{E}\{s_{ij}^2\} = \frac{1}{2}\mathbb{E}\{\|\text{skew}(\mathbf{Z}_{ij})\|_F^2\}$ . Based on [13], the PDF of the isotropic Langevin distribution on  $SO(n)$  meets the following property, which is translated by the extended bi-invariance property:

$$\forall \mathbf{U}_1, \mathbf{U}_2 \in \bar{O}(n) \text{ s.t. } \det(\mathbf{U}_1 \mathbf{U}_2) = 1, \int_{SO(n)} \widehat{f}_{ij}(\mathbf{U}_1 \mathbf{Z}_{ij} \mathbf{U}_2) d\mu(\mathbf{U}_1 \mathbf{Z}_{ij} \mathbf{U}_2) = \int_{SO(n)} \widehat{f}_{ij}(\mathbf{Z}_{ij}) d\mu(\mathbf{Z}_{ij}). \quad (83)$$

Based on the isotropic Langevin density (66) and property (83), let  $\mathbf{Z} = \mathbf{U}_1 \mathbf{Z}_{ij} \mathbf{U}_2 = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$ , we have:

$$\begin{aligned} & \mathbb{E}\{s_{ij}^2\} \\ &= \frac{1}{2} \mathbb{E}\{\|\text{skew}(\mathbf{Z}_{ij})\|^2\} \\ &= \frac{1}{8} \int_{SO(2)} \|\mathbf{Z} - \mathbf{Z}^\top\|^2 f(\mathbf{Z}) d\mu(\mathbf{Z}) \\ &= \frac{\kappa_{ij}^2}{8} \frac{1}{2\pi} \int_{-\pi}^{-\pi} \left\| \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} - \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}^\top \right\|^2 f\left(\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}\right) d\theta \\ &= \frac{1}{8} \frac{1}{2\pi} \int_{-\pi}^{-\pi} \left\| \begin{bmatrix} 0 & -2\sin(\theta) \\ 2\sin(\theta) & 0 \end{bmatrix} \right\|^2 f\left(\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}\right) d\theta \\ &= \frac{1}{8} \frac{1}{2\pi} \int_{-\pi}^{-\pi} 8\sin^2(\theta) f\left(\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}\right) d\theta \\ &= \frac{1}{8} \frac{1}{2\pi} \int_{-\pi}^{-\pi} 8\sin^2(\theta) \frac{1}{c_2(\kappa_{ij})} \exp(\kappa_{ij} \text{trace}\left(\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}\right)) d\theta \\ &= \frac{1}{8} \frac{1}{2\pi} \int_{-\pi}^{-\pi} 8\sin^2(\theta) \frac{1}{c_2(\kappa_{ij})} \exp(2\kappa_{ii} \cos(\theta)) d\theta \\ &= \frac{1}{2\pi c_2(\kappa_{ij})} \int_{-\pi}^{-\pi} \sin^2(\theta) \exp(2\kappa_{ij} \cos(\theta)) d\theta \\ &= \frac{1}{2\pi I_0(2\kappa_{ij})} \int_{-\pi}^{-\pi} \sin^2(\theta) \exp(2\kappa_{ij} \cos(\theta)) d\theta \\ &= \frac{1}{2\pi I_0(2\kappa_{ij})} \int_{-\pi}^{-\pi} (1 - \cos^2(\theta)) \exp(2\kappa_{ij} \cos(\theta)) d\theta \end{aligned} \quad (84)$$

We know that:

$$\cos(0\theta) = 1, \quad \cos(2\theta) = 2\cos^2(\theta) - 1. \quad (85)$$

We get:

$$\begin{aligned}
\mathbb{E}\{s_{ij}^2\} &= \frac{1}{2\pi I_0(2\kappa_{ij})} \int_{-\pi}^{-\pi} (1 - \cos^2(\theta)) \exp(2\kappa_{ij} \cos(\theta)) d\theta \\
&= \frac{1}{I_0(2\kappa_{ij})} \frac{1}{2\pi} \int_{-\pi}^{-\pi} \left(1 - \frac{1 + \cos(2\theta)}{2}\right) \exp(2\kappa_{ij} \cos(\theta)) d\theta \\
&= \frac{1}{I_0(2\kappa_{ij})} \frac{1}{2\pi} \int_{-\pi}^{-\pi} \left(\frac{1}{2} - \frac{\cos(2\theta)}{2}\right) \exp(2\kappa_{ij} \cos(\theta)) d\theta \\
&= \frac{1}{I_0(2\kappa_{ij})} \frac{1}{2\pi} \int_{-\pi}^{-\pi} \left(\frac{\cos(0\theta)}{2} - \frac{\cos(2\theta)}{2}\right) \exp(2\kappa_{ij} \cos(\theta)) d\theta
\end{aligned} \tag{86}$$

We know that:

$$\begin{aligned}
I_0(2\kappa_{ij}) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(2\kappa_{ij} \cos(\theta)) \cos(0\theta) d\theta \\
I_1(2\kappa_{ij}) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(2\kappa_{ij} \cos(\theta)) \cos(1\theta) d\theta \\
I_1(2\kappa_{ij}) &= \kappa_{ij} (I_0(2\kappa_{ij}) - I_2(2\kappa_{ij}))
\end{aligned} \tag{87}$$

Based on  $I_1(2\kappa_{ij}) = \kappa_{ij} (I_0(2\kappa_{ij}) - I_2(2\kappa_{ij}))$  and the PDF (130), we have:

$$\mathbb{E}\{s_{ij}^2\} = \frac{1}{I_0(2\kappa_{ij})} \left( \frac{I_0(2\kappa_{ij})}{2} - \frac{I_2(2\kappa_{ij})}{2} \right) = \frac{I_1(2\kappa_{ij})}{2\kappa_{ij} I_0(2\kappa_{ij})}. \tag{88}$$

So we have:  $\mathbb{E}\{s_{ij}^2\} = \frac{1}{2} \mathbb{E}\{\|\text{skew}(\mathbf{Z}_{ij})\|_F^2\}$

Let's consider equation (78). When  $i = i_1$ , we have:

$$\mathbf{F}_{ii_1} = \sum_{j \in V_i} \sum_{j \in V_i} 4\kappa_{ij}^2 \frac{I_1(2\kappa_{ij})}{2\kappa_{ij} I_0(2\kappa_{ij})} = \sum_{j \in V_i} 2\kappa_{ij} \frac{I_1(2\kappa_{ij})}{I_0(2\kappa_{ij})} \tag{89}$$

When  $i \neq i_1$  and  $(i, i_1) \in \varepsilon$ , use  $s_{ii_1} = -s_{i_1 i}$ , we have:

$$\mathbf{F}_{ii_1} = -2\kappa_{ii_1} \frac{I_1(2\kappa_{ii_1})}{I_0(2\kappa_{ii_1})} \tag{90}$$

Finally, we have:

$$\mathbf{F}_{ii_1} = \begin{cases} \sum_{j \in V_i} 2\kappa_{ij} \frac{I_1(2\kappa_{ij})}{I_0(2\kappa_{ij})}, & i = i_1 \\ -2\kappa_{ii_1} \frac{I_1(2\kappa_{ii_1})}{I_0(2\kappa_{ii_1})}, & \{i, i_1\} \in \varepsilon \\ -2\kappa_{i_1 i} \frac{I_1(2\kappa_{i_1 i})}{I_0(2\kappa_{i_1 i})}, & \{i, i_1\} \in \varepsilon \\ 0, & \text{else} \end{cases} \tag{91}$$

It is easy to find that it is a weighted Laplacian matrix.

## 7 Synchronization on $SO(3)$

### 7.1 Synchronization on $SO(3)$

Synchronization on  $SO(3)$  is the problem of estimating a set of rotations  $\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_{n_p} \in SO(3)$  from noisy measurements of some relative rotations  $\mathbf{R}_j \mathbf{R}_i^\top$ . The operations  $\oplus$  and  $\ominus$  are defined as:  $\mathbf{R}_i \oplus \mathbf{R}_j = \mathbf{R}_i \mathbf{R}_j$  and  $\mathbf{R}_j \ominus \mathbf{R}_i = \mathbf{R}_j \mathbf{R}_i^\top$ .

In our estimation problem, the target quantities (the parameters) are the rotation matrices  $\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_{n_p} \in SO(3)$ . The natural parameter space is thus:

$$\mathcal{P} = \{SO(3) \times \dots \times SO(3)\}_{n_p} \tag{92}$$

We can get a measurement directed graph  $\mathcal{G} = ([n_p], \mathcal{F})$ , we have one measurement  $\mathbf{H}_{ij} \in SO(3)$  of a form:

$$\mathbf{H}_{ij} = \mathbf{Z}_{ij} \mathbf{R}_j \mathbf{R}_i^\top \tag{93}$$

where  $\mathbf{Z}_{ij}$  is a random value whose distributed function  $\hat{f}_{ij} : SO(3) \rightarrow \mathbb{R}^+$  meets an isotropic Langevin distribution:

$$\begin{aligned}\hat{f}(\mathbf{Z}_{ij}) &= \frac{1}{c_n(\kappa_{ij})} \exp(\kappa_{ij} \operatorname{trace}(\mathbf{Z}_{ij})) \\ c_3(\kappa_{ij}) &= \exp(\kappa_{ij})(I_0(2\kappa_{ij}) - I_1(2\kappa_{ij})) \\ I_0(2\kappa_{ij}) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(2\kappa_{ij} \cos(\theta)) d\theta \\ I_1(2\kappa_{ij}) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(2\kappa_{ij} \cos(\theta)) \cos(\theta) d\theta\end{aligned}\tag{94}$$

Hence,

$$L(\mathbf{y}; \boldsymbol{\theta}) = \frac{1}{2} \sum_{(i,j) \in \mathcal{F}} \log \hat{f}_{ij}(\mathbf{H}_{ij} \mathbf{R}_i \mathbf{R}_j^\top) = \frac{1}{2} \sum_{(i,j) \in \mathcal{F}} \log \left( \frac{1}{c_3(\kappa_{ij})} \exp(\kappa_{ij} \operatorname{trace}(\mathbf{H}_{ij} \mathbf{R}_i \mathbf{R}_j^\top)) \right)\tag{95}$$

## 7.2 FIM for $SO(3)$

The synchronization on space  $SO(3)$  is the problem of estimating a set of variables (nodes in the graph except the anchor  $R_0$ )  $\theta = \{\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_{n_p}\} \in SO(3)$  from noisy measurements  $\mathbf{H}_{ij}$  of relative information between  $\mathbf{R}_i$  and  $\mathbf{R}_j$  (edges in the rotation graph). The orthonormal basis  $\mathbf{E} = (\mathbf{E}_{0,1}^R, \mathbf{E}_{0,2}^R, \mathbf{E}_{0,3}^R, \dots, \mathbf{E}_{n_p,1}^R, \mathbf{E}_{n_p,2}^R, \mathbf{E}_{n_p,3}^R)$  of the tangent space  $\mathcal{T}_{\mathbf{R}} SO(3)$  as:  $\mathbf{R}_1 \mathbf{E}_1, \mathbf{R}_1 \mathbf{E}_2, \mathbf{R}_1 \mathbf{E}_3, \dots, \mathbf{R}_{n_p} \mathbf{E}_1, \mathbf{R}_{n_p} \mathbf{E}_2, \mathbf{R}_{n_p} \mathbf{E}_3$ , where:

$$\mathbf{E}_{i,k}^R = (\mathbf{0}, \dots, \mathbf{0}, \underbrace{\mathbf{R}_i \mathbf{E}_k}_{3j+k-\text{th}}, \mathbf{0}, \dots, \mathbf{0})_{3 \times 3(n_p+1)}\tag{96}$$

$$\mathbf{E}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \mathbf{E}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \mathbf{E}_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\tag{97}$$

It is noted that, in many applications,  $\sqrt{2}\mathbf{E}_1, \sqrt{2}\mathbf{E}_2, \sqrt{2}\mathbf{E}_3$  are used as the bases of the manifold  $SO(3)$ , which are not the orthonormal bases. In that case, the weight for rotation group part in the final FIM will be  $2\omega_{ij}$ , where  $\omega_{ij}$  will be shown in (120). For the isometric coordinate transformation, the FIM is invariant.

The inner product is defined by  $\langle \operatorname{grad} L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_{i,k}^R \rangle_{\mathbf{R}} = \operatorname{trace}(\operatorname{grad} L(\mathbf{y}; \boldsymbol{\theta})^\top \cdot \mathbf{E}_{i,k}^R)$ . The FIM of the estimation problem w.r.t. the basis  $e$  is defined as:

$$\begin{aligned}\mathbf{F}_{ij,kl} &= \mathbb{E}\{\langle \operatorname{grad} L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_{i,k}^R \rangle_{\mathbf{R}} \cdot \langle \operatorname{grad} L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_{j,l}^R \rangle_{\mathbf{R}}\} \\ &= \mathbb{E}\{\langle \operatorname{grad}_{\mathbf{R}_i} L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{R}_i \mathbf{E}_k \rangle_{\mathbf{R}} \cdot \langle \operatorname{grad}_{\mathbf{R}_{i_1}} L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{R}_j \mathbf{E}_l \rangle_{\mathbf{R}}\}\end{aligned}\tag{98}$$

where  $\mathbf{F}_{ij,kl} \in \mathbb{R}$  means the  $(k, l)$ -th element of the  $(i, j)$ -th block of the FIM.

Based on the PDF (67) for the measurement error  $\mathbf{Z}_{ij}$ , solving (129) for  $\mathbf{Z}_{ij}$  and substituting into the PDF (67), we have:

$$\log \hat{f}_{ij}(\mathbf{Z}_{ij}) = -\log(c_n(\kappa_{ij})) + \kappa_{ij} \operatorname{trace}(\mathbf{H}_{ij} \mathbf{R}_i \mathbf{R}_j^\top).\tag{99}$$

Since the inner product on  $\mathbb{R}^{n \times n}$  is  $\langle \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2 \rangle_{\mathbf{R}} = \operatorname{trace}(\boldsymbol{\Omega}_1^\top \boldsymbol{\Omega}_2)$ , the gradient (99) implies that the Euclidean gradients of the log-likelihood function are:

$$\nabla_{\mathbf{R}_i} \log \hat{f}_{ij}(\mathbf{Z}_{ij}) = \kappa_{ij} \mathbf{H}_{ij}^\top \mathbf{R}_j.\tag{100}$$

Finally, based on the gradient on  $SO(n)$  (1), we have the Riemannian gradient of the log-likelihood function is:

$$\begin{aligned}\operatorname{grad}_{\mathbf{R}_i} \sum_{j \in V_i} \log \hat{f}_{ij}(\mathbf{Z}_{ij}) &= \sum_{j \in V_i} \kappa_{ij} \mathbf{R}_i \operatorname{skew}(\mathbf{R}_i^\top \mathbf{H}_{ij}^\top \mathbf{R}_j) \\ &= \sum_{j \in V_i} \kappa_{ij} \mathbf{R}_i \operatorname{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j).\end{aligned}\tag{101}$$

So the FIM will be:

$$\mathbf{F}_{ii_1,kl} = \mathbb{E}\{\langle \sum_{j \in V_i} \kappa_{ij} \mathbf{R}_i \operatorname{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j), \mathbf{R}_i \mathbf{E}_k \rangle_{\mathbf{R}} \cdot \langle \sum_{j \in V_i} \kappa_{ij} \mathbf{R}_i \operatorname{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j), \mathbf{R}_{i_1} \mathbf{E}_l \rangle_{\mathbf{R}}\}\tag{102}$$

For  $SO(3)$ , we have:

$$\begin{aligned}
& < \sum_{j \in V_i} \kappa_{ij} \mathbf{R}_i \text{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j), \mathbf{R}_i \mathbf{E}_k >_{\mathbf{R}} \\
& = < \sum_{j \in V_i} \kappa_{ij} \mathbf{R}_i \text{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j) \mathbf{R}_i^\top, \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top >_{\mathbf{R}} \\
& = \sum_{j \in V_i} \kappa_{ij} < \mathbf{R}_i \text{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j) \mathbf{R}_i^\top, \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top >_{\mathbf{R}}
\end{aligned} \tag{103}$$

Let's consider the following mapping  $G_{ij}(\mathbf{Z}_{ij}) = \mathbf{R}_i \text{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j) \mathbf{R}_i^\top$  and its three bases  $\mathbf{R}_i \mathbf{E}_1 \mathbf{R}_i^\top$ ,  $\mathbf{R}_i \mathbf{E}_2 \mathbf{R}_i^\top$ , and  $\mathbf{R}_i \mathbf{E}_3 \mathbf{R}_i^\top$ . We have:

$$\begin{aligned}
\mathbf{R}_i \text{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j) \mathbf{R}_i^\top &= \sum_{k=1}^3 < \mathbf{R}_i \text{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j) \mathbf{R}_i^\top, \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top >_{\mathbf{R}} \cdot \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top \\
&= \sum_{k=1}^3 \Phi_k^{ij} \cdot \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top
\end{aligned} \tag{104}$$

So the minimal representations of the mapping  $G_{ij}(\mathbf{Z}_{ij})$  are

$$\Phi_k^{ij} = < \mathbf{R}_i \text{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j) \mathbf{R}_i^\top, \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top >_{\mathbf{R}} \tag{105}$$

Let's discuss the isotropic Langevin distribution first. Based on [13], the PDF of the isotropic Langevin distribution on  $SO(n)$  meets the following property, which is translated by the extended bi-invariance property:

$$\begin{aligned}
& \forall \mathbf{U}_1, \mathbf{U}_2 \in \bar{O}(n) \text{ s.t. } \det(\mathbf{U}_1 \mathbf{U}_2) = 1, \\
& \int_{SO(n)} \widehat{f}_{ij}(\mathbf{Z}) d\mu(\mathbf{Z}) = \int_{SO(n)} \widehat{f}_{ij}(\mathbf{Z}_{ij}) d\mu(\mathbf{Z}_{ij}).
\end{aligned} \tag{106}$$

where  $\mathbf{Z} = \mathbf{U}_1 \mathbf{Z}_{ij} \mathbf{U}_2 = \begin{bmatrix} \mathbf{R}(\theta) & 0 \\ 0 & 1 \end{bmatrix}$ ,  $\mathbf{R}(\theta) \triangleq \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \in SO(3)$

So the trace appearing in the isotropic Langevin density (94) can be equivalently expressed as:

$$\text{trace}(\mathbf{Z}) = \text{trace}(\mathbf{U}^\top \mathbf{Z}_{ij} \mathbf{U}) = 1 + 2 \cos(\theta) \tag{107}$$

It is easy to know that  $\theta$  is independent of the choose of basis.  $\theta$  is depend on the magnitude of the rotation. It is isotropic based on the basis. This property is also called invariance.

$$\mathbb{E}\{\Phi_k^{ij}\} = \int_{SO(3)} < \mathbf{R}_i \text{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j) \mathbf{R}_i^\top, \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top >_{\mathbf{R}} \widehat{f}_{ij}(\mathbf{Z}_{ij}) d\mu(\mathbf{Z}_{ij}) \tag{108}$$

The mapping  $\text{skew}(\mathbf{Z}_{ij})$  satisfies:  $\forall \mathbf{U} \in O(n)$ ,  $\text{skew}(\mathbf{U} \mathbf{Z}_{ij} \mathbf{U}^\top) = \mathbf{U} \text{skew}(\mathbf{Z}_{ij}) \mathbf{U}^\top$ . So we have:

$$\begin{aligned}
\mathbb{E}\{\Phi_k^{ij}\} &= \mathbb{E}\{ < \mathbf{R}_i \text{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j) \mathbf{R}_i^\top, \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top >_{\mathbf{R}} \} \\
&= \mathbb{E}\{ < \mathbf{R}_i \mathbf{R}_j^\top \text{skew}(\mathbf{Z}_{ij}^\top) \mathbf{R}_j \mathbf{R}_i^\top, \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top >_{\mathbf{R}} \}
\end{aligned} \tag{109}$$

**Lemma 2.**  $\mathbf{X}$  is random matrix as well as  $\mathbf{A}$  and  $\mathbf{B}$  are two constant matrices, we have  $\mathbb{E}\{\mathbf{A} \mathbf{X} \mathbf{B}\} = \mathbf{A} \mathbb{E}\{\mathbf{X}\} \mathbf{B}$ .

**Lemma 3.**  $\mathbf{X}$  and  $\mathbf{Y}$  are two random matrices, we have  $\mathbb{E}\{\mathbf{X} \pm \mathbf{Y}\} = \mathbb{E}\{\mathbf{X}\} \pm \mathbb{E}\{\mathbf{Y}\}$ .

Based on above two lemmas, it is easy to know that  $\mathbb{E}\{\Phi_k^{ij}\} = \mathbb{E}\{ < \mathbf{R}_i \mathbf{R}_j^\top \text{skew}(\mathbf{Z}_{ij}^\top) \mathbf{R}_j \mathbf{R}_i^\top, \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top >_{\mathbf{R}} \} = < \mathbf{R}_i \mathbf{R}_j^\top \mathbb{E}\{\text{skew}(\mathbf{Z}_{ij}^\top)\} \mathbf{R}_j \mathbf{R}_i^\top, \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top >_{\mathbf{R}} = < \mathbf{R}_i \mathbf{R}_j^\top \text{skew}(\mathbb{E}\{\mathbf{Z}_{ij}^\top\}) \mathbf{R}_j \mathbf{R}_i^\top, \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top >_{\mathbf{R}} = < \mathbf{R}_i \mathbf{R}_j^\top \text{skew}(\mathbf{I}_{2 \times 2}) \mathbf{R}_j \mathbf{R}_i^\top, \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top >_{\mathbf{R}} = 0$ , so we have  $\mathbb{E}\{\Phi_k^{ij}\} = \mathbb{E}\{\Phi_l^{ij}\} = 0$ .

**Lemma 4.** There exists a permutation matrix  $\mathbf{P}_{kl} \in O(n)$  such that  $\mathbf{P}_{kl}^\top \mathbf{E}_k \mathbf{P}_{kl} = \mathbf{E}_l$  and  $\mathbf{P}_{kl}^\top \mathbf{E}_l \mathbf{P}_{kl} = -\mathbf{E}_k$ .

Based on Lemma 4, we have:

$$\begin{aligned}
& \mathbb{E}\{\Phi_k^{ij}\Phi_l^{ij}\} \\
&= \mathbb{E}\{<\mathbf{R}_i\mathbf{R}_j^\top \text{skew}(\mathbf{Z}_{ij}^\top)\mathbf{R}_j\mathbf{R}_i^\top, \mathbf{R}_i\mathbf{E}_k\mathbf{R}_i^\top>_{\mathbf{R}} <\mathbf{R}_i\mathbf{R}_j^\top \text{skew}(\mathbf{Z}_{ij}^\top)\mathbf{R}_j\mathbf{R}_i^\top, \mathbf{R}_i\mathbf{E}_l\mathbf{R}_i^\top>_{\mathbf{R}}\} \\
&= \int_{SO(3)} <\mathbf{R}_i\mathbf{R}_j^\top \text{skew}(\mathbf{Z}_{ij}^\top)\mathbf{R}_j\mathbf{R}_i^\top, \mathbf{R}_i\mathbf{E}_k\mathbf{R}_i^\top>_{\mathbf{R}} <\mathbf{R}_i\mathbf{R}_j^\top \text{skew}(\mathbf{Z}_{ij}^\top)\mathbf{R}_j\mathbf{R}_i^\top, \mathbf{R}_i\mathbf{E}_l\mathbf{R}_i^\top>_{\mathbf{R}} \hat{f}_{ij}(\mathbf{Z}_{ij})d\mu(\mathbf{Z}_{ij}) \\
&= \int_{SO(3)} - <\mathbf{R}_i\mathbf{R}_j^\top \text{skew}(\mathbf{Z}_{ij}^\top)\mathbf{R}_j\mathbf{R}_i^\top, \mathbf{R}_i\mathbf{P}_{kl}^\top\mathbf{E}_l\mathbf{P}_{kl}\mathbf{R}_i^\top>_{\mathbf{R}} <\mathbf{R}_i\mathbf{R}_j^\top \text{skew}(\mathbf{Z}_{ij}^\top)\mathbf{R}_j\mathbf{R}_i^\top, \mathbf{R}_i\mathbf{P}_{kl}^\top\mathbf{E}_k\mathbf{P}_{kl}\mathbf{R}_i^\top>_{\mathbf{R}} \\
&\quad \hat{f}_{ij}(\mathbf{Z}_{ij})d\mu(\mathbf{Z}_{ij}) \\
&= \int_{SO(3)} - <\mathbf{R}_i\mathbf{R}_j^\top \text{skew}(\mathbf{Z}_{ij}^\top)\mathbf{R}_j\mathbf{R}_i^\top, \mathbf{R}_i\mathbf{P}_{kl}^\top\mathbf{R}_i^\top\mathbf{R}_i\mathbf{E}_l\mathbf{R}_i\mathbf{R}_i^\top\mathbf{P}_{kl}\mathbf{R}_i^\top>_{\mathbf{R}} \\
&\quad <\mathbf{R}_i\mathbf{R}_j^\top \text{skew}(\mathbf{Z}_{ij}^\top)\mathbf{R}_j\mathbf{R}_i^\top, \mathbf{R}_i\mathbf{P}_{kl}^\top\mathbf{R}_i^\top\mathbf{R}_i\mathbf{E}_k\mathbf{R}_i\mathbf{R}_i^\top\mathbf{P}_{kl}\mathbf{R}_i^\top>_{\mathbf{R}} \hat{f}_{ij}(\mathbf{Z}_{ij})d\mu(\mathbf{Z}_{ij}) \\
&= \int_{SO(3)} - <\mathbf{R}_j^\top \text{skew}(\mathbf{Z}_{ij}^\top)\mathbf{R}_j, \mathbf{P}_{kl}^\top\mathbf{R}_i^\top\mathbf{R}_i\mathbf{E}_l\mathbf{R}_i\mathbf{R}_i^\top\mathbf{P}_{kl}>_{\mathbf{R}} \\
&\quad <\mathbf{R}_j^\top \text{skew}(\mathbf{Z}_{ij}^\top)\mathbf{R}_j, \mathbf{P}_{kl}^\top\mathbf{R}_i^\top\mathbf{R}_i\mathbf{E}_k\mathbf{R}_i\mathbf{R}_i^\top\mathbf{P}_{kl}>_{\mathbf{R}} \hat{f}_{ij}(\mathbf{Z}_{ij})d\mu(\mathbf{Z}_{ij}) \\
&= \int_{SO(3)} - <\mathbf{R}_i\mathbf{P}_{kl}\mathbf{R}_j^\top \text{skew}(\mathbf{Z}_{ij}^\top)\mathbf{R}_j\mathbf{P}_{kl}^\top\mathbf{R}_i^\top, \mathbf{R}_i\mathbf{E}_l\mathbf{R}_i>_{\mathbf{R}} \\
&\quad <\mathbf{R}_i\mathbf{P}_{kl}\mathbf{R}_j^\top \text{skew}(\mathbf{Z}_{ij}^\top)\mathbf{R}_j\mathbf{P}_{kl}^\top\mathbf{R}_i^\top, \mathbf{R}_i\mathbf{E}_k\mathbf{R}_i>_{\mathbf{R}} \hat{f}_{ij}(\mathbf{Z}_{ij})d\mu(\mathbf{Z}_{ij}) \\
&= \int_{SO(3)} - <\mathbf{R}_i\text{skew}(\mathbf{P}_{kl}\mathbf{R}_j^\top\mathbf{Z}_{ij}^\top\mathbf{R}_j\mathbf{P}_{kl}^\top)\mathbf{R}_i^\top, \mathbf{R}_i\mathbf{E}_l\mathbf{R}_i>_{\mathbf{R}} \\
&\quad <\mathbf{R}_i\text{skew}(\mathbf{P}_{kl}\mathbf{R}_j^\top\mathbf{Z}_{ij}^\top\mathbf{R}_j\mathbf{P}_{kl}^\top)\mathbf{R}_i^\top, \mathbf{R}_i\mathbf{E}_k\mathbf{R}_i>_{\mathbf{R}} \hat{f}_{ij}(\mathbf{Z}_{ij})d\mu(\mathbf{Z}_{ij}) \\
&= \int_{SO(3)} - <\mathbf{R}_i\mathbf{R}_j^\top\mathbf{R}_j\text{skew}(\mathbf{P}_{kl}\mathbf{R}_j^\top\mathbf{Z}_{ij}^\top\mathbf{R}_j\mathbf{P}_{kl}^\top)\mathbf{R}_j\mathbf{R}_j^\top\mathbf{R}_i^\top, \mathbf{R}_i\mathbf{E}_l\mathbf{R}_i>_{\mathbf{R}} \\
&\quad <\mathbf{R}_i\mathbf{R}_j^\top\mathbf{R}_j\text{skew}(\mathbf{P}_{kl}\mathbf{R}_j^\top\mathbf{Z}_{ij}^\top\mathbf{R}_j\mathbf{P}_{kl}^\top)\mathbf{R}_j\mathbf{R}_j^\top\mathbf{R}_i^\top, \mathbf{R}_i\mathbf{E}_k\mathbf{R}_i>_{\mathbf{R}} \hat{f}_{ij}(\mathbf{Z}_{ij})d\mu(\mathbf{Z}_{ij}) \\
&= \int_{SO(3)} - <\mathbf{R}_i\mathbf{R}_j^\top \text{skew}(\mathbf{R}_j\mathbf{P}_{kl}\mathbf{R}_j^\top\mathbf{Z}_{ij}^\top\mathbf{R}_j\mathbf{P}_{kl}^\top\mathbf{R}_j^\top)\mathbf{R}_j\mathbf{R}_i^\top, \mathbf{R}_i\mathbf{E}_l\mathbf{R}_i>_{\mathbf{R}} \\
&\quad <\mathbf{R}_i\mathbf{R}_j^\top \text{skew}(\mathbf{R}_j\mathbf{P}_{kl}\mathbf{R}_j^\top\mathbf{Z}_{ij}^\top\mathbf{R}_j\mathbf{P}_{kl}^\top\mathbf{R}_j^\top)\mathbf{R}_j\mathbf{R}_i^\top, \mathbf{R}_i\mathbf{E}_k\mathbf{R}_i>_{\mathbf{R}} \hat{f}_{ij}(\mathbf{Z}_{ij})d\mu(\mathbf{Z}_{ij}) \tag{110}
\end{aligned}$$

Because the isotropic Langvin distribution follows the bi-invariance, we can replace the variable by  $\mathbf{R}_j\mathbf{P}_{kl}\mathbf{R}_j^\top\mathbf{Z}_{ij}^\top\mathbf{R}_j\mathbf{P}_{kl}^\top\mathbf{R}_j^\top$ .

So we have:

$$\begin{aligned}
& \mathbb{E}\{\Phi_k^{ij}\Phi_l^{ij}\} \\
&= \int_{SO(3)} - <\mathbf{R}_i\mathbf{R}_j^\top \text{skew}(\mathbf{R}_j\mathbf{P}_{kl}\mathbf{R}_j^\top\mathbf{Z}_{ij}^\top\mathbf{R}_j\mathbf{P}_{kl}^\top\mathbf{R}_j^\top)\mathbf{R}_j\mathbf{R}_i^\top, \mathbf{R}_i\mathbf{E}_l\mathbf{R}_i>_{\mathbf{R}} <\mathbf{R}_i\mathbf{R}_j^\top \text{skew}(\mathbf{R}_j\mathbf{P}_{kl}\mathbf{R}_j^\top\mathbf{Z}_{ij}^\top\mathbf{R}_j\mathbf{P}_{kl}^\top\mathbf{R}_j^\top)\mathbf{R}_j\mathbf{R}_i^\top, \mathbf{R}_i\mathbf{E}_k\mathbf{R}_i>_{\mathbf{R}} \hat{f}_{ij}(\mathbf{R}_j\mathbf{P}_{kl}\mathbf{R}_j^\top\mathbf{Z}_{ij}^\top\mathbf{R}_j\mathbf{P}_{kl}^\top\mathbf{R}_j^\top)d\mu(\mathbf{R}_j\mathbf{P}_{kl}\mathbf{R}_j^\top\mathbf{Z}_{ij}^\top\mathbf{R}_j\mathbf{P}_{kl}^\top\mathbf{R}_j^\top) \tag{111} \\
&= -\mathbb{E}\{\Phi_l^{ij}\Phi_k^{ij}\} \\
&= -\mathbb{E}\{\Phi_k^{ij}\Phi_l^{ij}\} \\
&= 0
\end{aligned}$$

which means that  $\Phi_l^{ij}$  and  $\Phi_k^{ij}$  are independent. Similarly, we can get:  $\mathbb{E}\{\Phi_l^{ij2}\} = \mathbb{E}\{\Phi_k^{ij2}\}$ . We can also easily get the following Lemma 5 based on the above arguments. So we omit its proving process and a statically verification is shown in <https://github.com/cyb1212/A-sub mittedpaper/blob/master/code.zip>.

Based on  $\mathbb{E}\{\Phi_l^{ij2}\} = \mathbb{E}\{\Phi_k^{ij2}\}$ , based on Lemma 5, we have:

**Lemma 5.** If  $\mathbf{R}_i$ ,  $\mathbf{R}_{i_1}$ , and  $\mathbf{R}_{i_2} \in SO(n)$  and  $\mathbf{Z}_{ij} \sim \text{Lang}(\mathbf{I}_{n \times n}, \kappa_{ij})$ , we have:  $\mathbb{E}\{\bar{\Phi}\} = \mathbb{E}\{\hat{\Phi}\} = 0$  and  $\mathbb{E}\{\bar{\Phi}^2\} = \mathbb{E}\{\hat{\Phi}^2\}$ .  $\bar{\Phi} = <\text{skew}(\mathbf{Z}_{ij}^\top), \mathbf{E}_k>_{\mathbf{R}}$ ,  $\hat{\Phi} = <\mathbf{R}_i\text{skew}(\mathbf{R}_{i_1}\mathbf{Z}_{ij}^\top\mathbf{R}_{i_1}^\top)\mathbf{R}_i^\top, \mathbf{R}_{i_2}\mathbf{E}_k\mathbf{R}_{i_2}^\top>_{\mathbf{R}}$ .

$$\begin{aligned}
& \mathbb{E}\{\|\text{skew}(\mathbf{Z}_{ij})\|_F^2\} \\
&= \mathbb{E}\left\{\left\|\sum_{k=1}^3 \langle \text{skew}(\mathbf{Z}_{ij}), \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top \rangle_{\mathbf{R}} \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top\right\|_F^2\right\} \\
&= \mathbb{E}\left\{\sum_{k=1}^3 \langle \text{skew}(\mathbf{Z}_{ij}), \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top \rangle_{\mathbf{R}}^2\right\} \\
&= \mathbb{E}\left\{\sum_{k=1}^3 \Phi_k^{ij2}\right\} \\
&= \mathbb{E}\{3\Phi_k^{ij2}\}
\end{aligned} \tag{112}$$

So we have:  $\mathbb{E}\{\Phi_k^{ij2}\} = \frac{1}{3}\mathbb{E}\{\|\text{skew}(\mathbf{Z}_{ij})\|_F^2\}$   
(1) When  $i = i_1$  and  $k = l$ , we have:

$$\begin{aligned}
F_{ii,kk} &= \mathbb{E}\left\{\sum_{j \in V_i} -\kappa_{ij} \langle \text{skew}(\mathbf{Z}_{ij}), \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top \rangle_{\mathbf{R}} \sum_{j \in V_i} -\kappa_{ij} \langle \text{skew}(\mathbf{Z}_{ij}), \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top \rangle_{\mathbf{R}}\right\} \\
&= \sum_{j \in V_i} \sum_{j \in V_i} \kappa_{ij}^2 \mathbb{E}\{\langle \text{skew}(\mathbf{Z}_{ij}), \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top \rangle_{\mathbf{R}}^2\} \\
&= \sum_{j \in V_i} \kappa_{ij}^2 \mathbb{E}\{\Phi_k^2\} \\
&= \sum_{j \in V_i} \kappa_{ij}^2 \frac{1}{3} \mathbb{E}\{\|\text{skew}(\mathbf{Z}_{ij})\|_F^2\} \\
&= \sum_{j \in V_i} \frac{\kappa_{ij}^2}{12} \mathbb{E}\{\|\mathbf{Z} - \mathbf{Z}^\top\|_F^2\} \\
&= \sum_{j \in V_i} \frac{1}{3} \omega_{ii,kk}
\end{aligned} \tag{113}$$

Let's compute  $\omega_{ii,kk}$  based on the isotropic Langevin density (94):

$$\begin{aligned}
\omega_{ii,kk} &= \frac{\kappa_{ij}^2}{4} \int_{SO(3)} \|\mathbf{Z} - \mathbf{Z}^\top\|_F^2 \widehat{f}_{ii}(\mathbf{Z}) d\mu(\mathbf{Z}) \\
&= \frac{\kappa_{ij}^2}{4} \frac{1}{2\pi} \int_{-\pi}^{-\pi} \left\| \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}^\top \right\|^2 \widehat{f}_{ii} \left( \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \\
&\quad (1 - \cos(\theta)) d\theta \\
&= \frac{\kappa_{ij}^2}{4} \frac{1}{2\pi} \int_{-\pi}^{-\pi} \left\| \begin{bmatrix} 0 & -2\sin(\theta) & 0 \\ 2\sin(\theta) & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right\|^2 \widehat{f}_{ii} \left( \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) (1 - \cos(\theta)) d\theta \\
&= \frac{\kappa_{ii}^2}{4} \frac{1}{2\pi} \int_{-\pi}^{-\pi} 8\sin^2(\theta) \widehat{f}_{ii} \left( \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) (1 - \cos(\theta)) d\theta \\
&= \frac{\kappa_{ii}^2}{4} \frac{1}{2\pi} \int_{-\pi}^{-\pi} 8\sin^2(\theta) \frac{1}{c_3(\kappa_{ii})} \exp(\kappa_{ii}\text{trace} \left( \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \right)) (1 - \cos(\theta)) d\theta \\
&= \frac{\kappa_{ii}^2}{4} \frac{1}{2\pi} \int_{-\pi}^{-\pi} 8\sin^2(\theta) \frac{1}{c_3(\kappa_{ii})} \exp(2\kappa_{ii}\cos(\theta) + \kappa_{ii})(1 - \cos(\theta)) d\theta \\
&= \frac{2\kappa_{ii}^2}{2\pi c_3(\kappa_{ii})} \int_{-\pi}^{-\pi} \sin^2(\theta) \exp(2\kappa_{ii}\cos(\theta) + \kappa_{ii})(1 - \cos(\theta)) d\theta \\
&= \frac{2\kappa_{ii}^2}{2\pi \exp(\kappa_{ii})(I_0(2\kappa_{ii}) - I_1(2\kappa_{ii}))} \int_{-\pi}^{-\pi} \sin^2(\theta) \exp(2\kappa_{ii}\cos(\theta) + \kappa_{ii})(1 - \cos(\theta)) d\theta \\
&= \frac{2\kappa_{ii}^2}{2\pi(I_0(2\kappa_{ii}) - I_1(2\kappa_{ii}))} \int_{-\pi}^{-\pi} \sin^2(\theta)(1 - \cos(\theta)) \exp(2\kappa_{ii}\cos(\theta)) d\theta \\
&= \frac{2\kappa_{ii}^2}{I_0(2\kappa_{ii}) - I_1(2\kappa_{ii})} \frac{1}{2\pi} \int_{-\pi}^{-\pi} (1 - \cos^2(\theta))(1 - \cos(\theta)) \exp(2\kappa_{ii}\cos(\theta)) d\theta \\
&= \frac{2\kappa_{ii}^2}{I_0(2\kappa_{ii}) - I_1(2\kappa_{ii})} \frac{1}{2\pi} \int_{-\pi}^{-\pi} (1 - \cos(\theta) - \cos^2(\theta) + \cos^3(\theta)) \exp(2\kappa_{ii}\cos(\theta)) d\theta
\end{aligned} \tag{114}$$

We know that:

$$\begin{aligned}
\cos(0\theta) &= 1 \\
\cos(1\theta) &= \cos(\theta) \\
\cos(2\theta) &= 2\cos^2(\theta) - 1 \\
\cos(3\theta) &= 4\cos^3(\theta) - 3\cos(\theta)
\end{aligned} \tag{115}$$

We get:

$$\begin{aligned}
&\frac{2\kappa_{ii}^2}{I_0(2\kappa_{ii}) - I_1(2\kappa_{ii})} \frac{1}{2\pi} \int_{-\pi}^{-\pi} (1 - \cos(\theta) - \cos^2(\theta) + \cos^3(\theta)) \exp(2\kappa_{ii}\cos(\theta)) d\theta \\
&= \frac{2\kappa_{ii}^2}{I_0(2\kappa_{ii}) - I_1(2\kappa_{ii})} \frac{1}{2\pi} \int_{-\pi}^{-\pi} (1 - \cos(\theta) - \frac{1 + \cos(2\theta)}{2} + \frac{\cos(3\theta) + 3\cos(\theta)}{4}) \exp(2\kappa_{ii}\cos(\theta)) d\theta \\
&= \frac{2\kappa_{ii}^2}{I_0(2\kappa_{ii}) - I_1(2\kappa_{ii})} \frac{1}{2\pi} \int_{-\pi}^{-\pi} (\frac{1}{2} - \frac{1}{4}\cos(\theta) - \frac{\cos(2\theta)}{2} + \frac{\cos(3\theta)}{4}) \exp(2\kappa_{ii}\cos(\theta)) d\theta \\
&= \frac{2\kappa_{ii}^2}{I_0(2\kappa_{ii}) - I_1(2\kappa_{ii})} \frac{1}{2\pi} \int_{-\pi}^{-\pi} (\frac{1}{2}\cos(0\theta) - \frac{1}{4}\cos(1\theta) - \frac{\cos(2\theta)}{2} + \frac{\cos(3\theta)}{4}) \exp(2\kappa_{ii}\cos(\theta)) d\theta
\end{aligned} \tag{116}$$

We know that:

$$\begin{aligned}
I_0(2\kappa_{ii}) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(2\kappa_{ii} \cos(\theta)) \cos(0\theta) d\theta \\
I_1(2\kappa_{ii}) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(2\kappa_{ii} \cos(\theta)) \cos(1\theta) d\theta \\
I_2(2\kappa_{ii}) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(2\kappa_{ii} \cos(\theta)) \cos(2\theta) d\theta \\
I_3(2\kappa_{ii}) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(2\kappa_{ii} \cos(\theta)) \cos(3\theta) d\theta
\end{aligned} \tag{117}$$

Based on (117), we have:

$$\begin{aligned}
&\frac{2\kappa_{ii}^2}{I_0(2\kappa_{ii}) - I_1(2\kappa_{ii})} \frac{1}{2\pi} \int_{-\pi}^{-\pi} \left( \frac{1}{2} \cos(0\theta) - \frac{1}{4} \cos(1\theta) - \frac{\cos(2\theta)}{2} + \frac{\cos(3\theta)}{4} \right) \exp(2\kappa_{ii} \cos(\theta)) d\theta \\
&= \frac{2\kappa_{ii}^2}{I_0(2\kappa_{ii}) - I_1(2\kappa_{ii})} \left( \frac{1}{2} I_0(2\kappa_{ii}) - \frac{1}{4} I_1(2\kappa_{ii}) - \frac{I_2(2\kappa_{ii})}{2} + \frac{I_3(2\kappa_{ii})}{4} \right) \\
&= \frac{\kappa_{ii}^2 (2I_0(2\kappa_{ii}) - I_1(2\kappa_{ii}) - 2I_2(2\kappa_{ii}) + I_3(2\kappa_{ii}))}{2I_0(2\kappa_{ii}) - 2I_1(2\kappa_{ii})}
\end{aligned} \tag{118}$$

So we have:

$$F_{ii,kk} = F_{ii,ll} = \frac{1}{3} \sum_{j \in V_i} \frac{\kappa_{ii}^2 (2I_0(2\kappa_{ii}) - I_1(2\kappa_{ii}) - 2I_2(2\kappa_{ii}) + I_3(2\kappa_{ii}))}{2I_0(2\kappa_{ii}) - 2I_1(2\kappa_{ii})}, k \text{ and } l \in \{1, 2, 3\} \tag{119}$$

(2) When  $i = i_1$  and  $k \neq l$ , because of independence of  $\mathbb{E}\{\Phi_k \Phi_l\} = \mathbb{E}\{\Phi_k\} \mathbb{E}\{\Phi_l\} = 0$ , we have  $F_{ii,kl} = 0$ ,

So we can get when  $i = i_1$

$$F_{ii} = \frac{1}{3} \left( \sum_{j \in V_i} \frac{\kappa_{ii}^2 (2I_0(2\kappa_{ii}) - I_1(2\kappa_{ii}) - 2I_2(2\kappa_{ii}) + I_3(2\kappa_{ii}))}{2I_0(2\kappa_{ii}) - 2I_1(2\kappa_{ii})} \right) \mathbf{I}_{3 \times 3} \tag{120}$$

(3) When  $i \neq i_1$ ,  $(i, i_1) \in \mathcal{E}$ , it is easy to know that  $\mathbf{Z}_{ij} = \mathbf{Z}_{ji}^\top$  and their corresponding least representations meet  $\Phi_k^{ij}(\mathbf{Z}_{ij}) = -\Phi_k^{ij}(\mathbf{Z}_{ji})$ .

$$\begin{aligned}
F_{ii_1,kl} &= \mathbb{E}\{-k_{ii_1} < \mathbf{R}_i \text{skew}(\mathbf{R}_{i_1}^\top \mathbf{Z}_{ii_1}^\top \mathbf{R}_{i_1}) \mathbf{R}_i^\top, \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top >_{\mathbf{R}} \\
&\quad - k_{ii_1} < \mathbf{R}_i \text{skew}(\mathbf{R}_{i_1}^\top \mathbf{Z}_{ii_1}^\top \mathbf{R}_{i_1}) \mathbf{R}_i^\top, \mathbf{R}_i \mathbf{E}_l \mathbf{R}_i^\top >_{\mathbf{R}}\}
\end{aligned} \tag{121}$$

When  $k = l$ , based on (113), we have:

$$\begin{aligned}
F_{ii_1,kl} &= \mathbb{E}\{k_{ii_1}^2 < \mathbf{R}_i \text{skew}(\mathbf{R}_{i_1}^\top \mathbf{Z}_{ii_1}^\top \mathbf{R}_{i_1}) \mathbf{R}_i^\top, \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top >_{\mathbf{R}} < \mathbf{R}_i \text{skew}(\mathbf{R}_{i_1}^\top \mathbf{Z}_{ii_1}^\top \mathbf{R}_{i_1}) \mathbf{R}_i^\top, \mathbf{R}_i \mathbf{E}_k \mathbf{R}_k^\top >_{\mathbf{R}}\} \\
&= -2k_{ii_1}^2 \mathbb{E}\{\Phi_k^{ii_1}\}^2 \\
&= -\frac{1}{3} \omega_{ii_1,kk} \\
&= -\frac{1}{3} \left( \frac{k_{ii_1}^2 (2I_0(2k_{ii_1}) - I_1(2k_{ii_1}) - 2I_2(2k_{ii_1}) + I_3(2k_{ii_1}))}{2I_0(2k_{ii_1}) - 2I_1(2k_{ii_1})} \right)
\end{aligned} \tag{122}$$

When  $k \neq l$

$$\mathbf{F}_{ii_1,kl} = 0 \tag{123}$$

So we have:

$$\mathbf{F}_{ii_1} = -\frac{1}{3} \left( \frac{k_{ii_1}^2 (2I_0(2k_{ii_1}) - I_1(2k_{ii_1}) - 2I_2(2k_{ii_1}) + I_3(2k_{ii_1}))}{2I_0(2k_{ii_1}) - 2I_1(2k_{ii_1})} \right) \mathbf{I}_{3 \times 3} \tag{124}$$

Combine (120) and (124), we have:

$$\mathbf{F}_{ii_1} = \begin{cases} \frac{1}{3} \left( \sum_{j \in V_i} \frac{k_{ij}^2 (2I_0(2k_{ij}) - I_1(2k_{ij}) - 2I_2(2k_{ij}) + I_3(2k_{ij}))}{2I_0(2k_{ij}) - 2I_1(2k_{ij})} \right) \mathbf{I}_{3 \times 3} & i = i_1 \\ -\frac{1}{3} \left( \frac{k_{ii_1}^2 (2I_0(2k_{ii_1}) - I_1(2k_{ii_1}) - 2I_2(2k_{ii_1}) + I_3(2k_{ii_1}))}{2I_0(2k_{ii_1}) - 2I_1(2k_{ii_1})} \right) \mathbf{I}_{3 \times 3} & (i, i_1) \in \mathcal{F} \\ 0 & \text{else} \end{cases} \tag{125}$$

It is easy to find that the  $SO(3)$  sub-matrix is a weighted Laplacian matrix after using the Kronecker product operation. This matrix can be written as  $\mathbf{L}_w^{SO(3)} = \mathbf{L}_{w_{SO(3)}} \otimes \mathbf{I}_{3 \times 3}$ . The  $(i, i_1)$ -th block of the weighted Laplacian matrix  $\mathbf{L}_{w_{SO(3)}}$  is:

$$(\mathbf{L}_{w_{SO(3)}})_{i,i_1} = \begin{cases} \sum_{j \in V_i} w_{ij}^{SO(3)} & i = i_1 \\ -w_{i_1 i}^{SO(3)} & (i, i_1) \in \mathcal{E} \\ -w_{i_1 i}^{SO(3)} & (i_1, i) \in \mathcal{E} \\ 0 & \text{else,} \end{cases} \quad (126)$$

where  $w_{ij}^{SO(3)} = \frac{\omega_{ij}}{3} = \frac{\omega_{ii,kk}}{3}$ .

## 8 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  $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{n_p} \in \mathbb{R}^n$  and rotations  $\mathbf{R}_0, \mathbf{R}_1, \dots, \mathbf{R}_{n_p} \in SO(n)$  from noisy measurements of some relative rotations  $\mathbf{R}_j \mathbf{R}_i^\top$  and relative coordinate transformations  $\mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i)$ .

In 2D/3D pose-graph SLAM, including the anchored pose, the parameter space is:  $\mathcal{P} = \{\mathbb{R}^n \times \dots \times \mathbb{R}^n\}_{n_p+1} \times \{SO(n) \times \dots \times SO(n)\}_{n_p+1}$ .

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

$$\mathbf{p}_{ij} = \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i) + \mathbf{y}_{ij}, \quad (127)$$

where  $\mathbf{y}_{ij} \sim \mathcal{N}(\mathbf{0}, \Sigma_{ij})$ , meaning that  $\mathbf{y}_{ij}$  is a random vector follows a zero mean Gaussian distribution with probability density function (PDF)  $f_{ij} : \mathbb{R}^n \rightarrow \mathbb{R}^+$ :

$$f_{ij}(\mathbf{y}_{ij}) = \frac{1}{(2\pi)^{n/2} \det(\Sigma_{ij})^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{y}_{ij}^\top \Sigma_{ij}^{-1} \mathbf{y}_{ij})\right), \quad (128)$$

where  $\Sigma_{ij} = \delta_{ij}^2 \mathbf{I}_{n \times n}$ ,  $\delta_{ij}^2$  means the isotropic variance value.

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

$$\mathbf{H}_{ij} = \mathbf{Z}_{ij} \mathbf{R}_j \mathbf{R}_i^\top, \quad (129)$$

where  $\mathbf{Z}_{ij}$  is a random rotation whose distributed function  $\widehat{f}_{ij} : SO(n) \rightarrow \mathbb{R}^+$  meets an isotropic Langevin distribution with mean  $\mathbf{I}_{n \times n}$  and concentration  $\kappa_{ij} \geq 0$  [13]:

$$\begin{aligned} \widehat{f}_{ij}(\mathbf{Z}_{ij}) &= \frac{1}{c_n(\kappa_{ij})} \exp(\kappa_{ij} \operatorname{trace}(\mathbf{Z}_{ij})), \\ c_2(\kappa_{ij}) &= I_0(2\kappa_{ij}), \\ c_3(\kappa_{ij}) &= \exp(\kappa_{ij})(I_0(2\kappa_{ij}) - I_1(2\kappa_{ij})), \\ I_v(2\kappa_{ij}) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(2\kappa_{ij} \cos(\theta)) \cos(v\theta) d\theta, \end{aligned} \quad (130)$$

where  $c_n(\kappa_{ij})$ ,  $n = 2, 3$  is a normalization constant such that  $\widehat{f}_{ij}$  has unit mass.  $I_v(2\kappa_{ij})$ ,  $v = \{0, 1, 2, \dots\} \in \mathbb{Z}$  means the modified Bessel functions [37]. We write  $\mathbf{Z}_{ij} \sim Lang(\mathbf{I}_{n \times n}, \kappa_{ij})$  to mean that  $\mathbf{Z}_{ij}$  is a random rotation matrix with PDF (130). This PDF meets the assumptions shown in [13]: 1. smoothness and support; 2. independence; 3. bi-invariance.

Under the independence assumption of the measurements, the log-likelihood function of an estimand (parameters to be estimated)  $\boldsymbol{\theta} = \mathbf{x} \times \mathbf{R} = (\mathbf{x}_0, \dots, \mathbf{x}_{n_p}, \mathbf{R}_0, \dots, \mathbf{R}_{n_p}) \in \mathcal{P}$ , given the measurements  $\mathbf{y} = \{\mathbf{p}_{ij}, \mathbf{H}_{ij}, (i, j) \in \mathcal{E}\}$ <sup>1</sup>, is given by:

$$L(\mathbf{y}; \boldsymbol{\theta}) = \sum_{(i,j) \in \mathcal{E}} \log f_{ij}(\mathbf{p}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i)) + \frac{1}{2} \sum_{(i,j) \in \mathcal{F}} \log \widehat{f}_{ij}(\mathbf{H}_{ij} \mathbf{R}_i \mathbf{R}_j^\top). \quad (131)$$

---

<sup>1</sup>If an edge  $(i, j) \in \mathcal{E}$ , then there exists a corresponding edge  $(i, j) \in \mathcal{F}$  for the rotation graph.

The coefficient  $\frac{1}{2}$  is used to balance the information in the un-directed rotation graph  $\mathcal{G}_1$  and the directed pose-graph  $\mathcal{G}$ .

By introducing the PDF functions (128), (130), finding the maximum of the log-likelihood function (131) is equivalent to:

$$\max_{\boldsymbol{\theta} \in \mathcal{P}} \sum_{(i,j) \in \mathcal{E}} \kappa_{ij} \text{trace}(\mathbf{H}_{ij} \mathbf{R}_i \mathbf{R}_j^\top) - \sum_{(i,j) \in \mathcal{E}} \frac{\delta_{ij}^{-2}}{2} \|\mathbf{p}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i)\|_2^2. \quad (132)$$

## 9 2D pose-graph SLAM

### 9.1 Synchronization on $\mathbb{R}^2 \times SO(2)$

Synchronization on  $\mathbb{R}^2 \times SO(2)$  is the problem of estimating a set of rotations  $\mathbf{R}_0, \mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_{n_p} \in SO(2)$  and positions  $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n_p} \in \mathbb{R}^2$  from noise measurements of some relative rotations  $\mathbf{R}_j \mathbf{R}_i^\top$  and the relative coordinate transformations  $\mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i)$ .

In our estimation problem, the parameter space is:  $\mathcal{P} = \{SO(2) \times \dots \times SO(2)\}_{n_p+1} \times \{\mathbb{R}^2 \times \dots \times \mathbb{R}^2\}_{n_p+1}$ . There are two kinds of edges in the whole graph: one only for  $SO(2)$  and another one for pose graph.

For rotation graph  $(i, j) \in \mathcal{F}$ , it is noted that  $\mathcal{F}$  is a un-directed edge set, we have one measurement  $\mathbf{H}_{ij} \in SO(2)$  of a form:

$$\mathbf{H}_{ij} = \mathbf{Z}_{ij} \mathbf{R}_j \mathbf{R}_i^\top \quad (133)$$

where  $\mathbf{Z}_{ij}$  is a random rotation whose distributed function  $\hat{f}_{ij} : SO(2) \rightarrow \mathbb{R}^+$  meets a isotropic Langevin distribution:

$$\begin{aligned} \hat{f}_{ij}(\mathbf{Z}_{ij}) &= \frac{1}{c_2(\kappa_{ij})} \exp(\kappa_{ij} \text{trace}(\mathbf{Z}_{ij})) \\ c_2(\kappa_{ij}) &= I_0(2\kappa_{ij}) \\ I_0(2\kappa_{ij}) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(2\kappa_{ij} \cos(\theta)) d\theta \end{aligned} \quad (134)$$

This PDF meets: 1. smoothness and support; 2.independence; 3. invariance assumptions.

For the SE(2) edge,  $(i, j) \in \mathcal{E}$ , it is noted that  $\mathcal{E}$  is a directed edge set, we have a measurement  $\mathbf{p}_{ij} \in \mathbb{R}^2$  of a form:

$$\mathbf{p}_{ij} = \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i) + \mathbf{y}_{ij} \quad (135)$$

where  $\mathbf{y}_{ij}$  is a random vector whose distributed function  $f_{ij} : \mathbb{R}^2 \rightarrow \mathbb{R}^+$  meets a isotropic Gaussian distribution:

$$\begin{aligned} f_{ij}(\mathbf{y}_{ij}) &= \frac{1}{(2\pi)^{\frac{1}{2}} \det(\boldsymbol{\Sigma}_{ij})^{1/2}} \exp(-\frac{1}{2}(\mathbf{y}_{ij}^\top \boldsymbol{\Sigma}_{ij}^{-1} \mathbf{y}_{ij})) \\ \boldsymbol{\Sigma}_{ij} &= \delta_{ij}^2 \mathbf{I}_{2 \times 2} \end{aligned} \quad (136)$$

Hence<sup>2</sup>,

$$L(\mathbf{y}; \boldsymbol{\theta}) = \frac{1}{2} \sum_{(i,j) \in \mathcal{F}} \log \hat{f}_{ij}(\mathbf{H}_{ij} \mathbf{R}_i \mathbf{R}_j^\top) + \sum_{(i,j) \in \mathcal{E}} \log f_{ij}(\mathbf{p}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i)) \quad (137)$$

### 9.2 Geometry of the parameter space

In 2D case, define:

$$\mathbf{E} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad (138)$$

---

<sup>2</sup>In the log-likelihood function (214), based on measurement function (127) and (129),  $\mathbf{y}_{ij}$  and  $\mathbf{Z}_{ij}$  can be seen as the function of the estimand  $\mathbf{x} \times \mathbf{R}$ , satisfying:  $\mathbf{y}_{ij} = \mathbf{y}_{ij}(\mathbf{R}_i, \mathbf{x}_i, \mathbf{x}_j) = \mathbf{p}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i)$  and  $\mathbf{Z}_{ij} = \mathbf{Z}_{ij}(\mathbf{R}_i, \mathbf{R}_j) = \mathbf{H}_{ij} \mathbf{R}_i \mathbf{R}_j^\top$ . So the function (131) can be written as  $L(\mathbf{y}; \boldsymbol{\theta}) = \sum_{(i,j) \in \mathcal{E}} \log f_{ij}(\mathbf{y}_{ij}) + \frac{1}{2} \sum_{(i,j) \in \mathcal{F}} \log \hat{f}_{ij}(\mathbf{Z}_{ij})$ .

the orthonormal basis  $\mathbf{E}^{\mathbf{x}, \mathbf{R}} = (\mathbf{E}_{0,1}^{\mathbf{x}}, \mathbf{E}_{0,2}^{\mathbf{x}}, \dots, \mathbf{E}_{n_p,1}^{\mathbf{x}}, \mathbf{E}_{n_p,2}^{\mathbf{x}}, \mathbf{E}_0^{\mathbf{R}}, \dots, \mathbf{E}_{n_p}^{\mathbf{R}})$  of the tangent space  $\mathcal{T}_{(\mathbf{x}, \mathbf{R})}\mathcal{P}$  can be fixed as:

$$\begin{aligned}\mathbf{E}_{i,k}^{\mathbf{x}} &= (\mathbf{E}_{i,k}^{X^\top}; \mathbf{0}_{2(n_p+1) \times 2}), \quad i \in \{0, 1, \dots, n_p\}, \quad k = 1, 2 \\ \mathbf{E}_j^{\mathbf{R}} &= (\mathbf{0}_{2(n_p+1) \times 1}; \mathbf{E}_j^{R^\top}), \quad j \in \{0, 1, \dots, n_p\}, \\ \mathbf{E}_{i,k}^X &= (\mathbf{0}_{1 \times 2}, \dots, \mathbf{0}_{1 \times 2}, \underbrace{1}_{\substack{k-\text{th} \\ i-\text{th}}}, 0, \mathbf{0}_{1 \times 2}, \dots, \mathbf{0}_{1 \times 2})_{1 \times 2(n_p+1)}, \\ \mathbf{E}_j^R &= (\mathbf{0}_{2 \times 2}, \dots, \mathbf{0}_{2 \times 2}, \underbrace{\mathbf{R}_j \mathbf{E}}_{j-\text{th}}, \mathbf{0}_{2 \times 2}, \dots, \mathbf{0}_{2 \times 2})_{2 \times 2(n_p+1)},\end{aligned}\tag{139}$$

where  $\mathbf{E}_{i,k}^{\mathbf{x}}$  and  $\mathbf{E}_j^{\mathbf{R}}$  are corresponding to the  $k$ -axis coordinate of the  $i$ -th pose  $(\mathbf{x}_i, \mathbf{R}_i)$  and the rotation parameter of the  $j$ -th pose  $(\mathbf{x}_j, \mathbf{R}_j)$ .

The inner product is defined by:

$$\langle \Omega_1, \Omega_2 \rangle_\theta = \begin{cases} \Omega_1^\top \cdot \Omega_2 & \theta = \mathbf{X} \\ \text{trace}(\Omega_1^\top \cdot \Omega_2) & \theta = \mathbf{R} \end{cases}.\tag{140}$$

The FIM of the estimation problem w.r.t. the basis  $e = \mathbf{E}^{\mathbf{x}, \mathbf{R}}$  is defined as:

$$F_{ij} = \mathbb{E}\{\langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), e_i \rangle_\theta \cdot \langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), e_j \rangle_\theta^\top\}\tag{141}$$

The information matrix can be divided into four parts: 1.  $\mathbb{R}^2$  sub-matrix, 2.  $SO(2)$  by  $\mathbb{R}^2$  coupling sub-matrix, 3.  $\mathbb{R}^2$  by  $SO(2)$  coupling sub-matrix, 4.  $SO(2)$  sub-matrix.

### 9.3 FIM for 2D pose-graph SLAM

#### 9.3.1 $\mathbb{R}^2$ sub-matrix

We deal with the inner product corresponding to the bases of the coordinate of the  $i$ -th pose as a  $\mathbb{R}^{n \times 1}$ ,  $n = 2, 3$  vector meeting:

$$\langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_i^{\mathbf{x}} \rangle_{\mathbf{X}}^* = \begin{cases} \begin{bmatrix} \langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_{i,1}^{\mathbf{x}} \rangle_{\mathbf{x}} \\ \langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_{i,2}^{\mathbf{x}} \rangle_{\mathbf{x}} \end{bmatrix} & n = 2 \\ \begin{bmatrix} \langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_{i,1}^{\mathbf{x}} \rangle_{\mathbf{x}} \\ \langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_{i,2}^{\mathbf{x}} \rangle_{\mathbf{x}} \\ \langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_{i,3}^{\mathbf{x}} \rangle_{\mathbf{x}} \end{bmatrix} & n = 3 \end{cases}\tag{142}$$

Then, based on the definition (21), we have:

$$\begin{aligned}F_{i,i_1} &= \mathbb{E}\{\langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_i^{\mathbf{x}} \rangle_{\mathbf{X}}^* \cdot \langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_{i_1}^{\mathbf{x}} \rangle_{\mathbf{X}}^*\} \\ &= \mathbb{E}\{\text{grad}_{\mathbf{x}_i} L(\mathbf{y}; \boldsymbol{\theta})^\top \cdot \text{grad}_{\mathbf{x}_{i_1}} L(\mathbf{y}; \boldsymbol{\theta})\},\end{aligned}\tag{143}$$

where  $\mathbf{F}_{i,i_1} \in \mathbb{R}^{2 \times 2}$ ,  $\text{grad}_{\mathbf{x}_i} L(\mathbf{y}; \boldsymbol{\theta})$  means the gradient of  $L(\mathbf{y}; \boldsymbol{\theta})$  with respect to parameter  $\mathbf{x}_i$ .

Based on (214), we have:

$$\begin{aligned}\text{grad}_{\mathbf{x}_i} L(\mathbf{y}; \boldsymbol{\theta}) &= \text{grad}_{\mathbf{x}_i} \frac{1}{2} \sum_{(i,j) \in \mathcal{F}} \log \hat{f}_{ij}(\mathbf{H}_{ij} \mathbf{R}_i \mathbf{R}_j^\top) + \text{grad}_{\mathbf{x}_i} \sum_{(i,j) \in \mathcal{E}} \log f_{ij}(\mathbf{p}_{ij} - \mathbf{R}_j^\top(\mathbf{x}_i - \mathbf{x}_j)) \\ &= 0 + \text{grad}_{\mathbf{x}_i} \sum_{(i,j) \in \mathcal{E}} \log f_{ij}(\mathbf{p}_{ij} - \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i)) \\ &= \sum_{(i,j) \in \mathcal{E}} \nabla_{\mathbf{y}_{ij}} \log \frac{1}{2\pi \det(\boldsymbol{\Sigma}_{ij})^{1/2}} \exp(-\frac{1}{2}(\mathbf{y}_{ij}^\top \delta_{ij}^{-2} \mathbf{I}_{2 \times 2} \mathbf{y}_{ij})) \nabla_{\mathbf{x}_i} \mathbf{y}_{ij}\end{aligned}\tag{144}$$

Introduce (144) into (217), we have:

$$\begin{aligned}
\mathbf{F}_{ii_1} &= \mathbb{E}\{\text{grad}_{\mathbf{x}_i} L(\theta) \cdot \text{grad}_{\mathbf{x}_{i_1}} L(\mathbf{y}; \boldsymbol{\theta})^\top\} \\
&= \mathbb{E}\left\{\sum_{(i,j) \in \mathcal{E}} \sum_{(i,j) \in \mathcal{E}} -\delta_{ij}^{-2} \mathbf{y}_{ij}^\top \nabla_{\mathbf{x}_i} \mathbf{y}_{ij} \cdot (-\delta_{i_1 j_1}^{-2} \mathbf{y}_{i_1 j_1}^\top \nabla_{\mathbf{x}_i} \mathbf{y}_{i_1 j_1})^\top\right\} \\
&= \mathbb{E}\left\{\sum_{(i,j) \in \mathcal{E}} \sum_{(i_1, j_1) \in \mathcal{E}} \delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \mathbf{y}_{ij}^\top \nabla_{\mathbf{x}_i} \mathbf{y}_{ij} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1}^\top \cdot \mathbf{y}_{i_1 j_1}\right\}
\end{aligned} \tag{145}$$

When  $i = i_1$ , based on two sets  $V_i^+$  and  $V_i^-$  (meet  $(i, j) \in \mathcal{E} \leftrightarrow j \in V_i^+$  and  $(j, i) \in \mathcal{E} \leftrightarrow j \in V_i^-$ ), we have:

$$\begin{aligned}
\mathbf{F}_{ii_1} &= \mathbb{E}\left\{\sum_{(i,j) \in \mathcal{E}} \sum_{(i_1, j_1) \in \mathcal{E}} \delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \mathbf{y}_{ij}^\top \nabla_{\mathbf{x}_i} \mathbf{y}_{ij} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1}^\top \mathbf{y}_{i_1 j_1}\right\} \\
&= \mathbb{E}\left\{\sum_{(i,j) \in \mathcal{E}} \sum_{(i_1, j_1) \in \mathcal{E}} P_{ij}(\mathbf{x}_i)^\top P_{i_1 j_1}(\mathbf{x}_{i_1 j_1})\right\} \\
&= \mathbb{E}\left\{\sum_{j \in V_i^+} \sum_{j_1 \in V_{i_1}^+} P_{ij}(\mathbf{x}_i)^\top P_{i_1 j_1}(\mathbf{x}_{i_1}) + \sum_{j \in V_i^+} \sum_{j_1 \in V_{i_1}^-} P_{ij}(\mathbf{x}_i)^\top P_{j_1 i_1}(\mathbf{x}_{i_1}) + \right. \\
&\quad \left. \sum_{j \in V_i^-} \sum_{j_1 \in V_{i_1}^+} P_{ji}(\mathbf{x}_i)^\top P_{i_1 j_1}(\mathbf{x}_{i_1}) + \sum_{j \in V_i^-} \sum_{j_1 \in V_{i_1}^-} P_{ji}(\mathbf{x}_i)^\top P_{j_1 i_1}(\mathbf{x}_{i_1})\right\}
\end{aligned} \tag{146}$$

Because  $\nabla_{\mathbf{x}_i} \mathbf{y}_{ij} = -\mathbf{R}_i^\top$  and  $\nabla_{\mathbf{x}_i} \mathbf{y}_{ji} = \mathbf{R}_i^\top$ , we have:

$$\begin{aligned}
&\mathbb{E}\left\{\sum_{j \in V_i^+} \sum_{j_1 \in V_{i_1}^+} P_{ij}(\mathbf{x}_i)^\top P_{i_1 j_1}(\mathbf{x}_{i_1}) + \sum_{j \in V_i^-} \sum_{j_1 \in V_{i_1}^-} P_{ji}(\mathbf{x}_i)^\top P_{j_1 i_1}(\mathbf{x}_{i_1})\right\} \\
&= \mathbb{E}\left\{\sum_{j \in V_i^+} \sum_{j_1 \in V_{i_1}^+} \delta_{ij}^{-4} \mathbf{y}_{ij}^\top \mathbf{R}_i (-1)(-1) \mathbf{R}_i^\top \mathbf{y}_{i_1 j_1} + \sum_{j \in V_i^-} \sum_{j_1 \in V_{i_1}^-} \delta_{ij}^{-4} \mathbf{y}_{ij}^\top \mathbf{R}_i \mathbf{R}_i^\top \mathbf{y}_{i_1 j_1}\right\} \\
&= \sum_{j \in V_i^+} \delta_{ij}^{-4} \mathbb{E}\{\mathbf{y}_{ij}^\top \mathbf{y}_{i_1 j_1}\} + \sum_{j \in V_i^-} \delta_{ij}^{-4} \mathbb{E}\{\mathbf{y}_{ij}^\top \mathbf{y}_{i_1 j_1}\} \\
&= \sum_{j \in V_i^+} \delta_{ij}^{-2} \mathbf{I}_{2 \times 2} + \sum_{j \in V_i^-} \delta_{ij}^{-2} \mathbf{I}_{2 \times 2}
\end{aligned} \tag{147}$$

When  $i \neq i_1$ , based on (146), we have:

$$\begin{aligned}
\mathbf{F}_{ii_1} &= \mathbb{E}\left\{\sum_{j \in V_i^+} \sum_{j_1 \in V_{i_1}^+} P_{ij}(\mathbf{x}_i)^\top P_{i_1 j_1}(\mathbf{x}_{i_1}) + \sum_{j \in V_i^+} \sum_{j_1 \in V_{i_1}^-} P_{ij}(\mathbf{x}_i)^\top P_{j_1 i_1}(\mathbf{x}_{i_1}) + \right. \\
&\quad \left. \sum_{j \in V_i^-} \sum_{j_1 \in V_{i_1}^+} P_{ji}(\mathbf{x}_i)^\top P_{i_1 j_1}(\mathbf{x}_{i_1}) + \sum_{j \in V_i^-} \sum_{j_1 \in V_{i_1}^-} P_{ji}(\mathbf{x}_i)^\top P_{j_1 i_1}(\mathbf{x}_{i_1})\right\} \\
&= \mathbb{E}\left\{\sum_{j \in V_i^+} \sum_{j_1 \in V_{i_1}^-} P_{ij}(\mathbf{x}_i)^\top P_{j_1 i_1}(\mathbf{x}_{i_1}) + \sum_{j \in V_i^-} \sum_{j_1 \in V_{i_1}^+} P_{ji}(\mathbf{x}_i)^\top P_{i_1 j_1}(\mathbf{x}_{i_1})\right\} \\
&= \begin{cases} \mathbb{E}\{\sum_{j \in V_i^+} \sum_{j_1 \in V_{i_1}^-} P_{ij}(\mathbf{x}_i)^\top P_{j_1 i_1}(\mathbf{x}_{i_1})\} + 0 & (i, i_1) \in \mathcal{E} \\ 0 + \mathbb{E}\{\sum_{j \in V_i^-} \sum_{j_1 \in V_{i_1}^+} P_{ji}(\mathbf{x}_i)^\top P_{i_1 j_1}(\mathbf{x}_{i_1})\} & (i_1, i) \in \mathcal{E} \end{cases} \\
&= -\delta_{ii_1}^{-4} \mathbb{E}\{\mathbf{y}_{ii_1}^\top \mathbf{y}_{i_1 i}\} \quad \text{if } (i, i_1) \in \mathcal{E} \text{ or } (i_1, i) \in \mathcal{E} \\
&= -\delta_{ii_1}^{-2} \mathbf{I}_{2 \times 2} \quad \text{if } (i, i_1) \in \mathcal{E} \text{ or } (i_1, i) \in \mathcal{E}
\end{aligned} \tag{148}$$

When  $(i, i_1) \notin \mathcal{E}$  and  $i \neq i_1$ , based on  $\mathbb{E}\{\mathbf{y}_{ij} \mathbf{y}_{i_1 j_1}\} = \mathbb{E}\{\mathbf{y}_{ij}\} \mathbb{E}\{\mathbf{y}_{i_1 j_1}\} = \mathbf{0}$ , we have  $\mathbf{F}_{i,i_1} = \mathbf{0}_{2 \times 2}$ .

Combine (147) and (148), let  $(\mathbf{L}_w^{\mathbb{R}^2})_{i,i_1} = \mathbf{F}_{i,i_1}$ , we have the  $(i, i_1)$ -th block of the sub-matrix  $\mathbf{L}_w^{\mathbb{R}^2}$  corresponding to the  $(i, i_1)$ -th block of the FIM and belonging the Euclidean space  $\mathbb{R}^2$ :

$$\mathbf{F}_{ii_1} = (\mathbf{L}_w^{\mathbb{R}^2})_{i,i_1} = \begin{cases} \sum_{j_1 \in V_i^+} \delta_{ij_1}^{-2} \mathbf{I}_{2 \times 2} + \sum_{j_2 \in V_i^-} \delta_{j_2 i}^{-2} \mathbf{I}_{2 \times 2}, & i = i_1 \\ -\delta_{ii_1}^{-2} \mathbf{I}_{2 \times 2}, & (i, i_1) \in \mathcal{E} \text{ or } (i_1, i) \in \mathcal{E} \\ 0, & \text{else} \end{cases} \tag{149}$$

where  $\mathbf{L}_w^{\mathbb{R}^2} = \mathbf{L}_{w_{\mathbb{R}}} \otimes \mathbf{I}_{2 \times 2}$ , satisfying:

$$(\mathbf{L}_{w_{\mathbb{R}}})_{i,i_1} = \begin{cases} \sum_{j \in V_i} w_{ij}^{\mathbb{R}} & i = i_1 \\ -w_{i_1 i}^{\mathbb{R}} & (i, i_1) \in \mathcal{E} \\ -w_{i_1 i}^{\mathbb{R}} & (i_1, i) \in \mathcal{E} \\ 0 & \text{else,} \end{cases} \quad (150)$$

where  $w_{ij}^{\mathbb{R}} = \delta_{ij}^{-2}$  and  $\mathbf{L}_{w_{\mathbb{R}}}$  is a weighted Laplacian matrix.

### 9.3.2 $SO(2)$ by $\mathbb{R}^2$ coupling sub-matrix

Based on the geometry of the parameter space and the definition in (21), we can get the block of the  $SO(2)$  by  $\mathbb{R}^2$  coupling sub-matrix:

$$\begin{aligned} \mathbf{F}_{n_p+1+i,i_1} &= \mathbb{E}\{\langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_i^R \rangle_{\mathbf{R}} \cdot \langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_{i_1}^x \rangle_{\mathbf{x}}^*\} \\ &= \mathbb{E}\{\text{trace}(\text{grad}_{\mathbf{R}_i} L(\mathbf{y}; \boldsymbol{\theta})^\top \cdot \mathbf{R}_i \mathbf{E}) \cdot \text{grad}_{\mathbf{x}_{i_1}} L(\mathbf{y}; \boldsymbol{\theta})\}, \end{aligned} \quad (151)$$

where  $\mathbf{F}_{n_p+1+i,i_1} \in \mathbb{R}^{1 \times 2}$ ,  $\text{grad}_{\mathbf{R}_i} L(\mathbf{y}; \boldsymbol{\theta})$  means the gradient of  $L(\mathbf{y}; \boldsymbol{\theta})$  with respect to parameter  $\mathbf{R}_i$ .

When  $i > N$ , we have:

$$\begin{aligned} \text{grad}_{\mathbf{R}_i} L(\mathbf{y}; \boldsymbol{\theta}) &= \text{grad}_{\mathbf{R}_i} \frac{1}{2} \sum_{(i,j) \in \mathcal{E}} \log \hat{f}_{ij}(\mathbf{H}_{ij} \mathbf{R}_i \mathbf{R}_j^\top) + \text{grad}_{\mathbf{R}_i} \sum_{(i,j) \in \mathcal{E}} \log f_{ij}(\mathbf{p}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i)) \\ &= \sum_{(i,j) \in \mathcal{E}} -k_{ij} \mathbf{R}_i \text{skew}(\mathbf{Z}_{ij}) + \text{grad}_{\mathbf{R}_i} \sum_{(i,j) \in \mathcal{E}} \log f_{ij}(\mathbf{p}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i)) \end{aligned} \quad (152)$$

Let's consider the second part. There are two ways to get this part:

(1) Based on the basis  $\mathbf{R}_i \mathbf{E}$ :

$$\begin{aligned} &\sum_{(i,j) \in \mathcal{E}} \text{grad}_{\mathbf{R}_i} \log f_{ij}(\mathbf{p}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i)) \\ &= \sum_{(i,j) \in \mathcal{E}} \frac{1}{2} \mathbf{R}_i \mathbf{E} \nabla_{\mathbf{R}_i} \log f_{ij}(\mathbf{p}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i)) \end{aligned} \quad (153)$$

And we also have:

$$\begin{aligned} &\nabla_{\mathbf{R}_i} \log f_{ji}(\mathbf{P}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i)) \\ &= \nabla_{\mathbf{R}_i} \log \frac{1}{2\pi|\Sigma|^{1/2}} \exp(-\frac{1}{2}((\mathbf{p}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i))^\top \Sigma^{-1}(\mathbf{p}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i)))) \\ &= \nabla_{\mathbf{R}_i} \log \frac{1}{2\pi|\Sigma|^{1/2}} + \log \exp(-\frac{1}{2}((\mathbf{p}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i))^\top \Sigma^{-1}(\mathbf{p}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i)))) \\ &= \nabla_{\mathbf{R}_i} \log \exp(-\frac{1}{2}((\mathbf{P}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i))^\top \Sigma^{-1}(\mathbf{p}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i)))) \\ &= \nabla_{\mathbf{R}_i} - \frac{1}{2}((\mathbf{P}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i))^\top \Sigma^{-1}(\mathbf{p}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i))) \\ &= \nabla_{\mathbf{R}_i} - \frac{1}{2}(\mathbf{y}_{ij}^\top \Sigma^{-1} \mathbf{y}_{ij}) \\ &= \nabla_{\mathbf{y}_{ij}} - \frac{1}{2}(\mathbf{y}_{ij}^\top \Sigma^{-1} \mathbf{y}_{ij}) \nabla_{\mathbf{R}_i} \mathbf{y}_{ij} \\ &= -\frac{1}{2} \mathbf{y}_{ij}^\top (\Sigma^{-1} + \Sigma^{-\top}) \nabla_{\mathbf{R}_i} \mathbf{y}_{ij} \\ &= -\delta_{ij}^{-2} \mathbf{y}_{ij}^\top \nabla_{\mathbf{R}_i} \mathbf{y}_{ij} \end{aligned} \quad (154)$$

For  $\nabla_{\mathbf{R}_i} \mathbf{y}_{ij}(\mathbf{R}_i)$ , we have:

$$\begin{aligned}
& \nabla_{\mathbf{R}_i} \mathbf{y}_{ji}(\mathbf{R}_i) \\
&= \frac{\partial \mathbf{p}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i)}{\partial \mathbf{R}_i} \\
&= 0 - \frac{\partial \exp(\delta)^\top \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i)}{\partial \delta} \\
&= -(\mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i))_\times \\
&= -\mathbf{E}^\top \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i)
\end{aligned} \tag{155}$$

Introduce (238) and (237) into (153), we can get:

$$\begin{aligned}
(153) &= \sum_{(i,j) \in E} \mathbf{R}_i \mathbf{E} \delta_{ij}^{-2} \mathbf{y}_{ji}^\top \nabla_{\mathbf{R}_i} \mathbf{y}_{ji} \\
&= \frac{1}{2} \sum_{(i,j) \in E} \delta_{ij}^{-2} \mathbf{R}_i \mathbf{E} \mathbf{y}_{ji}^\top \mathbf{E}^\top \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i)
\end{aligned} \tag{156}$$

(2) Using the general computation way for gradient on  $SO(2)$ : Let  $\tilde{f}: R_{2 \times 2} \rightarrow R$  be a differentiable function and let  $f = \tilde{f}|SO(2)$  be its restriction to  $SO(2)$ . Let  $\nabla f(\mathbf{Q})$  be the usual gradient of  $f$  in  $R_{2 \times 2}$ . Then, the gradient of  $f$  is easily computed as:

$$\text{grad } f(\mathbf{Q}) = \mathbf{Q} \text{skew}(\mathbf{Q}^\top \cdot \nabla \tilde{f}(\mathbf{Q})), \quad \text{skew}(\star) = (\star - \star^\top)/2 \tag{157}$$

Firstly let's get the usual gradient of  $\tilde{f}$  in  $R_{n \times n}$ :

$$\begin{aligned}
\nabla \tilde{f}(\mathbf{Q}) &= \nabla_{\mathbf{R}_i \in R_{n \times n}} \log \frac{1}{2\pi|\Sigma|^{1/2}} \exp(-\frac{1}{2}((\mathbf{p}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i))^\top \Sigma^{-1}(\mathbf{p}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i)))) \\
&= \nabla_{\mathbf{R}_i \in R_{n \times n}} \frac{1}{2}(-(\mathbf{p}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i))^\top \Sigma^{-1}(\mathbf{p}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i))) \\
&= \nabla_{\mathbf{R}_i \in R_{n \times n}} -\delta_{ij}^{-2} \frac{1}{2}((\mathbf{p}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i))^\top (\mathbf{p}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i))) \\
&= \nabla_{\mathbf{R}_i \in R_{n \times n}} -\delta_{ij}^{-2} \frac{1}{2} \mathbf{y}_{ij}^\top \mathbf{y}_{ij} \\
&= \delta_{ij}^{-2} (\mathbf{y}_{ij} (\mathbf{x}_j - \mathbf{x}_i)^\top)^\top \\
&= \delta_{ij}^{-2} (\mathbf{x}_j - \mathbf{x}_i) \mathbf{y}_{ij}^\top
\end{aligned} \tag{158}$$

Let  $\mathbf{Q} = \mathbf{R}_i$ , we have:

$$\begin{aligned}
\text{grad}_{\mathbf{R}_i} f(\mathbf{R}_i) &= \mathbf{R}_i \text{skew}(\mathbf{R}_i^\top \cdot \nabla \tilde{f}(\mathbf{R}_i)) \\
&= \delta_{ij}^{-2} \mathbf{R}_i (\mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i) \mathbf{y}_{ij}^\top - \mathbf{y}_{ij} (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i) \\
&= \delta_{ij}^{-2} ((\mathbf{x}_j - \mathbf{x}_i) \mathbf{y}_{ij}^\top - \mathbf{R}_i \mathbf{y}_{ij} (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i)/2
\end{aligned} \tag{159}$$

Using simulation, we can find that Eq.(159) is equal to Eq.(156).

For  $\text{grad}_{\mathbf{x}_{i_1}} L(\mathbf{y}; \boldsymbol{\theta})$ , based on (144) we have:

$$\begin{aligned}
& \text{grad}_{\mathbf{x}_{i_1}} L(\mathbf{y}; \boldsymbol{\theta}) \\
&= \text{grad}_{\mathbf{x}_{i_1}} \frac{1}{2} \sum_{(i_1, j_1) \in \mathcal{F}} \log \hat{f}_{i_1 j_1}(\mathbf{H}_{i_1 j_1} \mathbf{R}_{i_1} \mathbf{R}_{j_1}^\top) + \nabla_{\mathbf{x}_{i_1}} \sum_{(i_1, j_1) \in \mathcal{E}} \log f_{i_1, j_1}(\mathbf{y}_{i_1 j_1}) \\
&= 0 + \nabla_{\mathbf{x}_{i_1}} \sum_{(i_1, j_1) \in E} \log f_{i_1 j_1}(\mathbf{y}_{i_1 j_1}) \\
&= \sum_{j_1 \in V_{i_1}^+} -\delta_{i_1 j_1}^{-2} \mathbf{y}_{i_1 j_1}^\top \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1} + \sum_{j_2 \in V_{i_1}^-} -\delta_{j_2 i_1}^{-2} \mathbf{y}_{j_2 i_1}^\top \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{j_2 i_1}
\end{aligned} \tag{160}$$

For 2D case, because of  $\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j = \mathbf{Z}_{ij}^\top$ , so we have:

$$\text{grad}_{\mathbf{R}_i} \sum_{j \in V_i} \log \hat{f}_{ij}(\mathbf{Z}_{ij}) = \sum_{j \in V_i} -\kappa_{ij} \mathbf{R}_i \text{skew}(\mathbf{Z}_{ij}). \tag{161}$$

Introduce (156) and (160) into (151), we have:

$$\begin{aligned}
& \mathbf{F}_{n_p+1+i,i_1} \\
&= \mathbb{E}\{\langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_i^R \rangle_{\mathbf{R}} \cdot \langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_{i_1}^x \rangle_{\mathbf{x}}^{\top}\} \\
&= \mathbb{E}\{\text{trace}(\text{grad}_{\mathbf{R}_i} L(\mathbf{y}; \boldsymbol{\theta})^{\top} \cdot \mathbf{R}_i \mathbf{E}) \cdot \mathbf{I}_{2 \times 2}^{\top} \cdot \text{grad}_{\mathbf{x}_{i_1}} L(\mathbf{y}; \boldsymbol{\theta})\} \\
&= \mathbb{E}\left\{\sum_{j \in V_i^+} \delta_{ij}^{-2} \text{trace}\left(\left(-\kappa_{ij} \mathbf{R}_i \text{skew}(\mathbf{Z}_{ij}) + \frac{1}{2} \mathbf{R}_i \mathbf{E} \mathbf{y}_{ij}^{\top} \mathbf{E}^{\top} \mathbf{R}_i^{\top} (\mathbf{x}_j - \mathbf{x}_i)\right)^{\top} \mathbf{R}_i \mathbf{E}\right) \sum_{(i_1, j_1) \in \mathcal{E}} \delta_{i_1 j_1}^{-2} \mathbf{y}_{i_1 j_1}^{\top} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1}\right\} \\
&= \mathbb{E}\left\{\sum_{j \in V_i^+} \sum_{(i_1, j_1) \in \mathcal{E}} \delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \text{trace}\left(k_{ij} \text{skew}(\mathbf{Z}_{ij}) \mathbf{R}_i^{\top} \mathbf{R}_i \mathbf{E} + \frac{1}{2} (\mathbf{x}_j - \mathbf{x}_i)^{\top} \mathbf{R}_i \mathbf{E}^{\top} \mathbf{y}_{ij} \mathbf{I}_{2 \times 2}\right) \mathbf{y}_{i_1 j_1}^{\top} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1}\right\} \\
&= \mathbb{E}\left\{\sum_{j \in V_i^+} \sum_{(i_1, j_1) \in \mathcal{E}} \delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \text{trace}\left(\kappa_{ij} \text{skew}(\mathbf{Z}_{ij}) \mathbf{E} + \frac{1}{2} (\mathbf{x}_j - \mathbf{x}_i)^{\top} \mathbf{R}_i \mathbf{E}^{\top} \mathbf{y}_{ij} \mathbf{I}_{2 \times 2}\right) \mathbf{y}_{i_1 j_1}^{\top} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1}\right\} \\
&= \sum_{j \in V_i^+} \sum_{(i_1, j_1) \in \mathcal{E}} \mathbb{E}\{\delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \text{trace}(\kappa_{ij} \text{skew}(\mathbf{Z}_{ij}) \mathbf{E} \mathbf{y}_{i_1 j_1} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1})\} \\
&\quad + \mathbb{E}\{\delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \text{trace}((\mathbf{x}_j - \mathbf{x}_i)^{\top} \mathbf{R}_i \mathbf{E}^{\top} \mathbf{y}_{ij}) \mathbf{y}_{i_1 j_1}^{\top} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1}\} \\
&= \sum_{j \in V_i^+} \sum_{(i_1, j_1) \in \mathcal{E}} 0 + \mathbb{E}\{\delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \text{trace}((\mathbf{x}_j - \mathbf{x}_i)^{\top} \mathbf{R}_i \mathbf{E}^{\top} \mathbf{y}_{ij}) \mathbf{y}_{i_1 j_1}^{\top} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1}\} \\
&= \sum_{j \in V_i^+} \sum_{(i_1, j_1) \in \mathcal{E}} \mathbb{E}\{\delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \text{trace}((\mathbf{x}_j - \mathbf{x}_i)^{\top} \mathbf{R}_i \mathbf{E}^{\top} \mathbf{y}_{ij}) \mathbf{y}_{i_1 j_1}^{\top} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1}\} \\
&= \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}((\mathbf{x}_j - \mathbf{x}_i)^{\top} \mathbf{R}_i \mathbf{E}^{\top} \mathbf{y}_{ij}) \mathbf{y}_{i_1 j_1}^{\top} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1}\} \\
&\quad + \sum_{j \in V_i^+} \sum_{j_2 \in V_{i_1}^-} \mathbb{E}\{\delta_{ij}^{-2} \delta_{j_2 i_1}^{-2} \text{trace}((\mathbf{x}_j - \mathbf{x}_i)^{\top} \mathbf{R}_i \mathbf{E}^{\top} \mathbf{y}_{ij}) \mathbf{y}_{j_2 i_1}^{\top} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{j_2 i_1}\}
\end{aligned} \tag{162}$$

When  $(i, j) = (i_1, j_1) \in \mathcal{E}$ , we know that  $(\mathbf{x}_j - \mathbf{x}_i)^{\top} \mathbf{R}_i \mathbf{E}^{\top} \mathbf{y}_{ij} \in \mathbb{R}^1$  and  $\nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1} = -\mathbf{R}_{i_1}^{\top}$ , we have:

$$\begin{aligned}
\mathbf{F}_{n_p+1+i,i_1} &= \sum_{j \in V_i^+} \sum_{j_1 \in V_{i_1}^+} \mathbb{E}\{\delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \text{trace}((\mathbf{x}_j - \mathbf{x}_i)^{\top} \mathbf{R}_i \mathbf{E}^{\top} \mathbf{y}_{i_1 j_1}) \mathbf{y}_{i_1 j_1}^{\top} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1}\} + 0 \\
&= \sum_{j \in V_i^+} \sum_{j_1 \in V_{i_1}^+} \mathbb{E}\{\delta_{ij}^{-4} (\mathbf{x}_j - \mathbf{x}_i)^{\top} \mathbf{E} \mathbf{R}_i \mathbf{y}_{ij} \mathbf{y}_{ij}^{\top} \nabla_{\mathbf{x}_i} \mathbf{y}_{ij}\} \\
&= \sum_{j \in V_i^+} \delta_{ij}^{-4} (\mathbf{x}_j - \mathbf{x}_i)^{\top} \mathbf{E} \mathbf{R}_i \mathbb{E}\{\mathbf{y}_{ij} \mathbf{y}_{ij}^{\top}\} \nabla_{\mathbf{x}_i} \mathbf{y}_{ij} \\
&= \sum_{j \in V_i^+} \delta_{ij}^{-2} (\mathbf{x}_i - \mathbf{x}_j)^{\top} \mathbf{E}
\end{aligned} \tag{163}$$

When  $(i, j) = (j_2, i_1) \in \mathcal{E}$ , we know that  $(\mathbf{x}_j - \mathbf{x}_i)^{\top} \mathbf{R}_i \mathbf{E}^{\top} \mathbf{y}_{ij} \in \mathbb{R}^1$  and  $\nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{j_2 i_1} = \mathbf{R}_{i_1}^{\top}$ , we have:

$$\begin{aligned}
\mathbf{F}_{n_p+1+i,i_1} &= \sum_{j \in V_i^+} \sum_{j_2 \in V_{i_1}^-} \mathbb{E}\{\delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \text{trace}((\mathbf{x}_j - \mathbf{x}_i)^{\top} \mathbf{R}_i \mathbf{E}^{\top} \mathbf{y}_{i_1 j_1}) \mathbf{y}_{j_2 i_1}^{\top} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{j_2 i_1}\} + 0 \\
&= \sum_{j \in V_i^+} \sum_{j_2 \in V_{i_1}^-} \mathbb{E}\{\delta_{ij}^{-4} (\mathbf{x}_j - \mathbf{x}_i)^{\top} \mathbf{E} \mathbf{R}_i \mathbf{y}_{ij} \mathbf{y}_{ij}^{\top} \nabla_{\mathbf{x}_i} \mathbf{y}_{j_2 i_1}\} \\
&= \delta_{ij}^{-4} (\mathbf{x}_j - \mathbf{x}_i)^{\top} \mathbf{E} \mathbf{R}_i \mathbb{E}\{\mathbf{y}_{ij} \mathbf{y}_{ij}^{\top}\} \nabla_{\mathbf{x}_i} \mathbf{y}_{j_2 i_1} \\
&= \delta_{ij}^{-2} (\mathbf{x}_j - \mathbf{x}_i)^{\top} \mathbf{E}
\end{aligned} \tag{164}$$

When  $(i, i_1) \notin \mathcal{E}$  and  $i \neq i_1$ , based on  $\mathbb{E}\{\mathbf{y}_{ij} \mathbf{y}_{i_1 j_1}\} = \mathbb{E}\{\mathbf{y}_{ij}\} \mathbb{E}\{\mathbf{y}_{i_1 j_1}\} = \mathbf{0}_{2 \times 2}$ , we have  $\mathbf{F}_{n_p+1+i,i_1} = \mathbf{0}_{1 \times 2}$ .

Combining the above situations, let  $(\Delta_w)_{i,i_1} = \mathbf{F}_{n_p+1+i,i_1}$ , we have the  $SO(2)$  by  $\mathbb{R}^2$  coupling sub-

matrix:

$$(\Delta_w)_{i,i_1} = \mathbf{F}_{n_p+1+i,i_1} = \begin{cases} \sum_{j \in V_i^+} \delta_{ij}^{-2} (\mathbf{x}_i - \mathbf{x}_j)^\top \mathbf{E} & (i = i_1, j = j_1) \in \mathcal{E} \\ \delta_{ii_1}^{-2} (\mathbf{x}_{i_1} - \mathbf{x}_i)^\top \mathbf{E} & (i, i_1) \in \mathcal{E} \\ \mathbf{0}_{1 \times 2} & \text{else} \end{cases}. \quad (165)$$

### 9.3.3 $\mathbb{R}^2$ by $SO(2)$ coupling sub-matrix

Because FIM is a symmetric matrix, we can get

$$\mathbf{F}_{i_1,n_p+1+i} = \mathbf{F}_{n_p+1+i,i_1}^\top = \begin{cases} \sum_{j \in V_i^+} \delta_{ij}^{-2} \mathbf{E}^\top (\mathbf{x}_i - \mathbf{x}_j) & (i = i_1, j = j_1) \in \mathcal{E} \\ \delta_{ii_1}^{-2} \mathbf{E}^\top (\mathbf{x}_{i_1} - \mathbf{x}_i) & (i, i_1) \in \mathcal{E} \\ \mathbf{0}_{2 \times 1} & \text{else} \end{cases}. \quad (166)$$

### 9.3.4 $SO(2)$ sub-matrix

Based on (152), we have:

$$\begin{aligned} & \mathbf{F}_{n_p+1+i,n_p+1+i_1} \\ &= \mathbb{E}\{\langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_i^R \rangle_{\mathbf{R}} \cdot \langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_{i_1}^R \rangle_{\mathbf{R}}^\top\} \\ &= \mathbb{E}\{\text{trace}(\text{grad}_{\mathbf{R}_i} L(\mathbf{y}; \boldsymbol{\theta})^\top \mathbf{R}_i \mathbf{E}) \cdot \text{trace}(\text{grad}_{\mathbf{R}_{i_1}} L(\mathbf{y}; \boldsymbol{\theta})^\top \mathbf{R}_{i_1} \mathbf{E})^\top\}, \\ &= \mathbb{E}\left\{\sum_{j \in V_i^+} \sum_{j_1 \in V_{i_1}^+} \delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \text{trace}\left(\kappa_{ij} \text{skew}(\mathbf{Z}_{ij}) \mathbf{E} - \frac{1}{2} (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}^\top \mathbf{y}_{ji} \mathbf{I}_{2 \times 2}\right)\right. \\ & \quad \cdot \text{trace}\left(\kappa_{i_1 j_1} \text{skew}(\mathbf{Z}_{ij}) \mathbf{E} - \frac{1}{2} (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}^\top \mathbf{y}_{ij} \mathbf{I}_{2 \times 2}\right)^\top\} \\ & \quad + \mathbb{E}\left\{\sum_{j \in V_i^-} \sum_{j_1 \in V_{i_1}^-} \delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \text{trace}(\kappa_{ij} \text{skew}(\mathbf{Z}_{ij}) \mathbf{E}) \text{trace}(\kappa_{ij} \text{skew}(\mathbf{Z}_{ij}) \mathbf{E})^\top\right\} \\ & \quad + \mathbb{E}\left\{\sum_{j \in V_i^-} \sum_{j_1 \in V_{i_1}^+} \delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \text{trace}(\kappa_{ij} \text{skew}(\mathbf{Z}_{ij}) \mathbf{E}) \text{trace}(\kappa_{ij} \text{skew}(\mathbf{Z}_{ij}) \mathbf{E})^\top\right\} \\ & \quad + \mathbb{E}\left\{\sum_{j \in V_i^+} \sum_{j_1 \in V_{i_1}^-} \delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \text{trace}(\kappa_{ij} \text{skew}(\mathbf{Z}_{ij}) \mathbf{E}) \text{trace}(\kappa_{ij} \text{skew}(\mathbf{Z}_{ij}) \mathbf{E})^\top\right\} \\ &= \mathbb{E}\left\{\langle \sum_{j \in V_i} -\kappa_{ij} \mathbf{R}_i \text{skew}(\mathbf{Z}_{ij}), \mathbf{R}_i \mathbf{E} \rangle_{\mathbf{R}} \cdot \langle \sum_{j \in V_{i_1}} -\kappa_{ij} \mathbf{R}_{i_1} \text{skew}(\mathbf{Z}_{i_1 j}), \mathbf{R}_{i_1} \mathbf{E} \rangle_{\mathbf{R}}\right\} \\ & \quad 0 + \mathbb{E} \sum_{j \in V_i^+} \sum_{j_1 \in V_{i_1}^+} \{\delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \text{trace}((\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}^\top \mathbf{y}_{ij}) \text{trace}((\mathbf{x}_{j_1} - \mathbf{x}_{i_1})^\top \mathbf{R}_{i_1} \mathbf{E}^\top \mathbf{y}_{i_1 j_1})^\top\} \end{aligned} \quad (167)$$

let:

$$\mathbf{D}_{ii_1} = \mathbb{E}\left\{\langle \sum_{j \in V_i} -\kappa_{ij} \mathbf{R}_i \text{skew}(\mathbf{Z}_{ij}), \mathbf{R}_i \mathbf{E} \rangle_{\mathbf{R}} \cdot \langle \sum_{j \in V_{i_1}} -\kappa_{ij} \mathbf{R}_{i_1} \text{skew}(\mathbf{Z}_{i_1 j}), \mathbf{R}_{i_1} \mathbf{E} \rangle_{\mathbf{R}}\right\} \quad (168)$$

Based on (91), we have:

$$\mathbf{D}_{ii_1} = \begin{cases} \sum_{j \in V_i} 2\kappa_{ij} \frac{I_1(2\kappa_{ij})}{I_0(2\kappa_{ij})}, & i = i_1 \\ -2\kappa_{ii_1} \frac{I_1(2\kappa_{ii_1})}{I_0(2\kappa_{ii_1})}, & \{i, i_1\} \in \mathcal{F} \\ 0, & \text{else} \end{cases} \quad (169)$$

Then, let's talk about the second part, we have:

$$\begin{aligned}
\Psi_{ii_1} &= \mathbb{E}\left\{\sum_{j \in V_i^+} \sum_{j_1 \in V_{i_1}^+} \delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \operatorname{trace}\left((\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}^\top \mathbf{y}_{ji}\right) \operatorname{trace}\left((\mathbf{x}_{j_1} - \mathbf{x}_{i_1})^\top \mathbf{R}_{i_1} \mathbf{E}^\top \mathbf{y}_{j_1 i_1}\right)^\top\right\} \\
&= \mathbb{E}\left\{\sum_{j \in V_i^+} \sum_{j \in V_{i_1}^+} \delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \left((\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}^\top \mathbf{y}_{ij}\right) \left((\mathbf{x}_{j_1} - \mathbf{x}_{i_1})^\top \mathbf{R}_{i_1} \mathbf{E}^\top \mathbf{y}_{i_1 j_1}\right)^\top\right\} \\
&= \mathbb{E}\left\{\sum_{j \in V_i^+} \sum_{j \in V_{i_1}^+} \delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}^\top \mathbf{y}_{ij} \mathbf{y}_{i_1 j_1}^\top \mathbf{R}_{i_1}^\top \mathbf{E} (\mathbf{x}_{j_1} - \mathbf{x}_{i_1})\right\} \\
&= \sum_{j \in V_i^+} \sum_{j \in V_{i_1}^+} \delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}^\top \mathbb{E}\{\mathbf{y}_{ij} \mathbf{y}_{i_1 j_1}^\top\} \mathbf{E} \mathbf{R}_{i_1}^\top (\mathbf{x}_{j_1} - \mathbf{x}_{i_1})
\end{aligned} \tag{170}$$

When  $i = i_1$  and  $j = j_1$ , we have:

$$\begin{aligned}
\Psi_{ii_1} &= \sum_{j \in V_i^+} \sum_{j \in V_i^+} \delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}^\top \mathbb{E}\{\mathbf{y}_{ij} \mathbf{y}_{i_1 j_1}^\top\} \mathbf{E} \mathbf{R}_{i_1}^\top (\mathbf{x}_{j_1} - \mathbf{x}_{i_1}) \\
&= \sum_{j \in V_i^+} \delta_{ij}^{-4} (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}^\top \delta_{ij}^2 \mathbf{I}_{2 \times 2} \mathbf{E} \mathbf{R}_{i_1}^\top (\mathbf{x}_{j_1} - \mathbf{x}_{i_1}) \\
&= \sum_{j \in V_i^+} \delta_{ij}^{-2} \|\mathbf{x}_j - \mathbf{x}_i\|^2
\end{aligned} \tag{171}$$

We have:

$$\Psi_{ii_1} = \begin{cases} \sum_{j \in V_i^+} \delta_{ij}^{-2} \|\mathbf{x}_j - \mathbf{x}_i\|^2, & i = i_1 \\ 0, & \text{else} \end{cases} \tag{172}$$

Combine (234) and (172), we have:

$$\mathbf{F}_{n_p+1+i, n_p+1+i_1} = D_{ii_1} + \Psi_{ii_1} = \begin{cases} \sum_{j \in V_i} 2\kappa_{ij} \frac{I_1(2\kappa_{ij})}{I_0(2\kappa_{ij})} + \psi_i, & i = i_1 \\ -2\kappa_{ii_1} \frac{I_1(2\kappa_{ii_1})}{I_0(2\kappa_{ii_1})}, & (i, i_1) \in \mathcal{F} \\ 0, & \text{else} \end{cases} \tag{173}$$

When  $\psi_i = \sum_{j \in V_i^+} \delta_{ij}^{-2} \|\mathbf{x}_j - \mathbf{x}_i\|^2$  is ignored, it is easy to find that the  $SO(2)$  sub-matrix is a weighted Laplacian matrix after using the Kronecker product operation. This weighted Laplacian matrix can be written as  $\mathbf{L}_w^{SO(2)} = \mathbf{L}_{w_{SO(2)}} \otimes \mathbf{I}_{1 \times 1}$ . The  $(i, i_1)$ -th block of the weighted Laplacian matrix  $\mathbf{L}_{w_{SO(2)}}$  is written as:

$$(\mathbf{L}_{w_{SO(2)}})_{i,i_1} = \begin{cases} \sum_{j \in V_i} w_{ij}^{SO(2)} & i = i_1 \\ -w_{i_1 i}^{SO(2)} & (i, i_1) \in \mathcal{E} \\ -w_{i_1, i}^{SO(2)} & (i_1, i) \in \mathcal{E} \\ 0 & \text{else,} \end{cases} \tag{174}$$

where  $w_{ij}^{SO(2)} = 2\kappa_{ij} \frac{I_1(2\kappa_{ij})}{I_0(2\kappa_{ij})}$ .

### 9.3.5 The whole FIM matrix for 2D pose graph

**Theorem 1.** For the 2D case of the pose graph problem (132), given the orthonormal basis (139), the FIM is:

$$\mathcal{I}_{2D} = \begin{bmatrix} \mathbf{L}_w^{\mathbb{R}^2} & \Delta_w^\top \\ \Delta_w & \mathbf{L}_w^{SO(2)} + \operatorname{diag}\{\psi_1, \dots, \psi_{n_p}\} \end{bmatrix}, \tag{175}$$

where  $\mathbf{L}_w^{\mathbb{R}^2}$  is the sub-FIM corresponding to the Euclidean space  $\mathbb{R}^2$ , satisfying  $\mathbf{L}_w^{\mathbb{R}^2} = \mathbf{L}_{w_{\mathbb{R}}} \otimes \mathbf{I}_{2 \times 2}$ .  $\mathbf{L}_{w_{\mathbb{R}}}$  is the weighted Laplacian matrix, of which weight value  $w_{ij}^{\mathbb{R}}$  for  $(i, j)$ -th edge is  $\delta_{ij}^{-2}$ ;  $\mathbf{L}_w^{SO(2)} + \operatorname{diag}\{\psi_1, \dots, \psi_{n_p}\}$  is the sub-FIM corresponding to the  $SO(2)$  Lie group, satisfying  $\mathbf{L}_w^{SO(2)} = \mathbf{L}_{w_{SO(2)}} \otimes \mathbf{I}_{d \times d}$ , where  $d$  is given in Section 3.3.1.  $\mathbf{L}_{w_{SO(2)}}$  is the weighted Laplacian matrix, of which weight value

$w_{ij}^{SO(2)}$  for  $(i, j)$ -th edge is  $2\kappa_{ij}\frac{I_1(2\kappa_{ij})}{I_0(2\kappa_{ij})}$ .  $\psi_i = \sum_{j \in V_i^+} \delta_{ij}^{-2} \|\mathbf{x}_i - \mathbf{x}_j\|_2^2$ ,  $i = 1, 2, \dots, n_p$ ; The  $(i, i_1)$ -th block of the  $SO(2)$  by  $\mathbb{R}^2$  coupling sub-matrix  $\Delta_w$  corresponding to the  $(n_p + 1 + i, i_1)$ -th block of the FIM is:

$$(\Delta_w)_{i,i_1} = \begin{cases} \sum_{j \in V_i^+} \delta_{ij}^{-2} (\mathbf{x}_i - \mathbf{x}_j)^\top \mathbf{E} & i = i_1 \\ \delta_{ii_1}^{-2} (\mathbf{x}_{i_1} - \mathbf{x}_i)^\top \mathbf{E} & (i, i_1) \in \mathcal{E} \\ \mathbf{0}_{1 \times 2} & \text{else.} \end{cases} \quad (176)$$

## 10 3D pose-graph SLAM

Synchronization on  $\mathbb{R}^3 \times SO(3)$  is the problem of estimating a set of rotations  $\mathbf{R}_0, \mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_{n_p} \in SO(3)$  and positions  $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n_p} \in \mathbb{R}^3$  from noisy measurements of some relative rotations  $\mathbf{R}_j \mathbf{R}_i^\top$  and the relative coordinate transformations  $\mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i)$ .

In our estimation problem, the parameter space is:  $\mathcal{P} = \{SO(3) \times \dots \times SO(3)\}_{n_p+1} \times \{\mathbb{R}^3 \times \dots \times \mathbb{R}^3\}_{n_p+1}$ . There are two kinds of edges in the whole graph: one only for  $SO(3)$  and other one for the pose graph.

For  $SO(3)$  graph  $(i, j) \in \mathcal{F}$ , it is noted that  $\mathcal{F}$  is a un-directed edge set, we have one measurement  $\mathbf{H}_{ij} \in SO(3)$  of a form:

$$\mathbf{H}_{ij} = \mathbf{Z}_{ij} \mathbf{R}_j \mathbf{R}_i^\top \quad (177)$$

where  $\mathbf{Z}_{ij}$  is a random value whose distributed function  $\hat{f}_{ij} : SO(3) \rightarrow \mathbb{R}^+$  meets a isotropic Langevin distribution:

$$\begin{aligned} \hat{f}_{ij}(\mathbf{Z}_{ij}) &= \frac{1}{c_3(\kappa_{ij})} \exp(\kappa_{ij} \operatorname{trace}(\mathbf{Z}_{ij})) \\ c_3(\kappa_{ij}) &= \exp(\kappa_{ij})(I_0(2\kappa_{ij}) - I_1(2\kappa_{ij})) \\ I_0(2\kappa_{ij}) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(2\kappa_{ij} \cos(\theta)) d\theta \\ I_1(2\kappa_{ij}) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(2\kappa_{ij} \cos(\theta)) \cos(\theta) d\theta \end{aligned} \quad (178)$$

This PDF meets: 1. smoothness and support; 2. independence; 3. invariance assumptions.

For the pose graph edge,  $(i, j) \in \mathcal{F}$ , it is noted that  $\mathcal{E}$  is a directed edge set, we have a measurement  $\mathbf{p}_{ij} \in \mathbb{R}^3$  of a form:

$$\mathbf{p}_{ij} = \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i) + \mathbf{y}_{ij} \quad (179)$$

where  $\mathbf{y}_{ij}$  is a random vector whose distributed function  $f_{ij} : \mathbb{R}^3 \rightarrow \mathbb{R}^+$  meets a isotropic Gaussian distribution:

$$\begin{aligned} f_{ij}(\mathbf{y}_{ij}) &= \frac{1}{(2\pi)^{3/2} \det(\Sigma_{ij})^{1/2}} \exp(-\frac{1}{2}(\mathbf{y}_{ij}^\top \Sigma_{ij}^{-1} \mathbf{y}_{ij})) \\ \Sigma_{ij} &= \delta_{ij}^2 \mathbf{I}_{3 \times 3} \end{aligned} \quad (180)$$

### 10.1 Synchronization on $\mathbb{R}^3 \times SO(3)$

The log-likelihood of the estimator  $\mathbf{R} = (\mathbf{R}_1, \dots, \mathbf{R}_{n_p})$  and  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_{n_p})$ , given the measurements  $\mathbf{H}_{ij}$  and  $\mathbf{p}_{ij}$ , is given by:

$$L(\mathbf{y}; \boldsymbol{\theta}) = \frac{1}{2} \sum_{(i,j) \in \mathcal{F}} \log \hat{f}_{ij}(\mathbf{H}_{ij} \mathbf{R}_j \mathbf{R}_i^\top) + \sum_{(i,j) \in \mathcal{E}} \log f_{ij}(\mathbf{p}_{ij} - \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i)) \quad (181)$$

### 10.2 Geometry of the parameter space

This section is to define the notions which will be used in the Fisher Information matrix (FIM).

#### 10.2.1 Tangent space

The dimension of the  $SO(3)$  is  $d = n(n-1)/2 = 3$ . The basic tangent space of the identity matrix is;

$$T_I SO(3) = so(3) \triangleq \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} = \boldsymbol{\Omega} \in \mathbb{R}^{3 \times 3} \quad (182)$$

The general tangent space of a rotation  $\mathbf{Q} \in SO(3)$ :

$$T_Q SO(3) = \mathbf{Q} so(3) \triangleq \mathbf{Q} \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} = \mathbf{Q}\boldsymbol{\Omega} \in SO(3) \quad (183)$$

The tangent space of the euclidean space  $\mathbb{R}^3$  is itself:

$$T_X \mathbb{R}^3 \triangleq \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3, \quad (184)$$

### 10.2.2 Orthonormal basis

For the orthonormal basis, we will show its difference with 2D orthonormal basis (139) because of the higher dimension.

In 3D case, define:

$$\mathbf{E}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}, \mathbf{E}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \mathbf{E}_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad (185)$$

the orthonormal basis  $\mathbf{E}^{x,R} = (\mathbf{E}_{0,1}^x, \mathbf{E}_{0,2}^x, \mathbf{E}_{0,3}^x, \dots, \mathbf{E}_{n_p,1}^x, \mathbf{E}_{n_p,2}^x, \mathbf{E}_{n_p,3}^x, \mathbf{E}_{0,1}^R, \mathbf{E}_{0,2}^R, \mathbf{E}_{0,3}^R, \dots, \mathbf{E}_{n_p,1}^R, \mathbf{E}_{n_p,2}^R, \mathbf{E}_{n_p,3}^R)$  of the tangent space  $\mathcal{T}_{(\mathbf{x}, \mathbf{R})}\mathcal{P}$  can be fixed as:

$$\begin{aligned} \mathbf{E}_{i,k}^x &= (\mathbf{E}_{i,k}^X)^\top; \mathbf{0}_{9(n_p+1) \times 3}, \quad i \in \{0, \dots, n_p\}, \quad k = 1, 2, 3, \\ \mathbf{E}_{j,k}^R &= (\mathbf{0}_{3(n_p+1) \times 1}; \mathbf{E}_{j,k}^R)^\top, \quad j \in \{0, \dots, n_p\}, \\ \mathbf{E}_{i,k}^X &= (\mathbf{0}_{1 \times 3}, \dots, \mathbf{0}_{1 \times 3}, \underbrace{1}_{k\text{-th}}, 0, 0, \mathbf{0}_{1 \times 3}, \dots, \mathbf{0}_{1 \times 3})_{1 \times 3(n_p+1)}, \\ \mathbf{E}_{j,k}^R &= (\mathbf{0}_{3 \times 3}, \dots, \mathbf{0}_{3 \times 3}, \underbrace{\mathbf{R}_j \mathbf{E}_k}_{3j+k\text{-th}}, \mathbf{0}_{3 \times 3}, \dots, \mathbf{0}_{3 \times 3})_{3 \times 9(n_p+1)}. \end{aligned} \quad (186)$$

The inner product is defined by:

$$\langle \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2 \rangle_\theta = \begin{cases} \boldsymbol{\Omega}_1^\top \cdot \boldsymbol{\Omega}_2 & \theta = \mathbf{X} \\ \text{trace}(\boldsymbol{\Omega}_1^\top \cdot \boldsymbol{\Omega}_2) & \theta = \mathbf{R} \end{cases}. \quad (187)$$

### 10.2.3 Gradient

Let  $h : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$  is a differential function, we can define the gradient of  $h$  based on a tangent vector field by:

$$\langle \text{grad } h(\mathbf{Q}), \mathbf{Q}\boldsymbol{\Omega} \rangle_R \quad (188)$$

where  $Q\boldsymbol{\Omega}$  is a tangent vector in a tangent space. The gradient function  $\text{grad } h(Q)$  is computed by:

$$\begin{aligned} \text{grad } h(\mathbf{Q}) &= \mathbf{Q} \text{skew}(\mathbf{Q}^\top \nabla h(\mathbf{Q})) \\ \text{skew}(\star) &\triangleq (\star - \star^\top)/2 \end{aligned} \quad (189)$$

## 10.3 FIM for 3D pose-graph SLAM

Definition [13]: Let  $\boldsymbol{\theta} \in \mathcal{P}$  be unknown parameter and  $f(\mathbf{y}; \boldsymbol{\theta})$  be the PDF of the measurement  $\mathbf{y}$  conditioned by  $\boldsymbol{\theta}$  (in this paper, the measurement noises are shown in the PDFs (178) and (180)). Based on the log-likelihood function  $L(\mathbf{y}; \boldsymbol{\theta}) = \log f(\mathbf{y}; \boldsymbol{\theta})$  shown in (181) and the orthonormal basis, the  $(i,j)$ -th element of the FIM is defined as:

$$\mathbf{F}_{(i,j)} = \mathbb{E} \{ \langle \text{grad } L(\mathbf{y}; \boldsymbol{\theta}), e_i \rangle \cdot \langle \text{grad } L(\mathbf{y}; \boldsymbol{\theta}), e_j \rangle \}, \quad (190)$$

where  $e_i$  and  $e_j$  are the  $i$ -th and  $j$ -th bases of the tangent space of the parameters. Expectations are taken w.r.t. the measurement  $\mathbf{y}$ . They will be defined based on the parameter space [2]. It is noted that the FIM is directly decided by the bases. For a parameter space, there may exist different bases.

### 10.3.1 $\mathbb{R}^3$ sub-matrix

For this situation, based on (149), we can directly obtain the sub-matrix corresponding to Euclidean space  $\mathbb{R}^3$ . We have the  $(i, i_1)$ -th block of the sub-matrix  $\mathbf{L}_w^{\mathbb{R}^3}$  corresponding to the  $(i, i_1)$ -th block of the FIM:

$$\mathbf{F}_{ii_1} = (\mathbf{L}_w^{\mathbb{R}^3})_{i,i_1} = (\mathbf{L}_{w_{\mathbb{R}}})_{i,i_1} \otimes \mathbf{I}_{3 \times 3} = \begin{cases} \sum_{j_1 \in V_i^+} \delta_{ij_1}^{-2} \mathbf{I}_{3 \times 3} + \sum_{j_2 \in V_i^-} \delta_{j_2 i}^{-2} \mathbf{I}_{3 \times 3}, & i = i_1 \\ -\delta_{ii_1}^{-2} \mathbf{I}_{3 \times 3}, & (i, i_1) \in \mathcal{E} \text{ or } (i_1, i) \in \mathcal{E} \\ \mathbf{0}, & \text{else} \end{cases} \quad (191)$$

### 10.3.2 $SO(3)$ by $\mathbb{R}^3$ coupling sub-matrix

$$\begin{aligned} \mathbf{F}_{n_p+i+1,i_1}(k) &= \mathbb{E}\{ \langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_i^R \rangle_{\mathbf{R}} \cdot \langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_{i_1}^x \rangle_{\mathbf{x}}^* \} \\ &= \mathbb{E}\{ \text{trace}(\text{grad}_{\mathbf{R}_i} L(\mathbf{y}; \boldsymbol{\theta})^\top \cdot \mathbf{R}_i \mathbf{E}) \cdot \text{grad}_{\mathbf{x}_{i_1}} L(\mathbf{y}; \boldsymbol{\theta}) \}, \end{aligned} \quad (192)$$

Based on (152), we have:

$$\begin{aligned} \text{grad}_{\mathbf{R}_i} L(\mathbf{y}; \boldsymbol{\theta}) &= \text{grad}_{\mathbf{R}_i} \frac{1}{2} \sum_{(i,j) \in \mathcal{F}} \log \hat{f}_{ij}(\mathbf{H}_{ij} \mathbf{R}_i \mathbf{R}_j^\top) + \text{grad}_{\mathbf{R}_i} \sum_{(i,j) \in \mathcal{E}} \log f_{ij}(\mathbf{p}_{ij} - \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i)) \\ &= \sum_{j \in V_i} \kappa_{ij} \mathbf{R}_i \text{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j) + \text{grad}_{\mathbf{R}_i} \sum_{(i,j) \in \mathcal{E}} \log f_{ij}(\mathbf{p}_{ij} - \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i)) \\ &= \sum_{j \in V_i} \kappa_{ij} \mathbf{R}_i \text{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j) + \sum_{(i,j) \in \mathcal{E}} \sum_{k=1}^3 \mathbf{R}_i \mathbf{E}_k \nabla_{\mathbf{R}_i \mathbf{E}_k} \log f_{ij}(\mathbf{p}_{ij} - \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i)) \end{aligned} \quad (193)$$

Let's consider the second part:

$$\begin{aligned} &\nabla_{\mathbf{R}_i \mathbf{E}_k} \log f_{ij}(\mathbf{p}_{ij} - \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i)) \\ &= \nabla_{\mathbf{R}_i \mathbf{E}_k} \log \frac{1}{2\pi \det(\boldsymbol{\Sigma})^{1/2}} \exp(-\frac{1}{2}((\mathbf{p}_{ij} - \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i))^\top \boldsymbol{\Sigma}^{-1}(\mathbf{p}_{ij} - \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i)))) \\ &= \nabla_{\mathbf{R}_i \mathbf{E}_k} \left( \log \frac{1}{2\pi \det(\boldsymbol{\Sigma})^{1/2}} + \log \exp(-\frac{1}{2}((\mathbf{p}_{ij} - \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i))^\top \boldsymbol{\Sigma}^{-1}(\mathbf{p}_{ij} - \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i)))) \right) \\ &= \nabla_{\mathbf{R}_i \mathbf{E}_k} \log \exp(-\frac{1}{2}((\mathbf{p}_{ij} - \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i))^\top \boldsymbol{\Sigma}^{-1}(\mathbf{p}_{ij} - \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i)))) \\ &= \nabla_{\mathbf{R}_i \mathbf{E}_k} -\frac{1}{2}((\mathbf{p}_{ij} - \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i))^\top \boldsymbol{\Sigma}^{-1}(\mathbf{p}_{ij} - \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i))) \\ &= \nabla_{\mathbf{R}_i \mathbf{E}_k} -\frac{1}{2}(\mathbf{y}_{ij}^\top \boldsymbol{\Sigma}^{-1} \mathbf{y}_{ij}) \\ &= \nabla_{\mathbf{y}_{ij}} -\frac{1}{2}(\mathbf{y}_{ij}^\top \boldsymbol{\Sigma}^{-1} \mathbf{y}_{ij}) \nabla_{\mathbf{R}_i \mathbf{E}_k} \mathbf{y}_{ij} \\ &= -\frac{1}{2} \mathbf{y}_{ij}^\top (\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Sigma}^{-T}) \nabla_{\mathbf{R}_i \mathbf{E}_k} \mathbf{y}_{ij} \\ &= -\delta_{ij}^{-2} \mathbf{y}_{ij}^\top \nabla_{\mathbf{R}_i \mathbf{E}_k} \mathbf{y}_{ij} \end{aligned} \quad (194)$$

(1) Based on the basis representation of the gradient function. For  $\nabla_{\mathbf{R}_i} \mathbf{y}_{ij}$ , based on the basis  $\mathbf{R}_i \mathbf{E}_1, \mathbf{R}_i \mathbf{E}_2, \mathbf{R}_i \mathbf{E}_3$ , we have:

$$\begin{aligned} &\nabla_{\mathbf{R}_i \mathbf{E}_1} \mathbf{y}_{ij} &\nabla_{\mathbf{R}_i \mathbf{E}_2} \mathbf{y}_{ji} &\nabla_{\mathbf{R}_i \mathbf{E}_3} \mathbf{y}_{ji} \\ &= \frac{\partial \mathbf{p}_{ij} - \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i)}{\partial \mathbf{R}_i} &= \frac{\partial \mathbf{p}_{ij} - \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i)}{\partial \mathbf{R}_i} &= \frac{\partial \mathbf{p}_{ij} - \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i)}{\partial \mathbf{R}_i} \\ &= \mathbf{0} - \frac{\partial \exp(\delta)^\top \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i)}{\partial \delta} = \mathbf{0} - \frac{\partial \exp(\delta)^\top \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i)}{\partial \delta} = \mathbf{0} - \frac{\partial \exp(\delta)^\top \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i)}{\partial \delta} && (195) \\ &= -(\mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i))_\times &= -(\mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i))_\times &= -(\mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i))_\times \\ &= -\mathbf{E}_1^\top \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i) &= -\mathbf{E}_2^\top \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i) &= -\mathbf{E}_3^\top \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i) \end{aligned}$$

Introduce (195) into (194), we can get:

$$(194) = \delta_{ij}^{-2} \mathbf{y}_{ij}^\top \nabla_{\mathbf{R}_i \mathbf{E}_k} \mathbf{y}_{ij} \\ = \delta_{ij}^{-2} \mathbf{y}_{ij}^\top \mathbf{E}_k^\top \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i) \quad (196)$$

(2) Using the general computation way for gradient on  $SO(3)$ : Let  $\tilde{f}: R_{3 \times 3} \rightarrow R$  be a differentiable function and let  $f = \tilde{f}|SO(3)$  be its restriction to  $SO(3)$ . Let  $\nabla \tilde{f}(\mathbf{Q})$  be the usual gradient of  $\tilde{f}$  in  $R_{3 \times 3}$ . Then, the gradient of  $f$  is easily computed as:

$$\text{grad } f(\mathbf{Q}) = \mathbf{Q} \text{skew}(\mathbf{Q}^\top \cdot \nabla \tilde{f}(\mathbf{Q})), \quad \text{skew}(\star) = (\star - \star^\top)/2 \quad (197)$$

Firstly let's get the usual gradient of  $\tilde{f}$  in  $R_{n \times n}$ :

$$\begin{aligned} \nabla \tilde{f}(\mathbf{Q}) &= \nabla_{\mathbf{R}_i \in R_{n \times n}} \log \frac{1}{2\pi|\Sigma|^{1/2}} \exp(-\frac{1}{2}((\mathbf{p}_{ij} - \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i))^\top \Sigma^{-1}(\mathbf{p}_{ij} - \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i)))) \\ &= \nabla_{\mathbf{R}_i \in R_{n \times n}} -\frac{1}{2}((\mathbf{p}_{ij} - \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i))^\top \Sigma^{-1}(\mathbf{p}_{ij} - \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i))) \\ &= \nabla_{\mathbf{R}_i \in R_{n \times n}} -\delta_{ij}^{-2} \frac{1}{2}((\mathbf{p}_{ij} - \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i))^\top (\mathbf{p}_{ij} - \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i))) \\ &= \nabla_{\mathbf{R}_i \in R_{n \times n}} -\delta_{ij}^{-2} \frac{1}{2} \mathbf{y}_{ij} \mathbf{y}_{ij}^\top \\ &= \delta_{ij}^{-2} (\mathbf{y}_{ij}(\mathbf{x}_j - \mathbf{x}_i)^\top)^\top \\ &= \delta_{ij}^{-2} (\mathbf{x}_j - \mathbf{x}_i) \mathbf{y}_{ij}^\top \end{aligned} \quad (198)$$

Let  $\mathbf{Q} = \mathbf{R}_i$ , we have:

$$\begin{aligned} \text{grad}_{\mathbf{R}_i} f(\mathbf{R}_i) &= \mathbf{R}_i \text{skew}(\mathbf{R}_i^\top \cdot \nabla \tilde{f}(\mathbf{R}_i)) \\ &= \delta_{ij}^{-2} \mathbf{R}_i (\mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i) \mathbf{y}_{ij}^\top - \mathbf{y}_{ij} (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i)/2 \\ &= \frac{1}{2} \delta_{ij}^{-2} ((\mathbf{x}_j - \mathbf{x}_i) \mathbf{y}_{ij}^\top - \mathbf{R}_i \mathbf{y}_{ij} (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i) \end{aligned} \quad (199)$$

Using simulation, we can find that Eq.(199) is equal to the second part of Eq.(200). Combine (193) and (196), we have:

$$\begin{aligned} \text{grad}_{\mathbf{R}_i} L(\mathbf{y}; \boldsymbol{\theta}) &= \text{grad}_{\mathbf{R}_i} \frac{1}{2} \sum_{(i,j) \in \varepsilon} \log \hat{f}_{ij}(\mathbf{H}_{ij} \mathbf{R}_i \mathbf{R}_j^\top) + \text{grad}_{\mathbf{R}_i} \sum_{(i,j) \in \mathcal{E}} \log f_{ij}(\mathbf{p}_{ij} - \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i)) \\ &= \sum_{j \in V_i} \kappa_{ij} \mathbf{R}_i \text{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j) + \sum_{(i,j) \in \mathcal{E}} \delta_{ij}^{-2} \sum_{k=1}^3 \mathbf{R}_i \mathbf{E}_k \mathbf{y}_{ij}^\top \mathbf{E}_k^\top \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i) \end{aligned} \quad (200)$$

For  $\text{grad}_{\mathbf{R}_{i_1}} L(\theta)$ ,  $i_1 < N$ , based on (193), we have:

$$\begin{aligned} \text{grad}_{\mathbf{R}_{i_1}} L(\mathbf{y}; \boldsymbol{\theta}) &= \text{grad}_{\mathbf{R}_{i_1}} \frac{1}{2} \sum_{(i_1, j_1) \in \varepsilon} \log \hat{f}_{i_1 j_1}(\mathbf{H}_{i_1 j_1} \mathbf{R}_{i_1} \mathbf{R}_{j_1}^\top) + \text{grad}_{\mathbf{R}_{i_1}} \sum_{(i_1, j_1) \in \mathcal{E}} \log f_{i_1 j_1}(\mathbf{p}_{i_1 j_1} - \mathbf{R}_{i_1}^\top(\mathbf{x}_{j_1} - \mathbf{x}_{i_1})) \\ &= 0 + \nabla_{\mathbf{x}_{i_1}} \sum_{(i_1, j_1) \in \mathcal{E}} \log f_{i_1 j_1}(\mathbf{y}_{i_1 j_1}) \\ &= \sum_{(i_1, j_1) \in \mathcal{E}} -\delta_{i_1 j_1}^{-2} \mathbf{y}_{i_1 j_1}^\top \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1} \end{aligned} \quad (201)$$

Introduce (200) and (201) into (192),  $\mathbb{E}\{\mathbf{Z}_{ij}\} = \mathbf{0}_{3 \times 3}$  and  $\mathbb{E}\{\mathbf{y}_{ij}\} = \mathbf{0}_{3 \times 1}$ , we have:

$$\begin{aligned}
\mathbf{F}_{n_p+i+1,i_1}(k) &= \mathbb{E}\{\langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), e_i \rangle_{\theta} \cdot \langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), e_{i_1} \rangle_{\theta}^{\top}\} \\
&= \mathbb{E}\left\{\sum_{(i,j) \in E} \text{trace}\left(\left(-k_{ij}\mathbf{R}_i \text{skew}(\mathbf{Z}_{ij}) + \delta_{ij}^{-2} \sum_{k=1}^3 \mathbf{R}_i \mathbf{E}_k \mathbf{y}_{ij}^{\top} \mathbf{E}^{\top} \mathbf{R}_i^{\top} (\mathbf{x}_j - \mathbf{x}_i)\right)^{\top} \mathbf{R}_i \mathbf{E}_k\right)\right. \\
&\quad \left.\sum_{(i_1,j_1) \in E} -\delta_{i_1 j_1}^{-2} \mathbf{y}_{i_1 j_1}^{\top} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1}\right\} \\
&= \mathbb{E}\left\{\sum_{(i,j) \in E} \text{trace}\left(\left(\delta_{ij}^{-2} \sum_{k=1}^3 \mathbf{R}_i \mathbf{E}_k \mathbf{y}_{ij}^{\top} \mathbf{E}^{\top} \mathbf{R}_i^{\top} (\mathbf{x}_j - \mathbf{x}_i)\right)^{\top} \mathbf{R}_i \mathbf{E}_k\right) \sum_{(i_1,j_1) \in E} -\delta_{i_1 j_1}^{-2} \mathbf{y}_{i_1 j_1}^{\top} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1}\right\} \\
&= \mathbb{E}\left\{\sum_{(i,j) \in E} \text{trace}\left(\left(\delta_{ij}^{-2} \sum_{k=1}^3 \mathbf{R}_i \mathbf{E}_k \mathbf{y}_{ij}^{\top} \mathbf{E}^{\top} \mathbf{R}_i^{\top} (\mathbf{x}_j - \mathbf{x}_i)\right)^{\top} \mathbf{R}_i \mathbf{E}_k\right) \sum_{j_1 \in V_{i_1}^+} -\delta_{i_1 j_1}^{-2} \mathbf{y}_{i_1 j_1}^{\top} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1}\right\} \\
&\quad + \mathbb{E}\left\{\sum_{(i,j) \in E} \text{trace}\left(\left(\delta_{ij}^{-2} \sum_{k=1}^3 \mathbf{R}_i \mathbf{E}_k \mathbf{y}_{ij}^{\top} \mathbf{E}^{\top} \mathbf{R}_i^{\top} (\mathbf{x}_j - \mathbf{x}_i)\right)^{\top} \mathbf{R}_i \mathbf{E}_k\right) \sum_{j_2 \in V_{i_1}^-} -\delta_{j_2 i_1}^{-2} \mathbf{y}_{j_2 i_1}^{\top} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{j_2 i_1}\right\}
\end{aligned} \tag{202}$$

When  $(i, j) = (i_1, j_1) \in \mathcal{E}$ , we know that  $(\mathbf{x}_i - \mathbf{x}_j)^{\top} \mathbf{R}_i \mathbf{E} \mathbf{y}_{ij} \in \mathbb{R}$  and  $\nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1}(\mathbf{x}_{i_1}) = \mathbf{R}_i^{\top}$  we have:

$$\begin{aligned}
\mathbf{F}_{n_p+i+1,i_1}(k) &= \mathbb{E}\{\langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_i^R \rangle_{\mathbf{R}} \cdot \langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), \mathbf{E}_{i_1}^x \rangle_{\mathbf{X}}^{\top}\} \\
&= \mathbb{E}\{\text{trace}(\text{grad}_{\mathbf{R}_i} L(\mathbf{y}; \boldsymbol{\theta})^{\top} \cdot \mathbf{R}_i \mathbf{E}) \cdot \text{grad}_{\mathbf{x}_{i_1}} L(\mathbf{y}; \boldsymbol{\theta})\}, \\
&= \mathbb{E}\left\{\sum_{(i,j) \in \mathcal{E}} \text{trace}\left(\left(\delta_{ij}^{-2} \sum_{k=1}^3 \mathbf{R}_i \mathbf{E}_k \mathbf{y}_{ij}^{\top} \mathbf{E}_k^{\top} \mathbf{R}_i^{\top} (\mathbf{x}_i - \mathbf{x}_j)\right)^{\top} \mathbf{R}_i \mathbf{E}_k\right) \delta_{i_1 j_1}^{-2} \mathbf{y}_{i_1 j_1}^{\top} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1}\right\} + 0 \\
&= \mathbb{E}\left\{\sum_{(i,j) \in \mathcal{E}} \text{trace}\left(\left(\delta_{ij}^{-2} \mathbf{R}_i \mathbf{E}_k \mathbf{y}_{ij}^{\top} \mathbf{E}_k^{\top} \mathbf{R}_i^{\top} (\mathbf{x}_i - \mathbf{x}_j)\right)^{\top} \mathbf{R}_i \mathbf{E}_k\right) \delta_{i_1 j_1}^{-2} \mathbf{y}_{i_1 j_1}^{\top} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1}\right\} \\
&= \mathbb{E}\left\{\sum_{(i,j) \in \mathcal{E}} \delta_{ij}^{-2} \text{trace}\left((\mathbf{x}_i - \mathbf{x}_j)^{\top} \mathbf{R}_i \mathbf{E}_k \mathbf{y}_{ij} (\mathbf{E}_k \mathbf{R}_i)^{\top} \mathbf{R}_i \mathbf{E}_k\right) \delta_{i_1 j_1}^{-2} \mathbf{y}_{i_1 j_1}^{\top} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1}\right\} \\
&= \mathbb{E}\left\{\sum_{(i,j) \in \mathcal{E}} \delta_{ij}^{-2} \text{trace}\left((\mathbf{x}_i - \mathbf{x}_j)^{\top} \mathbf{R}_i \mathbf{E}_k \mathbf{y}_{ij} \mathbf{E}_k^{\top} \mathbf{E}_k\right) \delta_{i_1 j_1}^{-2} \mathbf{y}_{i_1 j_1}^{\top} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1}\right\} \\
&= \mathbb{E}\left\{\sum_{(i,j) \in \mathcal{E}} \delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} (\mathbf{x}_i - \mathbf{x}_j)^{\top} \mathbf{R}_i \mathbf{E}_k \mathbf{y}_{ij} \mathbf{y}_{i_1 j_1}^{\top} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1}\right\} \\
&= \sum_{(i,j) \in \mathcal{E}} \delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} (\mathbf{x}_i - \mathbf{x}_j)^{\top} \mathbf{R}_i \mathbf{E}_k \mathbb{E}\{\mathbf{y}_{ij} \mathbf{y}_{i_1 j_1}^{\top}\} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1} \\
&= \sum_{(i,j) \in \mathcal{E}} \delta_{ij}^{-2} (\mathbf{x}_i - \mathbf{x}_j)^{\top} \mathbf{R}_i \mathbf{E}_k \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{i_1 j_1} \\
&= \sum_{(i,j) \in \mathcal{E}} \delta_{ij}^{-2} (\mathbf{x}_i - \mathbf{x}_j)^{\top} \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^{\top}
\end{aligned} \tag{203}$$

When  $(i, j) = (j_1, i_1) \in \mathcal{E}$ , we know that  $(\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E} \mathbf{y}_{ij} \in \mathbb{R}$  and  $\nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{j_1 i_1} = -\mathbf{R}_i^\top$  we have:

$$\begin{aligned}
\mathbf{F}_{n_p+i+1,i_1}(k) &= \mathbb{E}\{\langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), e_i \rangle_\theta \cdot \langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), e_{i_1} \rangle_\theta^\top\} \\
&= 0 + \mathbb{E}\{\text{trace}\left(\left(\delta_{ij}^{-2} \sum_{k=1}^3 \mathbf{R}_i \mathbf{E}_k \mathbf{y}_{ij}^\top \mathbf{E}_k^\top \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i)\right)^\top \mathbf{R}_i \mathbf{E}_k\right) \delta_{j_1 i_1}^{-2} \mathbf{y}_{j_1 i_1}^\top \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{j_1 i_1}\} \\
&= -\mathbb{E}\{\text{trace}\left((\delta_{ij}^{-2} \mathbf{R}_i \mathbf{E}_k \mathbf{y}_{ij}^\top \mathbf{E}_k^\top \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i))^\top \mathbf{R}_i \mathbf{E}_k\right) \delta_{j_1 i_1}^{-2} \mathbf{y}_{j_1 i_1}^\top \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{j_1 i_1}\} \\
&= -\mathbb{E}\{\delta_{ij}^{-2} \text{trace}\left((\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}_k \mathbf{y}_{ij} (\mathbf{E}_k \mathbf{R}_i)^\top \mathbf{R}_i \mathbf{E}_k\right) \delta_{j_1 i_1}^{-2} \mathbf{y}_{j_1 i_1}^\top \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{j_1 i_1}\} \quad (204) \\
&= -\mathbb{E}\{\delta_{ij}^{-2} \text{trace}\left((\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}_k \mathbf{y}_{ij}\right) \delta_{j_1 i_1}^{-2} \mathbf{y}_{j_1 i_1}^\top \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{j_1 i_1}\} \\
&= -\mathbb{E}\{\delta_{ij}^{-2} \delta_{j_1 i_1}^{-2} (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}_k \mathbf{y}_{ij} \mathbf{y}_{j_1 i_1}^\top \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{j_1 i_1}\} \\
&= -\delta_{ij}^{-2} \delta_{j_1 i_1}^{-2} (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}_k \mathbb{E}\{\mathbf{y}_{ij} \mathbf{y}_{j_1 i_1}^\top\} \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{j_1 i_1} \\
&= -\delta_{ij}^{-2} (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}_k \nabla_{\mathbf{x}_{i_1}} \mathbf{y}_{j_1 i_1} \\
&= \delta_{ij}^{-2} (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top
\end{aligned}$$

So, let  $(\Delta_w^{3D})_{i,i_1} = \mathbf{F}_{n_p+1+i,i_1}$ , we have the  $(i, i_1)$ -th block of the  $SO(3)$  by  $\mathbb{R}^3$  coupling sub-matrix  $\Delta_w$  corresponding to the  $(n_p + 1 + i, i_1)$ -th block of the FIM:

$$(\Delta_w^{3D})_{i,i_1} = \begin{cases} \begin{bmatrix} \sum_{j \in V_i^+} \delta_{ij}^{-2} (\mathbf{x}_i - \mathbf{x}_j)^\top \mathbf{R}_i \mathbf{E}_1 \mathbf{R}_i^\top \\ \sum_{j \in V_i^+} \delta_{ij}^{-2} (\mathbf{x}_i - \mathbf{x}_j)^\top \mathbf{R}_i \mathbf{E}_2 \mathbf{R}_i^\top \\ \sum_{j \in V_i^+} \delta_{ij}^{-2} (\mathbf{x}_i - \mathbf{x}_j)^\top \mathbf{R}_i \mathbf{E}_3 \mathbf{R}_i^\top \end{bmatrix} & i = i_1 \\ \begin{bmatrix} \delta_{i_1 i_1}^{-2} (\mathbf{x}_{i_1} - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}_1 \mathbf{R}_i^\top \\ \delta_{i_1 i_1}^{-2} (\mathbf{x}_{i_1} - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}_2 \mathbf{R}_i^\top \\ \delta_{i_1 i_1}^{-2} (\mathbf{x}_{i_1} - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}_3 \mathbf{R}_i^\top \end{bmatrix} & (i, i_1) \in \mathcal{E} \\ \mathbf{0}_{3 \times 3} & \text{else,} \end{cases} \quad (205)$$

### 10.3.3 $\mathbb{R}^3$ by $SO(3)$ coupling sub-matrix

Because FIM is a symmetric matrix, we can get:

$$\mathbf{F}_{1,n_p+i+i_1} = \mathbf{F}_{n_p+i+1,i_1}^\top \quad (206)$$

#### 10.3.4 $SO(3)$ sub-matrix

For this situation, we have:

$$\begin{aligned}
& \mathbf{F}_{n_p+i+i,n_p+i+i_1}(k, l) \\
&= \mathbb{E}\{\langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), e_i \rangle_\theta \cdot \langle \text{grad}L(\mathbf{y}; \boldsymbol{\theta}), e_{i_1} \rangle_\theta^\top\} \\
&= \mathbb{E}\{\text{trace}(\text{grad}_{\mathbf{R}_i} L(\mathbf{y}; \boldsymbol{\theta})^\top \cdot \mathbf{R}_i \mathbf{E}_k) \cdot \text{trace}(\text{grad}_{\mathbf{R}_{i_1}} L(\mathbf{y}; \boldsymbol{\theta})^\top \cdot \mathbf{R}_{i_1} \mathbf{E}_l)^\top\} \\
&= \mathbb{E}\left\{\sum_{(i,j) \in \mathcal{E}} \text{trace}\left(\left(\mathbf{R}_i \text{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j) + \delta_{ij}^{-2} \sum_{k=1}^3 \mathbf{R}_i \mathbf{E}_k \mathbf{y}_{ij}^\top \mathbf{E}_k^\top \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i)\right)^\top \mathbf{R}_i \mathbf{E}_k\right)\right. \\
&\quad \left.\sum_{(i_1,j_1) \in \mathcal{E}} \text{trace}\left(\left(\mathbf{R}_i \text{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j) + \delta_{i_1 j_1}^{-2} \sum_{k=1}^3 \mathbf{R}_{i_1} \mathbf{E}_k \mathbf{y}_{i_1 j_1}^\top \mathbf{E}_k^\top \mathbf{R}_{i_1}^\top (\mathbf{x}_{j_1} - \mathbf{x}_{i_1})\right)^\top \mathbf{R}_{i_1} \mathbf{E}_l\right)\right\} \\
&= \mathbb{E}\left\{\sum_{j \in V_i} \sum_{j_1 \in V_{i_1}} \langle \mathbf{R}_i \text{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j) \mathbf{R}_i^\top, \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top \rangle_R \langle \mathbf{R}_i \text{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j) \mathbf{R}_i^\top, \mathbf{R}_i \mathbf{E}_l \mathbf{R}_i^\top \rangle_R + 0 + \right. \\
&\quad \mathbb{E} \sum_{j \in V_i^+} \sum_{j_1 \in V_{i_1}^+} \left\{\delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \text{trace}\left(\sum_{k=1}^3 ((\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}_k \mathbf{y}_{ij} \mathbf{E}_k^\top \mathbf{R}_i^\top) \mathbf{R}_i \mathbf{E}_k\right) \right. \\
&\quad \left.\left.\text{trace}\left(\sum_{k=1}^3 ((\mathbf{x}_{j_1} - \mathbf{x}_{i_1})^\top \mathbf{R}_{i_1} \mathbf{E}_k \mathbf{y}_{i_1 j_1} \mathbf{E}_k^\top \mathbf{R}_{i_1}^\top) \mathbf{R}_{i_1} \mathbf{E}_l\right)\right)\right\} \\
&= \mathbb{E}\left\{\sum_{j \in V_i} \sum_{j_1 \in V_{i_1}} \langle \mathbf{R}_i \text{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j) \mathbf{R}_i^\top, \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top \rangle_R \langle \mathbf{R}_i \text{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j) \mathbf{R}_i^\top, \mathbf{R}_i \mathbf{E}_l \mathbf{R}_i^\top \rangle_R\right\} + 0 \\
&\quad 0 + \mathbb{E} \sum_{j \in V_i^+} \sum_{j_1 \in V_{i_1}^+} \left\{\delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \text{trace}\left((\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}_k \mathbf{y}_{ij}\right) \text{trace}\left((\mathbf{x}_{j_1} - \mathbf{x}_{i_1})^\top \mathbf{R}_{i_1} \mathbf{E}_l^\top \mathbf{y}_{i_1 j_1}\right)^\top\right\} \\
&\tag{207}
\end{aligned}$$

Let:

$$\mathbf{D}_{ii_1}(k, l) = \mathbb{E}\left\{\sum_{j \in V_i} \sum_{j_1 \in V_{i_1}} \text{trace} \langle \mathbf{R}_i \text{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j) \mathbf{R}_i^\top, \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top \rangle_R \langle \mathbf{R}_i \text{skew}(\mathbf{R}_j^\top \mathbf{Z}_{ij}^\top \mathbf{R}_j) \mathbf{R}_i^\top, \mathbf{R}_i \mathbf{E}_l \mathbf{R}_i^\top \rangle_R\right\}
\tag{208}$$

Based on (125), we have:

$$\mathbf{D}_{ii_1} = \begin{cases} \frac{1}{3} \left( \sum_{j \in V_i} \frac{k_{ij}^2 (2I_0(2k_{ij}) - I_1(2k_{ij}) - 2I_2(2k_{ij}) + I_3(2k_{ij}))}{2I_0(2k_{ij}) - 2I_1(2k_{ij})} \right) \mathbf{I}_{3 \times 3} & i = i_1 \\ -\frac{1}{3} \left( \frac{k_{i_1}^2 (2I_0(2k_{i_1}) - I_1(2k_{i_1}) - 2I_2(2k_{i_1}) + I_3(2k_{i_1}))}{2I_0(2k_{i_1}) - 2I_1(2k_{i_1})} \right) \mathbf{I}_{3 \times 3} & (i, i_1) \in \mathcal{F} \\ \mathbf{0} & \text{else} \end{cases}
\tag{209}$$

Then, let's talk about the second part, we have:

$$\begin{aligned}
\Psi_{ii_1} &= \mathbb{E} \sum_{j \in V_i^+} \sum_{j_1 \in V_{i_1}^+} \left\{ \delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \text{trace}\left((\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}_k \mathbf{y}_{ij}\right) \text{trace}\left((\mathbf{x}_{j_1} - \mathbf{x}_{i_1})^\top \mathbf{R}_{i_1}^\top \mathbf{E}_l^\top \mathbf{y}_{i_1 j_1}\right)^\top \right\} \\
&= \mathbb{E} \sum_{j \in V_i^+} \sum_{j_1 \in V_{i_1}^+} \left\{ \delta_{ij}^{-2} \delta_{i_1 j_1}^{-2} \left( (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}_k \mathbf{y}_{ij} \right) \left( (\mathbf{x}_{j_1} - \mathbf{x}_{i_1})^\top \mathbf{R}_{i_1}^\top \mathbf{E}_l^\top \mathbf{y}_{i_1 j_1} \right)^\top \right\}
\end{aligned}
\tag{210}$$

When  $i = i_1$  and  $j = j_1$ , we have:

$$\begin{aligned}
\Psi_i &= \mathbb{E} \sum_{j \in V_i^+} \left\{ \delta_{ij}^{-4} \left( (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}_k \mathbf{y}_{ij} \right) \mathbf{y}_{i_1 j_1}^\top \mathbf{E}_l^\top \mathbf{R}_{i_1}^\top (\mathbf{x}_{j_1} - \mathbf{x}_{i_1}) \right\} \\
&= \sum_{j \in V_i^+} \delta_{ij}^{-4} (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}_k \mathbb{E}\{\mathbf{y}_{ij} \mathbf{y}_{i_1 j_1}^\top\} \mathbf{E}_l^\top \mathbf{R}_{i_1}^\top (\mathbf{x}_{j_1} - \mathbf{x}_{i_1}) \\
&= \sum_{j \in V_i^+} \delta_{ij}^{-2} (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}_k \mathbf{E}_l^\top \mathbf{R}_{i_1}^\top (\mathbf{x}_{j_1} - \mathbf{x}_{i_1})
\end{aligned}
\tag{211}$$

When  $k = l$ ,  $\mathbf{E}_k \mathbf{E}_l^T = \mathbf{I}_{3 \times 3,k} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ ,  $k = 3$ ,  $\mathbf{E}_k \mathbf{E}_l^T = \mathbf{I}_{3 \times 3,k} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ ,  $k = 2$  and  
 $\mathbf{E}_k \mathbf{E}_l^T = \mathbf{I}_{3 \times 3,k} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ ,  $k = 1$ ; when  $k \neq l$ ,  $\mathbf{E}_k \mathbf{E}_l^T = \mathbf{I}_{3 \times 3,kl} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ ,  $k = 1$  and  $l = 3$ ,  
 $\mathbf{E}_k \mathbf{E}_l^T = \mathbf{I}_{3 \times 3,kl} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}$ ,  $k = 3$  and  $l = 1$ ,  $\mathbf{E}_k \mathbf{E}_l^T = \mathbf{I}_{3 \times 3,kl} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ ,  $k = 2$  and  $l = 1$ ,  
 $\mathbf{E}_k \mathbf{E}_l^T = \mathbf{I}_{3 \times 3,kl} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ ,  $k = 1$  and  $l = 2$ ,  $\mathbf{E}_k \mathbf{E}_l^T = \mathbf{I}_{3 \times 3,kl} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}$ ,  $k = 2$  and  $l = 3$   
and  $\mathbf{E}_k \mathbf{E}_l^T = \mathbf{I}_{3 \times 3,kl} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ ,  $k = 3$  and  $l = 2$ , we have:  

$$\Psi_{ii_1} = \begin{cases} \begin{bmatrix} \psi_i^{11} & \psi_i^{12} & \psi_i^{13} \\ \psi_i^{21} & \psi_i^{22} & \psi_i^{23} \\ \psi_i^{31} & \psi_i^{32} & \psi_i^{33} \end{bmatrix} & (i, i_1) \in \mathcal{E} \\ 0 & \text{else} \end{cases}$$

$$\begin{aligned} \psi_i^{11} &= \sum_{j \in V_i^+} \delta_{ij}^{-2} (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{I}_{3 \times 3,1} \mathbf{R}_{i_1}^\top (\mathbf{x}_{j_1} - \mathbf{x}_{i_1}) \\ \psi_i^{22} &= \sum_{j \in V_i^+} \delta_{ij}^{-2} (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{I}_{3 \times 3,2} \mathbf{R}_{i_1}^\top (\mathbf{x}_{j_1} - \mathbf{x}_{i_1}) \\ \psi_i^{33} &= \sum_{j \in V_i^+} \delta_{ij}^{-2} (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{I}_{3 \times 3,3} \mathbf{R}_{i_1}^\top (\mathbf{x}_{j_1} - \mathbf{x}_{i_1}) \\ \psi_i^{21} &= \psi_i^{12} = \sum_{j \in V_i^+} \delta_{ij}^{-2} (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{I}_{3 \times 3,12} \mathbf{R}_{i_1}^\top (\mathbf{x}_{j_1} - \mathbf{x}_{i_1}) \\ \psi_i^{31} &= \psi_i^{13} = \sum_{j \in V_i^+} \delta_{ij}^{-2} (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{I}_{3 \times 3,13} \mathbf{R}_{i_1}^\top (\mathbf{x}_{j_1} - \mathbf{x}_{i_1}) \\ \psi_i^{32} &= \psi_i^{23} = \sum_{j \in V_i^+} \delta_{ij}^{-2} (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{I}_{3 \times 3,23} \mathbf{R}_{i_1}^\top (\mathbf{x}_{j_1} - \mathbf{x}_{i_1}) \end{aligned} \tag{212}$$

Combine (209) and (212), we have:

$$\begin{aligned} \mathbf{F}_{n_p+i+i, n_p+i+i_1} &= \mathbf{D}_{ii_1} + \Psi_{ii_1} \\ &= \begin{cases} \frac{1}{3} \left( \sum_{j \in V_i} \frac{k_{ij}^2 (2I_0(2k_{ij}) - I_1(2k_{ij}) - 2I_2(2k_{ij}) + I_3(2k_{ij}))}{2I_0(2k_{ij}) - 2I_1(2k_{ij})} \right) \mathbf{I}_{3 \times 3} + \begin{bmatrix} \psi_i^{11} & \psi_i^{12} & \psi_i^{13} \\ \psi_i^{21} & \psi_i^{22} & \psi_i^{23} \\ \psi_i^{31} & \psi_i^{32} & \psi_i^{33} \end{bmatrix} & i = i_1 \\ -\frac{1}{3} \left( \frac{k_{ii_1}^2 (2I_0(2k_{ii_1}) - I_1(2k_{ii_1}) - 2I_2(2k_{ii_1}) + I_3(2k_{ii_1}))}{2I_0(2k_{ii_1}) - 2I_1(2k_{ii_1})} \right) \mathbf{I}_{3 \times 3} & (i, i_1) \in \mathcal{F} \\ \mathbf{0}_{3 \times 3} & \text{else} \end{cases} \end{aligned} \tag{213}$$

Finally, the  $(n_p + 1 + i, n_p + 1 + i_1)$ -th block of the FIM corresponding to the  $SO(3)$  sub-matrix are written as:

$$\mathbf{F}_{n_p+1+i, n_p+1+i_1} = \begin{cases} \sum_{j \in V_i} \frac{\omega_{ij}}{3} \mathbf{I}_{3 \times 3} + \Psi_i & i = i_1 \\ -\frac{\omega_{ii_1}}{3} \mathbf{I}_{3 \times 3} & (i, i_1) \in \mathcal{E} \\ -\frac{\omega_{i_1i}}{3} \mathbf{I}_{3 \times 3} & (i_1, i) \in \mathcal{E} \\ \mathbf{0}_{3 \times 3} & \text{else}, \end{cases} \tag{214}$$

where  $\mathbf{F}_{n_p+1+i, n_p+1+i_1} \in \mathbb{R}^{3 \times 3}$ ,  $\Psi_i$  is shown in the equation (218). When  $\Psi_i$  is ignored, it is easy to find that the  $SO(3)$  sub-matrix is a weighted Laplacian matrix after using the Kronecker product operation. This matrix can be written as  $\mathbf{L}_w^{SO(3)} = \mathbf{L}_{w_{SO(3)}} \otimes \mathbf{I}_{3 \times 3}$ . The  $(i, i_1)$ -th block of the weighted Laplacian

matrix  $\mathbf{L}_{w_{SO(3)}}$  is:

$$(\mathbf{L}_{w_{SO(3)}})_{i,i_1} = \begin{cases} \sum_{j \in V_i} w_{ij}^{SO(3)} & i = i_1 \\ -w_{i_1,i_1}^{SO(3)} & (i, i_1) \in \mathcal{E} \\ -w_{i_1,i}^{SO(3)} & (i_1, i) \in \mathcal{E} \\ 0 & \text{else,} \end{cases} \quad (215)$$

where  $w_{i,j}^{SO(3)} = \frac{\omega_{ij}}{3}$ .

### 10.3.5 The whole FIM matrix for 3D pose graph

We can get the final FIM by:

$$\mathbf{I}_{3D} = \begin{pmatrix} \mathbf{L}_w^{\mathbb{R}^3} & \Delta_w^{3D\top} \\ \Delta_w^{3D} & \mathbf{L}_w^{SO(3)} + \text{diag}\{\Psi_1, \dots, \Psi_{n_p}\} \end{pmatrix} \quad (216)$$

**Theorem 2.** For the 3D case of the pose graph problem (132), given the orthonormal basis (186), the FIM is:

$$\mathbf{I}_{3D} = \begin{bmatrix} \mathbf{L}_w^{\mathbb{R}^3} & \Delta_w^{3D\top} \\ \Delta_w^{3D} & \mathbf{L}_w^{SO(3)} + \text{diag}\{\Psi_1, \dots, \Psi_{n_p}\} \end{bmatrix}, \quad (217)$$

where  $\mathbf{L}_w^{\mathbb{R}^3}$  is the sub-FIM corresponding to the Euclidean space  $\mathbb{R}^3$ , satisfying  $\mathbf{L}_w^{\mathbb{R}^3} = \mathbf{L}_{w_{\mathbb{R}}} \otimes \mathbf{I}_{3 \times 3}$ .  $\mathbf{L}_{w_{\mathbb{R}}}$  is the same as that in Theorem 1;  $\mathbf{L}_w^{SO(3)} + \text{diag}\{\Psi_1, \dots, \Psi_{n_p}\}$  is the sub-FIM corresponding to the  $SO(3)$  Lie group, satisfying  $\mathbf{L}_w^{SO(3)} = \mathbf{L}_{w_{SO(3)}} \otimes \mathbf{I}_{d \times d}$ , where  $d$  is given in Section 3.3.1.  $\mathbf{L}_{w_{SO(3)}}$  is the weighted Laplacian matrix, of which weight value  $w_{ij}^{SO(3)}$  for  $(i, j)$ -th edge is  $\frac{1}{3} \frac{\kappa_{ij}^2 (2I_0(2\kappa_{ij}) - I_1(2\kappa_{ij}) - 2I_2(2\kappa_{ij}) + I_3(2\kappa_{ij}))}{2I_0(2\kappa_{ij}) - 2I_1(2\kappa_{ij})}$ .  $\Psi_i$  satisfies:

$$\begin{aligned} \Psi_i &= \begin{bmatrix} \psi_i^{11} & \psi_i^{12} & \psi_i^{13} \\ \psi_i^{12} & \psi_i^{22} & \psi_i^{21} \\ \psi_i^{13} & \psi_i^{21} & \psi_i^{33} \end{bmatrix}, \quad i = 1, 2, \dots, n_p \\ \psi_i^{kl} &= \sum_{j \in V_i^+} \delta_{ij}^{-2} (\mathbf{x}_i - \mathbf{x}_j)^\top \mathbf{R}_i \mathbf{I}_{3 \times 3}^{k,l} \mathbf{R}_i^\top (\mathbf{x}_i - \mathbf{x}_j); \end{aligned} \quad (218)$$

Let  $\boldsymbol{\varsigma}_{ij}^k = (\delta_{ij}^{-2} (\mathbf{x}_j - \mathbf{x}_i)^\top \mathbf{R}_i \mathbf{E}_k \mathbf{R}_i^\top)^\top$ ,  $k = 1, 2, 3$ , we have the  $(i, i_1)$ -th block of the  $SO(3)$  by  $\mathbb{R}^3$  coupling sub-matrix  $\Delta_w^{3D}$  corresponding to the  $(n_p + 1 + i, i_1)$ -th block of the FIM:

$$(\Delta_w^{3D})_{i,i_1} = \begin{cases} \begin{bmatrix} \sum_{j \in V_i^+} \boldsymbol{\varsigma}_{ij}^1 & \sum_{j \in V_i^+} \boldsymbol{\varsigma}_{ij}^2 & \sum_{j \in V_i^+} \boldsymbol{\varsigma}_{ij}^3 \end{bmatrix}^\top & i = i_1 \\ \begin{bmatrix} -\boldsymbol{\varsigma}_{ii_1}^1 & -\boldsymbol{\varsigma}_{ii_1}^2 & -\boldsymbol{\varsigma}_{ii_1}^3 \end{bmatrix}^\top & (i, i_1) \in \mathcal{E} \\ \mathbf{0}_{3 \times 3} & \text{else.} \end{cases} \quad (219)$$

## 11 CRLB for pose-graph SLAM

Classical CRLB gives a lower bound on the covariance matrix  $\mathbf{C}$  of any unbiased estimator for an estimation problem in  $\mathbb{R}^n$ . In terms of the FIM  $\mathbf{F} = \mathbf{I}_{nD}$ ,  $n = 2, 3$  of that problem, the classical result reads  $\mathbf{C} \succeq \mathbf{F}^{-1}$ . However, because our parameter space  $\mathcal{P}$  is a manifold instead of a flat Euclidean space, the CRLB takes up the more general form  $\mathbf{C} \succeq \mathbf{F}^{-1} + \text{curvature terms}$  [25].

Inspired by [13], we also show that when the signal-to-noise ratio (SNR) is large enough, the curvature terms will become negligible. The CRLB is the asymptotic bound, which means only the leading-order curvature term has been computed. For SLAM, the non-flat property of the parameter space comes from the rotation group, so, in order to limit the Taylor truncation error, the SNR  $snr$  is defined as<sup>3</sup>:

$$snr = \frac{(n_p + 1 - |\mathcal{A}|) \mathbb{E}\{\text{dist}^2(\mathbf{Z}_{uni}^{\mathbf{R}}, \mathbf{I}_{3 \times 3})\}}{d^2 \text{trace}(\mathbf{I}^{SO(3)-1})}, \quad (220)$$

---

<sup>3</sup>Because the SNR will be used in the 3D case only, we only show the formulation for 3D pose-graph.

where  $\mathbf{I}^{SO(3)^{-1}}$  is the sub-matrix of the inverse function of the FIM  $\mathbf{I}_{3D}^{-1}$  corresponding to the  $SO(3)$  Lie group,  $|\mathcal{A}|$  means the number of anchored nodes (in general  $|\mathcal{A}| = 1$  for SLAM), the expectation is taken w.r.t.  $\mathbf{Z}_{uni}^R$ , uniformly distributed over  $SO(3)$ , it is easy to find that the SNR is inversely proportional to the trace of the inverse of the information matrix. In other words, when the uncertainty corresponding to the rotation group is small (large-scale  $\mathbf{I}^{SO(3)}$ ), the SNR  $snr$  will be large. For the molecule of the SNR  $snr$ ,  $n_p + 1 - |\mathcal{A}|$  means the number of the estimated poses and  $\mathbb{E}\{\text{dist}^2(\mathbf{Z}_{uni}^R, \mathbf{I}_{3 \times 3})\}$  shows a suitable constant to balance the denominator.

## 11.1 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 introduce a lemma [42]:

**Lemma 6.** *Let  $\mathcal{M} = \mathcal{M}_1 \times \mathcal{M}_2$  be the product of two Riemannian manifolds,  $R$  be its curvature tensor, and  $R_1, R_2$  be curvature tensors for  $\mathcal{M}_1$  and  $\mathcal{M}_2$  respectively, then one can relate  $R$ ,  $R_1$  and  $R_2$  by:*

$$R(X, Y) = R_1(X_1, Y_1) + R_2(X_2, Y_2), \quad (221)$$

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 of  $\star$ .

Based on Lemma 6, we can prove that 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  $\{SO(2) \times \cdots \times SO(2)\}_{n_p}$ . As the (product) Lie group, the curvature tensors of the parameter space  $\mathcal{P}_1$  on the tangent space  $\mathcal{T}_{(\mathbf{x}, \mathbf{R})}\mathcal{P}_1$  is given by a simple formula [13]:

$$\langle R(\bar{\mathbf{X}}, \bar{\Omega})\bar{\Omega}, \bar{\mathbf{X}} \rangle = \frac{1}{4} \|\bar{[\mathbf{X}, \Omega]}\|_F^2, \quad (222)$$

where  $[\bar{\mathbf{X}}, \bar{\Omega}]$  is the Lie bracket of  $\bar{\mathbf{X}}$  and  $\bar{\Omega}$ , two vectors (not necessarily orthonormal) in the tangent space  $\mathcal{T}_{(\mathbf{x}, \mathbf{R})}\mathcal{P}_1$ . Because of the Lie brackets and the bases (139), it is easy to find that these two groups are both flat and their curvature tensors are  $\mathbf{0}$ . So the curvature tensor of the space  $\mathcal{P}_1$  is  $\mathbf{0}$ .

Based on Appendix B in [13], the curvature tensor for CRLB of synchronization on  $\{SO(n) \times \cdots \times SO(n)\}$  is computed by:

$$\begin{aligned} \mathbf{R}_m[\hat{\Omega}, \hat{\Omega}] &\triangleq \mathbb{E} \left\{ \langle R(\mathbf{X}, \hat{\Omega})\hat{\Omega}, \mathbf{X} \rangle \right\} \\ &= \mathbb{E} \left\{ \frac{1}{4} \|\mathbf{[X, \hat{\Omega}]}\|_F^2 \right\} \\ &= \mathbb{E} \left\{ \frac{1}{4} \left\| \sum_{s,l} \alpha_{kl} \beta_{ks} [\mathbf{E}_l, \mathbf{E}_s] \right\|_F^2 \right\} \end{aligned} \quad (223)$$

where  $\hat{\Omega}$ ,  $\mathbf{X}$  are two vectors in tangent space of the parameter space, meeting  $\hat{\Omega} = \sum_{k,l} \alpha_{kl} \mathbf{R}_k \mathbf{E}_l$ ,  $\mathbf{X} = \sum_{k,l} \beta_{kl} \mathbf{R}_k \mathbf{E}_l$ .  $\mathbf{R}_k \mathbf{E}_l$  is the orthonormal basis.

For  $SO(2)$ , its orthonormal basis is  $\mathbf{R}_k \mathbf{E}$ ,  $\mathbf{E} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ . Let's only consider the bracket operation:  $[\mathbf{E}, \mathbf{E}] = \mathbf{E}\mathbf{E} - \mathbf{E}\mathbf{E} = 0$ , so its curvature terms  $\mathbf{R}_m[\Omega, \Omega]$  is 0, which means that it is flat. For the Euclidean space, the result is also equal to 0.

The CRLB for 2D pose-graph SLAM on  $\mathbb{R}^2 \times SO(2)$  is:

$$\mathbf{C} \succeq \mathbf{F}^{-1} + \text{curvature terms} = \mathbf{F}^{-1} + \mathbf{0} = \mathbf{F}^{-1}. \quad (224)$$

So we can see that, for the parameter space  $\mathcal{P}_1$ , its CRLB formula is the same as the classical CRLB result for the flat Euclidean space shown as  $\mathbf{C} \succeq \mathbf{F}^{-1}$ .

## 11.2 CRLB for 3D pose-graph SLAM

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 6, 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  $\mathcal{P}_2$ .

Based on Theorem 4 in [25], CRLB follows:  $\mathbf{C} \geq \mathbf{F}^{-1} - \frac{1}{3}(R_m(\mathbf{F}^{-1})\mathbf{F}^{-1} + \mathbf{F}^{-1}R_m(\mathbf{F}^{-1}))^4$ , 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  $R_m(\mathbf{F}^{-1}) = \frac{1}{4}\text{diag}\{\mathbf{0}, \text{ddiag}(\mathbf{I}^{SO(3)-1})\}$  of the parameter space  $\mathcal{P}_2$  for the anchored rotation synchronization situation has been shown by [13].

Following the same formulation as in [13], we can get the curvature tensor of 3D pose-graph SLAM:

$$\begin{aligned} \mathbf{C} &\succeq \mathbf{F}^{-1} - \frac{1}{12}(\text{ddiag}(\tilde{\mathbf{L}})\mathbf{F}^{-1} + \mathbf{F}^{-1}\text{ddiag}(\tilde{\mathbf{L}})), \\ \tilde{\mathbf{L}} &= \left[ \begin{array}{cc} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}^{SO(3)-1} \end{array} \right]_{6n_p \times 6n_p}. \end{aligned} \quad (225)$$

It is easy to find that when  $snr$  is large enough, compared with  $\mathbf{F}^{-1}$ , the function  $\text{ddiag}(\tilde{\mathbf{L}})\mathbf{F}^{-1} + \mathbf{F}^{-1}\text{ddiag}(\tilde{\mathbf{L}})$  is much smaller. So it is negligible.

## 12 Optimal Experimental Design

We know that the Theory of Optimal ExpIt is known that the TOED [30], including T-/D-optimality, for the FIM of the estimated state vector is a common way to evaluate the results in terms of estimation accuracy. In this section, we will discuss and compare the optimal experimental design metrics of 2D/3D pose-graph SLAM and show the tight bounds of these metrics, which are much easier to compute than the accurate metrics for the FIM.

### 12.1 T-optimality design metric

#### 12.1.1 T-optimality design metric for the synchronization on $\mathbb{R}^2 \times SO(2)$

T-optimality criterion maximizes the trace of the FIM. Based on the 2D FIM (175), we can get its T-optimality design metric:

$$\text{trace}(\mathcal{I}_{2D}) = \text{trace}(\mathbf{L}_w^{\mathbb{R}^2}) + \text{trace}(\mathbf{L}_w^{SO(2)}) + \sum_{i=1}^{n_p} \sum_{j \in V_i^+} \delta_{ij}^{-2} \|\mathbf{x}_j - \mathbf{x}_i\|_2^2. \quad (226)$$

#### 12.1.2 T-optimality design metric for the synchronization on $\mathbb{R}^3 \times SO(3)$

Based on the 3D FIM (217), we can get the T-optimality design metric of the FIM:

$$\text{trace}(\mathcal{I}_{3D}) = \text{trace}(\mathbf{L}_w^{\mathbb{R}^3}) + \text{trace}(\mathbf{L}_w^{SO(3)}) + \sum_{i=1}^{n_p} \text{trace}(\Psi_i). \quad (227)$$

We can find that, in these three parts, only one part  $\sum_{i=1}^{n_p} \text{trace}(\Psi_i)$  includes the state vector  $\mathbf{x}_i$ ,  $\mathbf{x}_j$ ,  $\mathbf{R}_i$ . In fact, because of the special structure about  $\Psi_i$ , we have:  $\sum_{i=1}^{n_p} \text{trace}(\Psi_i) = \sum_{i=1}^{n_p} \sum_{j \in V_i^+} \delta_{ij}^{-2} \|\mathbf{x}_j - \mathbf{x}_i\|_2^2$ .

Based on the definition (214),  $\Psi_i$  can be written as:

$$\begin{aligned} \Psi_i &= \sum_{j \in V_i^+} \Psi_{(i,j)}, \quad \Psi_{(i,j)} = \begin{bmatrix} \psi_{(i,j)}^{11} & \psi_{(i,j)}^{12} & \psi_{(i,j)}^{13} \\ \psi_{(i,j)}^{12} & \psi_{(i,j)}^{22} & \psi_{(i,j)}^{21} \\ \psi_{(i,j)}^{13} & \psi_{(i,j)}^{21} & \psi_{(i,j)}^{33} \end{bmatrix}, \\ \psi_{(i,j)}^{kl} &= \delta_{ij}^{-2} (\mathbf{x}_i - \mathbf{x}_j)^\top \mathbf{R}_i \mathbf{I}_{3 \times 3}^{k,l} \mathbf{R}_i^\top (\mathbf{x}_i - \mathbf{x}_j). \end{aligned} \quad (228)$$

<sup>4</sup>It is noted that the original conclusion using the Moore–Penrose pseudoinverse for the anchored situation based on all poses. Because, in the pose-graph SLAM, we commonly delete the anchored pose in the estimated problem. So the Moore–Penrose pseudoinverse will change into the inverse of the reduced matrix of the FIM by deleting the corresponding anchored poses.

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

$$\Psi_{(i,j)} = \delta_{ij}^{-2} \begin{bmatrix} p_2^2 + p_3^2 & -p_1 p_2 & -p_1 p_3 \\ -p_1 p_2 & p_1^2 + p_3^2 & -p_2 p_3 \\ -p_1 p_3 & -p_2 p_3 & p_1^2 + p_2^2 \end{bmatrix}. \quad (229)$$

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

$$\begin{aligned} \lambda^3 + k_1 \lambda^2 + k_2 \lambda + k_3 &= 0, \\ k_1 &= -b_{11} - b_{22} - b_{33}, \\ k_2 &= b_{11}b_{22} + b_{11}b_{33} + b_{22}b_{33} - b_{12}^2 - b_{23}^2 - b_{33}^2, \\ k_3 &= -\det(\mathbf{B}). \end{aligned} \quad (230)$$

We can find that  $\det(\Psi_{(i,j)}) = 0$  easily. So  $k_3 = -\det(\Psi_{(i,j)}) = -\det(\mathbf{B}) = 0$ . Substitute the matrix (229) into the equations (230), we have:

$$\begin{aligned} \lambda^2 + k_1 \lambda + k_2 &= 0, \\ k_1 &= -2(p_1^2 + p_2^2 + p_3^2), \Rightarrow (\lambda - (p_1^2 + p_2^2 + p_3^2))^2 = 0. \\ k_2 &= (p_1^2 + p_2^2 + p_3^2)^2, \end{aligned} \quad (231)$$

Based on  $p_1^2 + p_2^2 + p_3^2 = (\mathbf{x}_i - \mathbf{x}_j)^\top \mathbf{R}_i \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \mathbf{R}_i^\top (\mathbf{x}_i - \mathbf{x}_j) = \frac{1}{2} \|\mathbf{x}_i - \mathbf{x}_j\|_2^2$ , we can known that the eigenvalues of  $\Psi_{(i,j)}$  are  $\lambda_1^{ij} = \lambda_2^{ij} = \frac{1}{2} \delta_{ij}^{-2} \|\mathbf{x}_j - \mathbf{x}_i\|_2^2$  and  $\lambda_3^{ij} = 0$ . So we have  $\sum_{i=1}^{n_p} \text{trace}(\Psi_i) = \sum_{i=1}^{n_p} \sum_{j \in V_i^+} \text{trace}(\Psi_{(i,j)}) = \sum_{i=1}^{n_p} \sum_{j \in V_i^+} (\lambda_1^{ij} + \lambda_2^{ij} + \lambda_3^{ij}) = \sum_{i=1}^{n_p} \sum_{j \in V_i^+} \delta_{ij}^{-2} \|\mathbf{x}_j - \mathbf{x}_i\|_2^2$  and  $\lambda_\infty(\Psi_{(i,j)}) = \frac{1}{2} \delta_{ij}^{-2} \|\mathbf{x}_j - \mathbf{x}_i\|_2^2$ .

### 12.1.3 Further analysis

We know that the measurement function is:  $\mathbf{p}_{ij} = \mathbf{R}_i^\top(\mathbf{x}_j - \mathbf{x}_i) + \mathbf{y}_{ij}$ . In general, we have:

$$\delta_{ij}^{-2} \|\mathbf{x}_j - \mathbf{x}_i\|_2^2 = \delta_{ij}^{-2} \|\mathbf{p}_{ij} - \mathbf{y}_{ij}\|_2^2 \approx \delta_{ij}^{-2} \|\mathbf{p}_{ij}\|_2^2. \quad (232)$$

Introduce (232) into the T-optimality metric (226) and (227), we have:

$$\text{trace}(\mathbf{I}_{nD}) \approx \text{trace}(\mathbf{L}_w^{\mathbb{R}^n}) + \text{trace}(\mathbf{L}_w^{SO(n)}) + \sum_{i=1}^{n_p} \sum_{j \in V_i^+} \delta_{ij}^{-2} \|\mathbf{p}_{ij}\|_2^2, \quad n = 2, 3. \quad (233)$$

Based on this approximation (233), we can see that the trace 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,  $\sum_{i=1}^{n_p} \sum_{j \in V_i^+} \delta_{ij}^{-2} \|\mathbf{p}_{ij}\|_2^2$  or  $\sum_{i=1}^{n_p} \sum_{j \in V_i^+} \delta_{ij}^{-2} \|\mathbf{x}_j - \mathbf{x}_i\|_2^2$  is relatively small. So we have:

$$\text{trace}(\mathbf{I}_{nD}) \rightarrow \text{trace}(\mathbf{L}_w^{\mathbb{R}^n}) + \text{trace}(\mathbf{L}_w^{SO(n)}) = \sum_{i=1}^{n_p} \sum_{j \in V_i} (dw_{i,j}^{SO(n)} + nw_{ij}^{\mathbb{R}}). \quad (234)$$

## 12.2 D-optimality design metric

D-optimality design is to use the log-determinant of the covariance matrix as an objective function. However, the determinant of a high-dimensional dense matrix is really expensive to compute. Due to the sparse structure of the FIM, we always compute  $\log(\det(\mathbf{C}))$  via  $\log(\det(\mathbf{F}))$ :  $\log(\det(\mathbf{C})) \approx \log(\det(\mathbf{F}^{-1})) = -\log(\det(\mathbf{F}))$ . Some references show that the D-optimality metric can keep monotonicity during the exploration [30]. Besides, the D-optimality metric and the entropy have an explicit relationship [31]. D-optimality is a useful metric for quantifying the uncertainty of the estimated robot poses and the generated map in an active SLAM problem. 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 [7]. Because the situation using the isotropic Langevin noise on  $\mathbb{R}^2 \times SO(2)$  is similar to the ones of [7], in this part, we only show the extended result in the group  $\mathbb{R}^3 \times SO(3)$ .

**Theorem 3.** Consider the 3D pose-graph SLAM problem in Section 8, its D-optimality design metric  $\log(\det(\mathcal{I}_{3D}))$  of the FIM has a lower bound  $lb$  and an upper bound  $ub$ :

$$\begin{aligned} lb &\leq \log(\det(\mathcal{I}_{3D})) \leq ub, \\ lb &= \log(\det(\mathbf{L}_w^{\mathbb{R}^3})) + \log(\det(\mathbf{L}_w^{SO(3)})) \\ ub &= \log(\det(\mathbf{L}_w^{\mathbb{R}^3})) + \sum_{i=1}^{3n_p} \log(\lambda_i(\mathbf{L}_w^{SO(3)}) + \lambda_\infty). \end{aligned} \quad (235)$$

where  $\lambda_\infty = \|\text{diag}\{\Psi_1, \dots, \Psi_{n_p}\}\|_{eig}$  and  $\lambda_i(\mathbf{L}_w^{SO(3)})$  means the  $i$ -th eigenvalue of  $\mathbf{L}_w^{SO(3)}$ .

*Proof.* Based on the Schur's determinant formula [7] and (217), because  $\mathbf{L}_w^{\mathbb{R}^3}$  is invertible, we have:

$$\log(\det(\mathcal{I}_{3D})) = \log(\det(\mathbf{L}_w^{\mathbb{R}^3})) + \log(\det(\mathbf{L}_w^{SO(3)} + \text{diag}\{\Psi_1, \dots, \Psi_{n_p}\} - \Delta_w^{3D} \mathbf{L}_w^{\mathbb{R}^3-1} \Delta_w^{3D\top})). \quad (236)$$

Similar to the proof of Theorem 3 in [7], we can show  $\Delta_w^{3D} \mathbf{L}_w^{\mathbb{R}^3-1} \Delta_w^{3D\top} \succeq 0$  and  $\text{diag}\{\Psi_1, \dots, \Psi_{n_p}\} - \Delta_w^{3D} \mathbf{L}_w^{\mathbb{R}^3-1} \Delta_w^{3D\top} \succeq 0$  are orthogonal projection matrices, then we have:

$$\begin{aligned} lb &= \log(\det(\mathbf{L}_w^{\mathbb{R}^3})) + \log(\det(\mathbf{L}_w^{SO(3)})) \leq \log(\det(\mathcal{I}_{3D})) \leq \log(\det(\mathbf{L}_w^{\mathbb{R}^3})) + \log(\det(\mathbf{L}_w^{SO(3)})) \\ &+ \text{diag}\{\Psi_1, \dots, \Psi_{n_p}\}) < ub = \log(\det(\mathbf{L}_w^{\mathbb{R}^3})) + \log(\det(\mathbf{L}_w^{SO(3)} + \lambda_\infty \mathbf{I}_{3n_p \times 3n_p})) \\ &= \log(\det(\mathbf{L}_w^{\mathbb{R}^3})) + \sum_{i=1}^{3n_p} \log(\lambda_i(\mathbf{L}_w^{SO(3)}) + \lambda_\infty). \end{aligned} \quad (237)$$

□

Some discussions about  $\lambda_\infty$  are shown as follows:

**Corollary 1.** The biggest eigenvalue  $\lambda_\infty$  in Theorem 3 is smaller than  $\max_{i=1,2,\dots,n_p} \frac{1}{2} \sum_{j \in V_i^+} \delta_{ij}^{-2} \|\mathbf{x}_j - \mathbf{x}_i\|_2^2$ .

*Proof.* Based on the Rayleigh quotients [43], we know that the variational description of the maximal eigenvalue  $\lambda_\infty(\Psi_i)$  of the real symmetric matrix  $\Psi_i$ ,  $i = 1, \dots, n_p$  is:

$$\begin{aligned} \lambda_\infty(\Psi_i) &= \max_{\mathbf{x} \neq 0} \frac{\mathbf{x}^\top \Psi_i \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} = \max_{\mathbf{x} \neq 0} \frac{\mathbf{x}^\top \sum_{j \in V_i^+} \Psi_{(i,j)} \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} \leq \sum_{j \in V_i^+} \max_{\mathbf{x} \neq 0} \frac{\mathbf{x}^\top \Psi_{(i,j)} \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} \\ &= \sum_{j \in V_i^+} \lambda_\infty(\Psi_{(i,j)}) = \sum_{j \in V_i^+} \frac{1}{2} \delta_{ij}^{-2} \|\mathbf{x}_j - \mathbf{x}_i\|_2^2. \end{aligned} \quad (238)$$

Because all eigenvalues of a block diagonal matrix are the eigenvalues of all block matrices on the diagonal [44], we have:

$$\lambda_\infty = \max_{i=1,2,\dots,n_p} \lambda_\infty(\Psi_i) \leq \max_{i=1,2,\dots,n_p} \frac{1}{2} \sum_{j \in V_i^+} \delta_{ij}^{-2} \|\mathbf{x}_j - \mathbf{x}_i\|_2^2. \quad (239)$$

□

## 12.3 Discussion and comparison

### 12.3.1 Efficiency of the metric

**T-optimality metric** The parameters  $w_{i,j}^{SO(n)}$  and  $w_{i,j}^{\mathbb{R}}$  are constant. To simplify the discussion, we assume that these two parameters are the same in every measurement process for each pair of  $(i, j)$ . Set:  $c^\omega = dw_{i,j}^{SO(n)} + nw_{i,j}^{\mathbb{R}}$ , based on (234), we have:

$$\text{trace}(\mathcal{I}_{nD}) \approx \sum_{i=1}^{n_p} \sum_{j \in V_i} c^\omega = c^\omega \sum_{i=1}^{n_p} |V_i|, \quad n = 2, 3, \quad (240)$$

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

Based on the approximation between the T-optimality metric and node degree (240), we can get a simple conclusion: for fixed number of pose, minimizing the T-optimality metric is close 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 distinguish their uncertainty levels. The two graphs shown in Figure 4 illustrate this.

```
Loading file: data/cubicle.g2o ...
Processed input file data/cubicle.g2o in 6.36707 seconds
Number of poses: 5750
Number of measurements: 16869
```

Figure 4: Two examples of pose-graphs (with  $\mathbf{x}_0$  as anchor)

Because of the same nodes and total node degree, graph 1 and graph 2 have the similar T-optimality objective function based on the equation (240). However, it is easy to know that graph 2 is much better than graph 1 because of the loop closure. The uncertainty of poses  $\mathbf{x}_4$  to  $\mathbf{x}_7$  in graph 1 is larger than the ones in graph 2. So in some situations, the T-optimality metric is not very valuable.

**D-optimality metric** Based on the bounds (237), we have the lower bound ( $lb = \log(\det(\mathbf{L}_w^{\mathbb{R}^n})) + \log(\det(\mathbf{L}_w^{SO(n)}))$ ) and the upper bound ( $ub = \log(\det(\mathbf{L}_w^{\mathbb{R}^n})) + \sum_{i=1}^{dn_p} \log(\lambda_i(\mathbf{L}_w^{SO(n)}) + \lambda_\infty)$ ). For most pose-graph SLAM problems, the relative translations  $\|\mathbf{x}_i - \mathbf{x}_j\|_2^2$  are usually small, so we have  $ub \rightarrow lb$ . Finally, we can get:  $\log(\det(\mathcal{I}_{nD})) \rightarrow lb = \log(\det(\mathbf{L}_w^{\mathbb{R}^n})) + \log(\det(\mathbf{L}_w^{SO(n)}))$ . We can get the following lemma:

**Theorem 4.** Consider the 2D/3D pose-graph SLAM problem in Section 8,  $\xi \triangleq \max_{i=1,2,\dots,n_p} \sum_{j \in V_i^+} \delta_{ij}^{-2} \|\mathbf{x}_j - \mathbf{x}_i\|_2^2$  (2D) or  $\xi \triangleq \max_{i=1,2,\dots,n_p} \frac{1}{2} \sum_{j \in V_i^+} \delta_{ij}^{-2} \|\mathbf{x}_j - \mathbf{x}_i\|_2^2$  (3D) and  $\lambda_{min}(\mathbf{L}_w^{SO(n)})$  is the minimal eigenvalue of  $\mathbf{L}_w^{SO(n)}$ , define  $\varepsilon = \log(\det(\mathcal{I}_{nD})) - \log(\det(\mathbf{L}_w^{\mathbb{R}^n})) - \log(\det(\mathbf{L}_w^{SO(n)}))$ . Then we have,

$$0 \leq \varepsilon \leq dn_p \log(1 + \xi / \lambda_{min}(\mathbf{L}_w^{SO(n)})). \quad (241)$$

*Proof.* Based on the equations (237) and (239), we have:

$$\begin{aligned} \varepsilon &= \log(\det(\mathcal{I}_{nD})) - \log(\det(\mathbf{L}_w^{\mathbb{R}^n})) - \log(\det(\mathbf{L}_w^{SO(n)})) \\ &\leq \log(\det(\mathbf{L}_w^{SO(n)} + \xi \mathbf{I}_{dn_p \times dn_p})) - \log(\det(\mathbf{L}_w^{SO(n)})) \\ &= \log \prod_{i=1}^{dn_p} \frac{\lambda_i(\mathbf{L}_w^{SO(n)}) + \xi}{\lambda_i(\mathbf{L}_w^{SO(n)})} \\ &\leq \log \left(1 + \xi / \lambda_{min}(\mathbf{L}_w^{SO(n)})\right)^{dn_p} \\ &= dn_p \log \left(1 + \xi / \lambda_{min}(\mathbf{L}_w^{SO(n)})\right), \end{aligned} \quad (242)$$

where  $\lambda_i(\mathbf{L}_w^{SO(n)})$ ,  $i = 1, 2, \dots, dn_p$  are the eigenvalues of  $\mathbf{L}_w^{SO(n)}$ ,  $\lambda_{min}(\mathbf{L}_w^{SO(n)})$  is the minimal eigenvalue of  $\mathbf{L}_w^{SO(n)}$ .  $\square$

Based on Kirchhoff's matrix tree theorem [45], it is easy to know that  $\det(\mathbf{L}_w^{\mathbb{R}^n})$  and  $\det(\mathbf{L}_w^{SO(n)})$  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 maximizing the weighted number of spanning trees of the graph  $\mathcal{G}$  (also named weighted tree connectivity).

**Remark 1.** The similar 2D pose-graph SLAM results with the Gaussian noise about the relationship between the D-optimality metric with the weighted tree connectivity have been discussed in [7]. 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 [7]) can be extended to 3D case.

---

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

---

```

1: procedure LB( $\mathbf{L}_{w_{SO(3)}}, \mathbf{L}_{w_{\mathbb{R}}}$ )
2:    $\mathbf{l} \leftarrow COLAMD(\mathbf{L}_{w_{SO(3)}})$                                  $\triangleright$  Column approximate minimum degree permutation
3:    $\mathbf{L}_1 \leftarrow SparseCholesky(\mathbf{L}_{w_{SO(3)}}(\mathbf{l}, \mathbf{l}))$        $\triangleright$  Sparse Cholesky factor based on  $\mathbf{l}$  for  $\mathbf{L}_{w_{SO(3)}}$ 
4:    $\mathbf{L}_2 \leftarrow SparseCholesky(\mathbf{L}_{w_{\mathbb{R}}}(\mathbf{l}, \mathbf{l}))$        $\triangleright$  Sparse Cholesky factor based on the same  $\mathbf{l}$  for  $\mathbf{L}_{w_{\mathbb{R}}}$ 
5:   return  $6 \cdot \sum_i (\log(\mathbf{L}_1)_{i,i} + \log(\mathbf{L}_2)_{i,i})$ 

```

---

### 12.3.2 Computational complexity

**T-optimality metric** After constructing the FIM, the computational complexity of the trace 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 in an incremental method. The complexity can be further reduced to  $O(L_p)$ , where  $L_p$  is the predicted horizon (MPC framework). It is easy to see that the T-optimality metric is also a computational cost-effective metric.

**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  $\mathbf{x}_i$  and  $\mathbf{R}_i$ , 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:  $\log(\det(\mathbf{L}_w^{\mathbb{R}^3}))$  and  $\log(\det(\mathbf{L}_w^{SO(3)}))$ . Using the sparse Cholesky decomposition with a good fill-reducing permutation (Algorithm 1 in [7]), they can be computed much faster than the log-determinant function of the dense matrix  $O(n_p^3)$ ; However, because  $\mathbf{L}_w^{\mathbb{R}^3}$  and  $\mathbf{L}_w^{SO(3)}$  have the same sparse structure, we can simply modify the original Algorithm 1 in [7] to a more efficient new algorithm (Algorithm 1).

In Algorithm 1,  $(\star)_{i,i}$  means the  $i$ -th diagonal element of the matrix  $\star$ . The same order  $\mathbf{l}$  in Algorithm 1 can also be used to compute the upper bound.

The above algorithm is similar to the Algorithm 1 in [7]. 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:  $\log(\det(\mathbf{L}_w^{\mathbb{R}^3}))$ ,  $\lambda_i(\mathbf{L}_w^{SO(3)})$  and  $\lambda_\infty$ . For  $\lambda_i(\mathbf{L}_w^{SO(3)})$ , we can use the Lanczos algorithm [46] and Fast Multi-pole Method [47] for the sparse Hermitian matrix (FIM is a sparse Hermitian matrix). The Lanczos algorithm can help to generate a tridiagonal real symmetric matrix from the matrix  $\mathbf{L}_w^{SO(3)}$  in complexity  $O(d_n m_p n_p)$  or  $O(d_n n_p^2)$  if  $m_p = n_p$  [48], where  $d_n$  is the average number of nonzero elements in a row of the matrix  $\mathbf{L}_w^{SO(3)}$ ,  $m_p$  is the number of iterations in the Lanczos algorithm (as default, let  $m_p = n_p$  [46]). For the generated tridiagonal matrices, the Fast Multi-pole Method computes all eigenvalues in just  $O(n_p \log n_p)$  operations. So the computational complexity of  $\lambda_i(\mathbf{L}_w^{SO(3)})$  is totally  $O(d_n n_p^2) + O(n_p \log n_p)$ ; For  $\lambda_\infty$ , we have shown the analytical results  $\lambda_\infty \approx \max_{i=1,2,\dots,n_p} \frac{1}{2} \sum_{j \in V_i^+} \delta_{ij}^{-2} \|\mathbf{p}_{ij}\|_2^2$ , which has the computation complexity  $O(n_p)$ .

The above discussion is based on the operations on the whole Laplacian matrix. 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 [35]. 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 is one of our future research directions.

## 13 Code and simulation for D-optimality bound

### 13.1 Code

In order to verify the results in this note. We write a matlab code. The main structure of the code is to verify the correctness of the upper and lower bounds of the log-determinant function of the Fisher information matrix (FIM). The structure of the code includes are: (1)Load dataset; (2)Construct FIM; (3)Compute log-determinant function and obtain bounds.

In order to verify the results of the CRLB in this note. We write a matlab code. The main purpose of the code is to verify the CRLB is reachable using David Rosen's SE(3)-sync pose-graph SLAM method (which belongs to maximum likelihood estimator MLE). The structure of the code includes are: (1)Load dataset; (2)Obtain ground truth; (3)Sample noise; (4)Obtain noisy measurement; (5)Obtain results using SE(3)-sync; (6)Repeat step (3)-(5) for many times; (7)Obtain covariance based on statistical way; (2)Construct FIM; (8)Compute CRLB based on FIM; (9)Compare and plot.

### 13.2 Load dataset

Because the results of the note are based on the Langevin distribution, we need to use the dataset of which noise follows this distribution. The way to obtain this noise in  $SO(2)$  and  $SO(3)$  is presented in [4]. We directly use their dataset in their matlab code. The dataset includes:

3D datasets: sphere2500, torus3D, grid3D, parking-garage, cubicle, rim, smallGrid3D, tinyGrid3D;  
 2D datasets: kitti0, kitti2, kitti5, kitti6, kitti7, kitti8, kitti9, CSAIL, manhattan, city10000, intel, ais2klinik;

All these dataset are saved in a fold named 'data'. After running the code, these datasets are stored in an architecture: measurements. It includes five parts: edges (the measurement direction), R (noisy rotation measurements), t (noisy translation measurements), kappa (concentration in Langevin distribution), tau (covariance of the isotropic Gaussian noise). In this part, we also output the number of poses and the number of measurements. The output is:

---

```
Loading file: data/cubicle.g2o ...
Processed input file data/cubicle.g2o in 6.36707 seconds
Number of poses: 5750
Number of measurements: 16869
```

Figure 5: Output of the loading dataset (cubicle)

Because in these process, we need to compute the inverse of the FIM, we only choose two small datasets (3D datasets: smallGrid3D and 2D datasets: intel). We directly use their dataset in their matlab code.

All these dataset are saved in a fold named 'data'. After running the code, these datasets are stored in an architecture: measurements. It includes five parts: edges (the measurement direction), R (noisy rotation measurements), t (noisy translation measurements), kappa (concentration in Langevin distribution), tau (covariance of the isotropic Gaussian noise). In this part, we also output the number of poses and the number of measurements.

### 13.3 Construct FIM

Based on the results shown in (165), (166), (240), (206) and (209), we use the sparse function to construct the FIM efficiently. It is divided into 2D FIM part and 3D FIM part. After the constructing process, it will visualize the sparsity pattern of the FIM by spy function. We also output the constructing time. The corresponding matlab file is: construct\_FIM.m. We build the whole FIM,  $\mathbb{R}^2$  and  $\mathbb{R}^3$  parts (A Laplacian matrix),  $SO(2)$  and  $SO(3)$  parts, the Laplacian matrix parts in  $SO(2)$  and  $SO(3)$  parts, two parts presenting the coupling information matrix with rotation and translation part, and the added non-laplacian part in  $SO(2)$  and  $SO(3)$  parts. They are saved and named as: I, LR2, LSO2, LSO2\_L, M\_R2\_SO, M\_SO\_R2 and LD. The simulation output is:

```
Try to construct Fisher Information matrix (FIM): data/cubicle.g2o ...
The FIM of the dataset data/cubicle.g2o is constructed using 7.49096 seconds
```

Figure 6: Output of the constructing FIM (cubicle)

### 13.4 Compute log-determinant function and Obtain bounds

After obtaining the FIM, we need to compute the objective function and its lower and upper bounds based on . For the objective function, we use Algorithm 1 shown in [?]. The corresponding matlab file is shown in logdet.m. For the upper and lower bound beside the file logdet.m, we also use the eigs.m, which is the library function, to compute the eigenvalues of LSO2\_L and LD. The bound and the objective function

are obtained in f\_det\_bound.m. Because of the special structure of the part  $D(SO(3))$  includes the state vector, the upper bound  $\log(\det(\mathbf{L}_w^{\mathbb{R}^3})) + \log(\det(\mathbf{L}_w^{SO(3)} + \text{diag}\{\Psi_1, \dots, \Psi_{n_p}\}))$  is changing with the state vector. We use a random state vector to compute the FIM and this upper bound. For the constant upper bound  $\log(\det(\mathbf{L}_w^{\mathbb{R}^3})) + \sum_{i=1}^{3n_p} \log\left(\lambda_i(\mathbf{L}_w^{SO(3)}) + \max_{i=1,2,\dots,n_p} \frac{1}{2} \sum_{j \in V_i^+} \delta_{ij}^{-2} \|\mathbf{p}_{ij}\|^2\right)$ , we can get the unchanged bound. The objective function, changing upper bound, lower bound and constant upper bound are respectively  $v_I$ , Upper\_bound, Lower\_bound and Upper\_bound\_new. The Simulation outputs are:

```
Compute the bound of the FIM for data/smallGrid3D.g2o ...
The bound of the determinant of the FIM for the dataset data/smallGrid3D.g2o is obtained costing 0.298972 seconds
The determinant value is: 4.490084e+03
The changing upper bound is: 4.545123e+03
The constant upper bound is: 4.497863e+03
The lower bound is: 4.381454e+03
```

Figure 7: Output of the constructing FIM (smallGrid3D)

### 13.5 Obtain ground truth

In order to compute the covariance, we need a ground truth and then sample noise to it. The ground truth is obtained by the optimization results of original datasets present in [?] using SE(3)-sync. Then we can get the estimated poses. We set these estimated poses as the ground truth.

### 13.6 Sample noise

After obtaining the ground truth, we need to generate the random noises follow the isotropic Gaussian distribution and isotropic Langevin distribution. We use 'normrnd' MATLAB function to generate the isotropic Gaussian distribution. For the isotropic Langevin distribution, the noises are generated using the Acceptance-Rejection Method (ARM) [54].

### 13.7 Obtain noisy measurement

Then we can add these noises into our relative measurements by the edge data using following equations:

$$\begin{aligned} \mathbf{H}_{ij} &= \mathbf{R}_i^\top \mathbf{R}_j \mathbf{Z}_{ij}, \mathbf{Z}_{ij} \sim \text{Lang}(\mathbf{I}_{n \times n}, k_c \cdot \kappa) \\ \mathbf{p}_{ij} &= \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i) + \mathbf{y}_{ij}, \mathbf{y}_{ij} \sim \mathcal{N}(\mathbf{0}, \delta \mathbf{I}_{n \times n}) \end{aligned} \quad (243)$$

These noisy measurement and covariance values are saved as R, t, tau and kappa, which construct a structure 'measurement'. We can generate many sets of 'measurement' dataset.

### 13.8 Obtain results using SE(3)-sync

For every dataset, we can use the SE(3)-sync to finish the pose-graph SLAM and obtain its estimating results. The results are the estimated pose for this measurement. We can obtain all estimated poses.

### 13.9 Obtain covariance based on statistical way

After obtaining the estimated result for every measurement, we can compare it with the ground truth and compute the covariance for every pose. For a estimated result, we can compute the trace function of covariance by:

$$\text{trace}(\mathbf{C}) = \mathbb{E}\left\{\sum_{i=1}^{N_p} \left( \text{dist}(\mathbf{R}_i, \widehat{\mathbf{R}}_i)^2 + \|\mathbf{x}_i - \widehat{\mathbf{x}}_i\|_2^2 \right) \right\} \quad (244)$$

where  $\text{dist}(\star, \cdot) = \|\log(\star^\top \cdot)\|_F$ ,  $\widehat{\star}$  means the ground truth of  $\star$ . Every measurement dataset can generate a value using (244), then repeat many times to obtain the exception by average.

### 13.10 Compute CRLB based on FIM

Based on the constructed FIM, we can compute the trace function of its inverse and its curvature terms by (199). Then compared with the obtained statistical covariance.

### 13.11 Simulation setting

In this section, we verify the correctness of CRLB. These experiments are based on 'small-3Dgrid' and 'intel' problems drawn from the pose-graph SLAM. The datasets and the estimation results in these simulation results are obtained based on the open code of the literature [?].

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.

In this section, we validate the correctness of our results 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 pose-graph SLAM. The datasets and the estimation results in these simulations are obtained based on the open sources of the literature [4].

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.

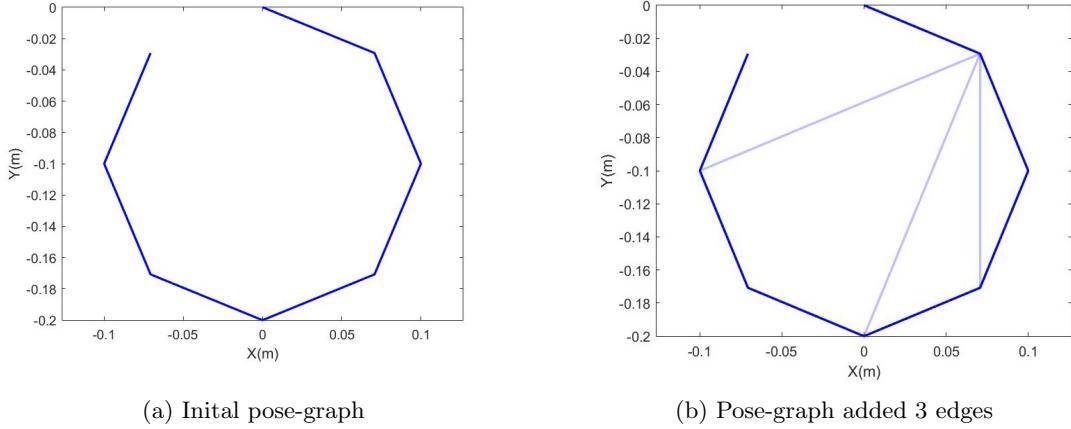


Figure 8: Pose-graph example

### 13.12 Relation between T-/D-optimality metrics with graph topology

In this experiment, we verify the results in Section 12.1.3 and 12.3.1, which show that the T-optimality metric has a strong and direct relationship with the total node degree of the pose-graph as well as the D-optimality metric has a better performance than the T-optimality metric. We use a small example with eight nodes to demonstrate this. A regular octagon is constructed based on these eight poses. The variances of the translation  $\delta_{ij}^2$  and the concentrations  $\kappa_{ij}$  of all measurements are set as  $1 \times 10^{-4}$  and  $5 \times 10^3$ , respectively. The initial pose graph only includes the odometer without the additional measurements, as shown in Fig. 8a. Then, we begin to add the measurements gradually from 1 edge to 21 edges without repetition. For example in Fig. 8b, we add 3 additional measurements. Then, for every graph, we can evaluate the trace and the log-determinant of the FIM after using SE-sync [4] to solve the SLAM problem. The results (black points) and their corresponding total node degree (red line) are shown in Fig. 9 and 10. Each black point means the optimality metric corresponding to a pose-graph with the same 8 poses and unique additional measurements. For example, if the measurement number is 10, which means that we need to add 3 new measurements to the original graph shown in Fig. 8a and Fig. 8b. Based on combinatorics, we can generate all possible pose graphs with different measurements but the same poses, and then evaluate every pose graph using the optimality metric.

In Fig. 9, we can find that the T-optimality metric is almost proportional to the total node degree of the whole pose graph, because the term  $\sum_{i=1}^{n_p} \sum_{j \in V_i^+} \delta_{ij}^{-2} \|p_{ij}\|^2$  is relatively small. The T-optimality metric values of the graphs, which have the same node degree, are similar. Directly using T-optimality metric may lead the obvious local minimal problem in its applications, such as the complex active SLAM problem. In Fig. 10, we can see 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 all possible graphs with  $m = 12$  and  $m = 13$  measurements. It can be seen that the D-optimality metric values with

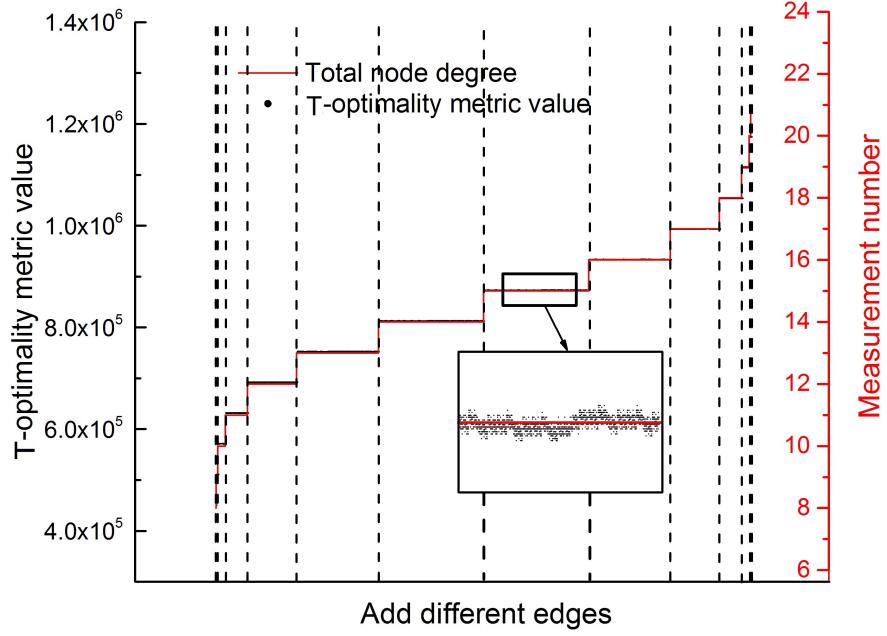


Figure 9: Direct relationship between T-optimality metric with total node degree

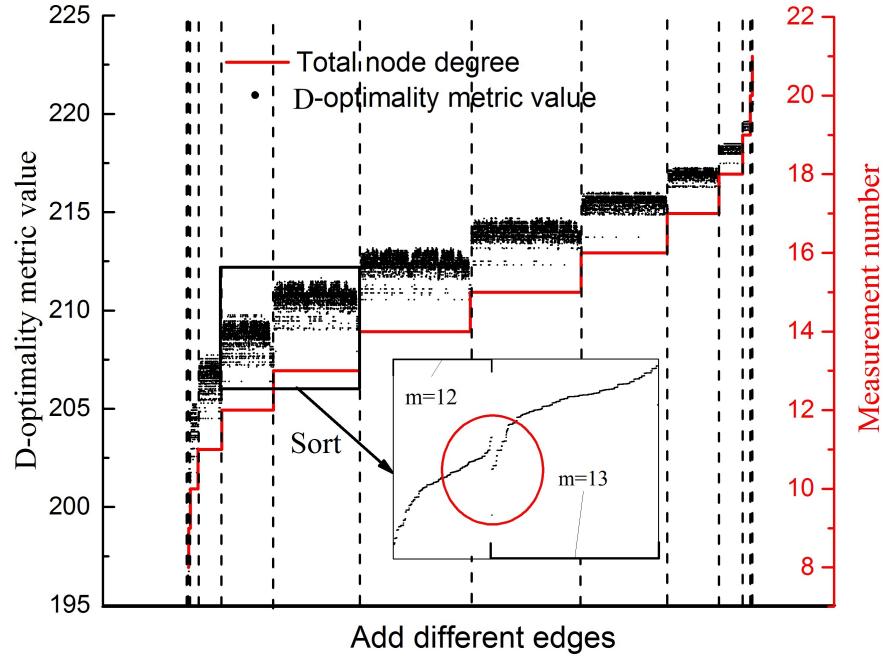


Figure 10: Part relationship between D-optimality metric with total node degree

13 measurements are smaller than the ones with 12 measurements (pay attention to the red circle part). This situation is common for many real pose-graph SLAM problems. Even with fewer measurements, the SLAM problem with better network structure can have a better D-optimality metric, whose physical meaning is the volume of the covariance uncertainty.

In this part, we only discuss the relationship between the node degree and the T-/D- optimality metrics. Their relationship with the weighted number of the spanning tree, which is equal to the lower bound of the D-optimality, will be presented in the following section.

### 13.13 T-/D-optimality metrics in active SLAM application

In this simulation, in order to further explore the applied efficiency of the T-/D-optimality metrics, we compare them in the real active SLAM task. An unmanned aerial vehicle (UAV) is used to pass several pre-defined way-points and perform the SLAM task. The features will be detected when they locate in the sensor range of the UAV. In the moving process, we can get the relative measurements based on the common features between two poses. By this way, the pose graph is built as shown in Fig. 11. Finally, we apply the T-/D-optimality metrics to evaluate 20 candidate paths with the same number of the additional poses and select the optimal one. Based on 100 random simulations with different designed way-points, we find that the optimal one selected by the T-optimality metric is commonly same as the one picked out by the D-optimality metric (77%). Fig. 11 shows a result of which black and blue lines are respectively selected by the D-optimality metric and the T-optimality metric. The black path will lead more loops than the blue path, which leads to the smaller covariance value of the poses (black line:  $1.2497 \times 10^{-3}$  m, blue line:  $1.2701 \times 10^{-3}$  m). However, we also find that, when we pick out top 5 paths which have the big T-optimality metrics in every simulation as a set  $\Xi$ , the optimal one selected by the D-optimality metric will belong to  $\Xi$  with a very high probability (100%, 100 simulations). The above results show that these two metrics have a strong positive correlation, but still have some differences. The D-optimality metric is better than the T-optimality metric.

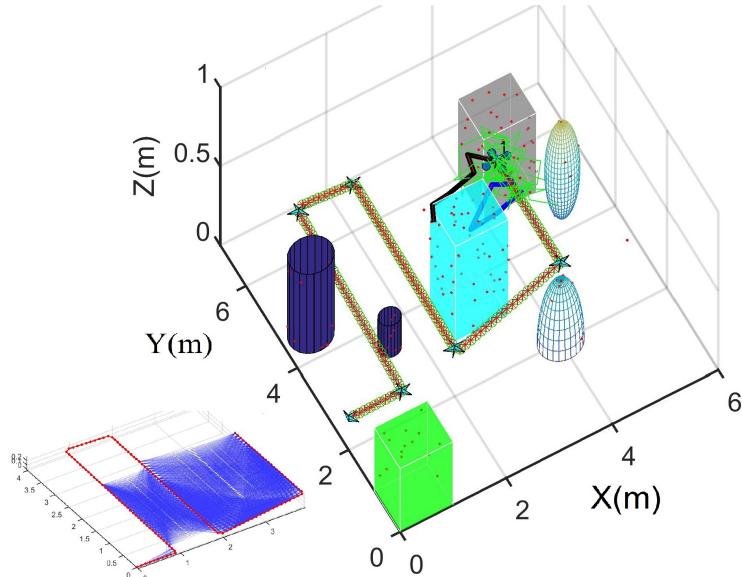


Figure 11: Active SLAM task using two metrics. The green circles, the red stars and the red points are respectively the real UAV trajectory, the estimated trajectory and the detected features. The green lines mean 20 randomly candidate paths. The black and blue lines are the random path selected by the D-/T-optimality metrics. The left figure is the pose graph with the red nodes and the blue edges.

Even though, sometimes, the T-optimality metric has the poorer performance than the D-optimality metric, it can be computed very fast. We can only use the diagonal elements of the FIM to compute the T-optimality metric. For the D-optimality metric, we need to construct the whole FIM and then compute the log-determinant function, which will obviously introduce the high computational complexity. We respectively cost 5.3934s and 0.4398s to compute the D-optimality metric and the T-optimality metric of 20 candidate paths with 170 poses and about 5206 measurements. We can find that the main advantage of the T-optimality metric is the better computational efficiency. So in our current work [52], we use the T-optimality metric or the weighted node degree to deal with the large-scale search for rough candidate actions, and the D-optimality or its lower bound is applied for sophisticated search within a small elite group.

### 13.14 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 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

Table 1: D-optimality 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	$1.9860 \times 10^4$	$2.3284 \times 10^{-3}$	$1.9876 \times 10^4$	$9.2391 \times 10^{-4}$	$1.9858 \times 10^4$	$3.6712 \times 10^{-4}$
Intel	1728	2512	$3.0194 \times 10^4$	$6.0744 \times 10^{-3}$	$3.0352 \times 10^4$	$4.0055 \times 10^{-3}$	$3.0155 \times 10^4$	$2.0743 \times 10^{-3}$
manhattan	3500	5453	$7.1124 \times 10^4$	$3.3530 \times 10^{-1}$	$7.1193 \times 10^4$	$7.3602 \times 10^{-3}$	$7.1104 \times 10^4$	$4.5105 \times 10^{-3}$
KITTI	4541	4677	$1.1697 \times 10^5$	$8.9039 \times 10^{-3}$	$1.1703 \times 10^5$	$4.6468 \times 10^{-3}$	$1.1695 \times 10^5$	$3.1092 \times 10^{-3}$
city10000	10000	20687	$1.6667 \times 10^5$	$9.0052 \times 10^{-2}$	$1.6816 \times 10^5$	$2.4292 \times 10^{-2}$	$1.6520 \times 10^5$	$2.0664 \times 10^{-2}$
ais2klinik	15115	16727	$2.3642 \times 10^5$	$3.6301 \times 10^{-2}$	$2.3975 \times 10^5$	$1.4693 \times 10^{-2}$	$2.3567 \times 10^5$	$1.3116 \times 10^{-2}$

Table 2: D-optimality 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	$2.4562 \times 10^2$	$2.1558 \times 10^{-4}$	$2.7561 \times 10^2$	$2.9866 \times 10^{-2}$	$2.4163 \times 10^2$	$1.4115 \times 10^{-4}$
small-grid	125	297	$4.4452 \times 10^3$	$3.6373 \times 10^{-3}$	$4.8488 \times 10^3$	$3.1836 \times 10^{-2}$	$4.3815 \times 10^3$	$4.3020 \times 10^{-4}$
garage	1661	6275	$2.0618 \times 10^4$	$3.9511 \times 10^{-2}$	$3.6618 \times 10^4$	$3.6117 \times 10^{-2}$	$1.5845 \times 10^4$	$2.2989 \times 10^{-3}$
sphere	2500	4949	$9.9105 \times 10^4$	$7.4856 \times 10^{-2}$	$1.0607 \times 10^5$	$7.4023 \times 10^{-2}$	$9.9054 \times 10^4$	$6.7317 \times 10^{-3}$
torus	5000	9048	$2.4986 \times 10^5$	$1.9488 \times 10^{-1}$	$2.6526 \times 10^5$	$1.0156 \times 10^{-1}$	$2.4985 \times 10^5$	$1.6390 \times 10^{-2}$
cubicle	5750	16869	$2.3729 \times 10^5$	$2.1579 \times 10^{-1}$	$3.1839 \times 10^5$	$5.2156 \times 10^{-2}$	$2.3685 \times 10^5$	$1.4141 \times 10^{-2}$
grid	8000	22236	$4.2613 \times 10^5$	$3.2318 \times 10^0$	$4.4902 \times 10^5$	$1.5065 \times 10^{-1}$	$4.2610 \times 10^5$	$9.0280 \times 10^{-2}$
rim	10195	29743	$4.7257 \times 10^5$	$4.1626 \times 10^{-1}$	$5.8622 \times 10^5$	$6.5298 \times 10^{-2}$	$4.7209 \times 10^5$	$4.1260 \times 10^{-2}$

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 1 (2D) and Table 2 (3D). We demonstrate the number of the poses and the measurements, the log-determinant of the FIM, its upper and lower bounds, and their computational time<sup>5</sup>.

On each of these examples, the log-determinant of the FIM is bounded within  $lb$  and  $ub$  correctly. At the same time,  $lb$  and  $ub$  are very close to original metric in many datasets, especially for the lower bound  $lb$ . We can also find that, benefiting from the great dimensionality reduction, the computation of the upper and lower bounds of the log-determinant of the FIM is much cheaper than the ones of the original D-optimality metric. 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<sup>6</sup>, the log-determinant of the whole FIM gets closed to its lower bound (in 103%), because of the small term  $\sum_{i=1}^{n_p} \sum_{j \in V_i^+} \delta_{ij}^{-2} \|\mathbf{x}_j - \mathbf{x}_i\|_2^2$ . So in the real applications of the D-optimality metric, such as the active pose-graph SLAM [53], we suggest using the tight lower bound, whose physical meaning is the sum of the weighted number of the spanning tree of two graphs, 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  $SO(3)$  parts. They are used to compute the metrics. Table 3 and Table 4 show the computational time. The results show that the sparser weighted Laplacian matrices for the lower bound are much easier to be constructed than the whole FIM.

<sup>5</sup>It is noted that these results do not include the computational time of constructing process of the FIM and Laplacian matrix.

<sup>6</sup>The ‘garage’ dataset has a poor measurement network, whose weighted values of measurement edges are small (large covariance), and the translation values between different poses are large. Both of these two properties make that the term  $\sum_{i=1}^{n_p} \sum_{j \in V_i^+} \delta_{ij}^{-2} \|\mathbf{x}_j - \mathbf{x}_i\|_2^2$  is relatively large compared with the term  $\log(\det(\mathbf{L}_w^{\mathbb{R}^n})) + \log(\det(\mathbf{L}_w^{SO(n)}))$ . This situation is not very common in the real world datasets.

Table 3: Matrix constructing time for the 2D SLAM datasets

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

Table 4: Matrix constructing time for the 3D SLAM datasets

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

### 13.15 Efficiency of CRLB

In this section, our main purpose is to validate that the CRLB is reachable using SE-sync method (which belongs to maximum likelihood estimator) [4]. 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 presented in [4] 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 Langevin distribution are generated. We use ‘normrnd’ MATLAB function to generate the isotropic Gaussian distribution. For the isotropic Langevin distribution, the noises are generated by the Acceptance-Rejection Method (ARM) [54]. Then we can add these noises into our relative measurements by the edge data using following equations:

$$\begin{aligned} \mathbf{H}_{ij} &= \mathbf{Z}_{ij} \mathbf{R}_j \mathbf{R}_i^\top, \mathbf{Z}_{ij} \sim \text{Lang}(\mathbf{I}_{n \times n}, k_c \cdot \kappa_{ij}) \\ \mathbf{p}_{ij} &= \mathbf{R}_i^\top (\mathbf{x}_j - \mathbf{x}_i) + \mathbf{y}_{ij}, \mathbf{y}_{ij} \sim \mathcal{N}(\mathbf{0}, \delta_{ij}^2 \mathbf{I}_{n \times n}), \end{aligned} \quad (245)$$

where  $k_c$  is the coefficient to determine the uncertainty level of the rotation measurement.  $k_c \cdot \kappa_{ij}$  is the concentration of the noise<sup>7</sup>.

For the estimated results, we can compute the trace of covariance by:

$$\text{trace}(\mathbf{C}) = \mathbb{E}\left\{\sum_{i=1}^{n_p} (\text{dist}(\mathbf{R}_i, \bar{\mathbf{R}}_i)^2 + \|\mathbf{x}_i - \bar{\mathbf{x}}_i\|_2^2)\right\}, \quad (246)$$

where  $\bar{\star}$  means the ground truth of  $\star$ . Every measurement dataset can generate a value using (246), then repeat many times to obtain the mathematical expectation  $\mathbb{E}\{\text{trace}(\mathbf{C})\}$  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  $\kappa_{ij}$  and  $\delta_{ij}^{-2}$  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 results are shown in Fig. 12.

For the 3D situation, we use the ‘tiny-3Dgrid’ dataset to obtain the covariance, CRLB with curvature terms and CRLB without curvature terms.  $\kappa_{ij}$  is set as 12.5 and 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. 13.

<sup>7</sup>It is noted that we only consider the uncertainty level of the rotation measurement in our SNR definition (220).

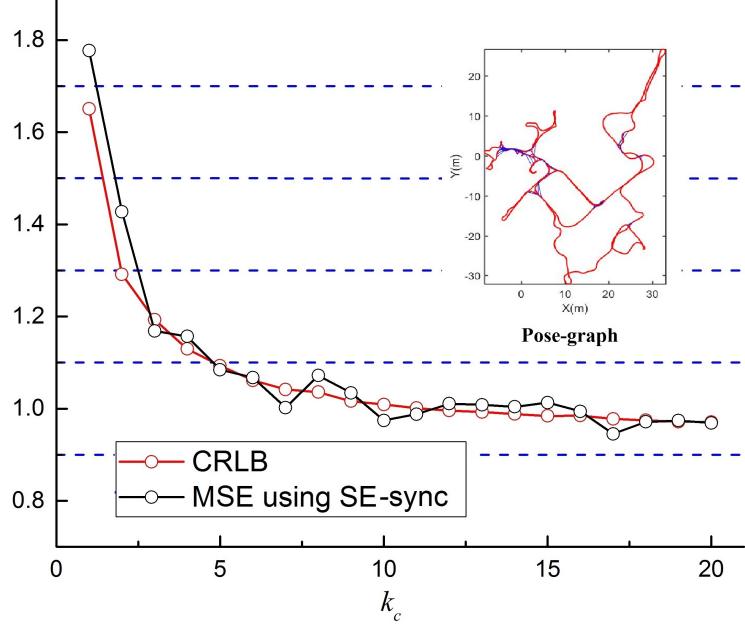


Figure 12: Comparison of the CRLB with the mean squared error ( $MSE = \frac{\text{trace}(C)}{n_p}$ ) of known estimators for the synchronization of  $n_p = 1045$  poses with  $m = 1172$  measurements and one anchor based on the ‘CSAIL’ dataset. The MSE using the SE-sync method appears to reach the CRLB.

The above two simulations show two key points. The first one is that the CRLB is reachable when the SNR is large enough. The other one is that, compared with the  $\mathbf{F}^{-1}$ , the curvature tensor of the parameter space  $\mathcal{P}_3$  is negligible.

## Acknowledgment

We would like to thank Dr. N. Boumal, who has answered many of our queries about his research [2] and [13], which helps us significantly in this research. We want to thank Dr. David M. Rosen for sharing us the code to generate the random noise for our simulation. We would like to thank Dr. K. Khosoussi, Dr. E. Zhang, Dr. J. Wang and Mr. F. Bai for the many productive conversations and input over the course of this work. Finally, we would like to thank the anonymous reviewers of the TRO submission for their constructive feedback, which contributed to Section 6.3 as well as improving the overall quality of the paper.

## References

- [1] Yongbo Chen, Shoudong Huang, Liang Zhao, and Gamini Dissanayake, Cramér-Rao bounds and optimal design metrics for pose-graph SLAM, *IEEE Transactions on Robotics*, submitted.
- [2] N. Boumal, “On intrinsic Cramér-Rao bounds for Riemannian submanifolds and quotient manifolds,” *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1809-1821, Jan. 2013.
- [3] 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 Trans. on Robot.*, vol. 32, no. 6, pp. 1309-1332, Dec. 2016.
- [4] 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,” *Int. J. of Robot. Res.*, vol. 38, no. 2-3, pp. 95-125, Mar. 2018.
- [5] J. Briales, J. Gonzalez-Jimenez, “Cartan-Sync: Fast and global SE(d)-synchronization,” *IEEE Robot. and Autom. Lett.*, vol. 2, no. 4, pp. 2127-2134, Jun. 2017.

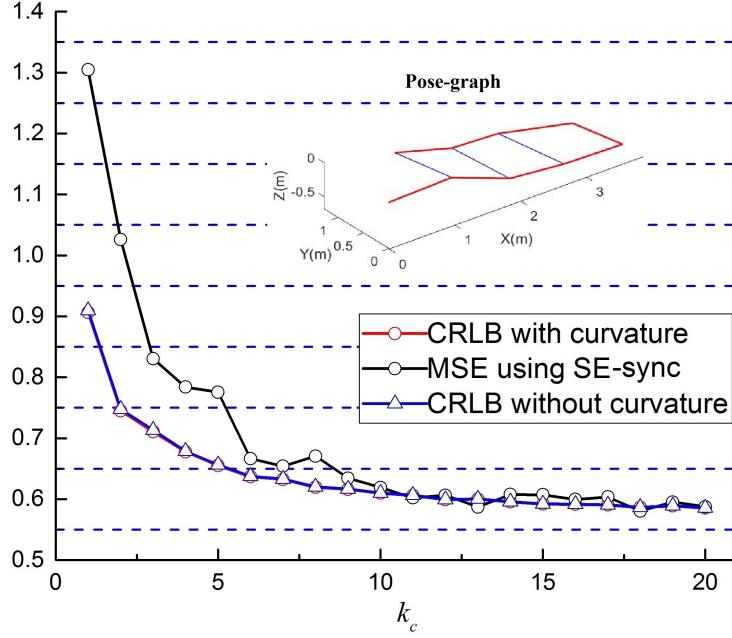


Figure 13: Comparison of the CRLB with the mean squared error ( $MSE = \frac{\text{trace}(\mathbf{C})}{n_p}$ ) of known estimators for the synchronization of  $n_p = 9$  poses with  $m = 11$  measurements and one anchor based on the ‘tiny-3Dgrid’ dataset. The MSE using the SE-sync method appears to reach the CRLB.

- [6] K. Khosoussi, S. Huang, and G. Dissanayake, “Tree-Connectivity: Evaluating the Graphical Structure of SLAM,” in *Proc. of IEEE Int. Conf. on Robot. and Autom. (ICRA)*, May 2016, pp. 1316-1322.
- [7] K. Khosoussi, M. Giamou, G. S. Sukhatme, S. Huang, G. Dissanayake, and J. P. How, “Reliable graph for SLAM,” *Int. J. of Robot. Res.*, vol. 38, no. 2-3, pp. 260-298, Jan. 2019.
- [8] H. Cramér, (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 *Proc. IEEE Conf. on Comp. Vision and Pattern Recognition (CVPR)*, Jun. 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 Trans. on Signal Processing*, vol. 62, no. 2, pp. 363-375, Nov. 2014.
- [11] S. Liu, S. P. Chepuri, M. Fardad, E. Maşazade, G. Leus, and P. K. Varshney, “Sensor selection for estimation with correlated measurement noise,” *IEEE Trans. on Signal Processing*, vol. 64, no. 13, pp. 3509-3522, Apr. 2016.
- [12] R. A. Bailey and P. J. Cameron, “Combinatorics of optimal designs,” *Surveys in Combinatorics*, vol. 365, pp. 19-73, 2009.
- [13] N. Boumal, A. Singer, P. A. Absil, and V. D. Blondel, “Cramér–Rao bounds for synchronization of rotations,” *Information and Inference: A J. of the IMA*, vol. 3, no. 1, pp. 1-39, Mar. 2013.
- [14] R. Kümmel, G. Grisetti, H. Strasdat, K. Konolige, and W. Burgard. “G2o: A General Framework for Graph Optimization,” in *Proc. of IEEE Int. Conf. on Robot. and Autom. (ICRA)*, May 2011, pp. 3607-3613.
- [15] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. J. Leonard, and F. Dellaert. “iSAM2: Incremental smoothing and mapping using the Bayes tree,” *Int. J. of Robot. Res.*, vol. 31, no. 2, pp. 216-235, Feb. 2012.
- [16] V. Ila, L. Polok, M. Solony, and P. Svoboda. “SLAM++-A highly efficient and temporally scalable incremental SLAM framework,” *Int. J. of Robot. Res.*, vol. 36, no. 2, pp. 210-230, Feb. 2017.

- [17] S. Agarwal and K. Mierle and Others, Ceres Solver, <http://ceres-solver.org>.
- [18] L. Carlone, G. C. Calafio, C. Tommolillo, and F. Dellaert. “Planar pose graph optimization: Duality, optimal solutions, and verification,” *IEEE Trans. on Robot.*, vol. 32, no. 3, pp. 545-565, May 2016.
- [19] O. Ozyesil, N. Sharon, and A. Singer. “Synchronization over Cartan motion groups via contraction,” *SIAM J. on Applied Algebra and Geometry*, vol. 2, no. 2, pp. 207-241, Apr. 2018.
- [20] L. Carlone, and G. C. Calafio. “Convex relaxations for pose graph optimization with outliers,” *IEEE Robot. and Autom. Lett.*, Vol. 3, no. 2, pp. 1160-1167, Jan. 2018.
- [21] J. Briales, and J. Gonzalez-Jimenez, “Convex Global 3D Registration with Lagrangian Duality,” in *Proc. IEEE Conf. on Comp. Vision and Pattern Recognition (CVPR)*, Jul. 2017, pp. 4960-4969.
- [22] O. Ozyesil, and A. Singer, “Robust Camera Location Estimation by Convex Programming,” in *Proc. IEEE Conf. on Comp. Vision and Pattern Recognition (CVPR)*, Jun. 2015, pp. 2674-2683.
- [23] S. M. Ross, (2014). Introduction to probability models. Academic press.
- [24] R. C. Rao, “Information and accuracy attainable in the estimation of statistical parameters,” *Breakthroughs in statistics*, vol. 37, pp. 81-91, 1945.
- [25] S. Smith, “Covariance, subspace, and intrinsic Cramér-Rao bounds,” *IEEE Trans. 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 Cramér-Rao bound to Riemannian manifolds,” in *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2005, pp. 1033-1036.
- [27] P. Vakili, H. Mirzaei, S. Zarbaftian, I. C. Paschalidis, D. Kozakov, and S. Vajda, “Optimization on the Space of Rigid and Flexible Motions: an Alternative Manifold Optimization Approach,” in *Proc. IEEE Annual Conf. on Decision and Control (CDC)*, Dec. 2014, pp. 5825-5830.
- [28] O. Brüls, M. Arnold and A. Cardona, “Two Lie Group Formulations for Dynamic Multibody Systems with Large Rotations,” in *Proc. ASME Int. Design Eng. Technical Conf. and Comp. and Inform. in Eng. Conf.*, Jan. 2011, pp. 85-94.
- [29] L. Carlone, R. Tron, K. Daniilidis, and F. Dellaert, “Initialization Techniques for 3D SLAM: a Survey on Rotation Estimation and its use in Pose Graph Optimization,” in *Proc. of IEEE Int. Conf. on Robot. and Autom. (ICRA)*, May 2015, pp. 4597-4604.
- [30] H. Carrillo, Y. Latif, M. L. Rodriguez, J. Neira, and J. A. Castellanos, “On the Monotonicity of Optimality Criteria during Exploration in Active SLAM,” in *Proc. of IEEE Int. Conf. on Robot. and Autom. (ICRA)*, May 2015, pp. 1476-1483.
- [31] H. Carrillo, I. Reid, and J. A. Castellanos, “On the Comparison of Uncertainty Criteria for Active SLAM,” in *Proc. of IEEE Int. Conf. on Robot. and Autom. (ICRA)*, May 2012, pp. 2080-2087.
- [32] Y. Kim and A. Kim, “On the uncertainty propagation: Why Uncertainty on Lie Groups Preserves Monotonicity?” in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2017, pp. 3425-3432.
- [33] M. L. Rodríguez-Arévalo, J. Néira, and J. A. Castellanos. “On the importance of uncertainty representation in active SLAM,” *IEEE Trans. on Robot.*, vol. 34, no. 3, pp. 829-834, Apr. 2018.
- [34] T. D. Barfoot, and P. T. Furgale. “Associating uncertainty with three-dimensional poses for use in estimation problems,” *IEEE Trans. on Robot.*, vol. 30, no.3, pp. 679-693, Jan. 2014.
- [35] 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,” *Int. J. of Robot. Res.*, vol. 36, no. 10, pp. 1088-1130, Sep. 2017.
- [36] K. Khosoussi, S. Huang, and G. Dissanayake, “Novel Insights into the Impact of Graph Structure on SLAM,” in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2014, pp. 2707-2714.

- [37] Wolfram (2001) Modified Bessel function of the first kind: Integral representations. <http://functions.wolfram.com/03.02.07.0007.01>.
- [38] J.C. Simo and L. Vu-Quoc, “On the dynamics in space of rods undergoing large motions—a geometrically exact approach,” *Comput. Methods in Appl. Mech. and Eng.*, vol. 66, no. 2, pp. 125-161, Feb. 1988.
- [39] Bar-Shalom, Yaakov, X. Rong Li, and Thiagalingam Kirubarajan. Estimation with applications to tracking and navigation: theory algorithms and software. John Wiley & Sons, 2004.
- [40] P.-A. Absil, R. Mahony, and R. Sepulchre, (2008) Optimization Algorithms on Matrix Manifolds. Princeton, NJ: Princeton University Press.
- [41] E. Eade (2011). Lie groups for 2d and 3d transformations, <http://ethaneade.com/lie.pdf>, revised May. 2017.
- [42] M. P. Do Carmo (1992). Riemannian geometry. Birkhauser.
- [43] J. R. Magnus and H. Neudecker, (1988). Matrix differential calculus with applications in statistics and econometrics. Wiley series in probability and mathematical statistics.
- [44] D. A. Harville, (1997). Matrix algebra from a statistician’s perspective, Springer.
- [45] S. Chaiken, and D. J. Kleitman. “Matrix tree theorems,” *J. of combinatorial theory, Series A*, vol. 24, no. 3, pp. 377-381, May 1978.
- [46] C. Lanczos, “An iteration method for the solution of the eigenvalue problem of linear differential and integral operators,” *J. of Res. of the National Bureau of Standards*, vol. 45, no. 4, pp. 255-282, Oct. 1950.
- [47] 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, May 2013.
- [48] Y. Saad. “Numerical methods for large eigenvalue problems: revised edition,” *Siam*, 2011.
- [49] Y. Chen, S. Huang, R. Fitch, and J. Yu, “Efficient Active SLAM based on Submap Joining, Graph Topology and Convex Optimization,” in *Proc. of IEEE Int. Conf. on Robot. and Autom. (ICRA)*, pp. 5159-5166, May 2018.
- [50] S. Huang, N. M. Kwok, G. Dissanayake, Q. P. Ha, and G. Fang, “Multi-step Look-ahead Trajectory Planning in SLAM: Possibility and necessity,” in *Proc. of IEEE Int. Conf. on Robot. and Autom. (ICRA)*, pp. 1091-1096, May 2005.
- [51] M. Giamou, K. Khosoussi, and J. P. How, “Talk Resource-Efficiently to Me: Optimal Communication Planning for Distributed Loop Closure Detection,” in *Proc. of IEEE Int. Conf. on Robot. and Autom. (ICRA)*, pp. 3841-3848, May 2018.
- [52] Y. Chen, S. Huang, R. Fitch, L. Zhao, H. Yu and D. Yang., On-line 3D active pose-graph SLAM based on key poses using graph topology and sub-maps. in *Proc. of IEEE Int. Conf. on Robot. and Autom. (ICRA)*, pp. 169-175, May 2019.
- [53] R. Valencia, M. Morta, J. Andrade-Cetto, and J. M. Porta. “Planning reliable paths with pose SLAM,” *IEEE Trans. on Robot.*, vol. 29, no. 4, pp. 1050-1059, Apr. 2013.
- [54] B. D. Flury. “Acceptance-rejection sampling made easy,” *SIAM Review*, vol. 32, no.3, pp. 474-476, Sep. 1990.
- [55] D. Bump, (2004). Lie Groups. Graduate Texts in Mathematics, vol. 225. Berlin: Springer.