POLITECNICO DI MILANO

Scuola di Ingegneria Industriale e dell'Informazione
Corso di Laurea Magistrale in Mathematical Engineering
Curriculum in Statistical Learning

On Intrinsic Cramér–Rao Bounds on the Lie Group $\mathrm{SE}(d)^n$ with an Application to Computer Vision

Relatore: **Prof. Roberto Notari**
Correlatore: **Prof. Marco Compagnoni**

Tesi di Laurea Magistrale di:

Michele Bertini
Codice persona **10679371**
Matricola **247099**

Anno Accademico 2024/2025

Non vediamo le cose come sono, ma come siamo.

Anaïs Nin

Abstract

This thesis investigates pose-graph estimation on the Lie group $\mathrm{SE}(d)^n$ from noisy relative measurements defined on a possibly disconnected measurement graph representing the available relative pose observations. The goal is to estimate a set of unknown absolute poses from relative measurements modeled as $H_{ij} = \mathrm{Exp}(\widehat{\xi}_{ij})\, T_i T_j^{-1}$, where $T_i \in \mathrm{SE}(d)$ are the absolute poses and ξ_{ij} is a zero-mean heteroscedastic Gaussian noise in the Lie algebra. This measurement model preserves group consistency, ensures noise-free exactness, and is invariant to global right actions on the poses.

Due to this invariance, maximum likelihood estimation does not admit a unique solution, but yields a globally consistent set of relative transformations over the measurement graph, enabling coherent inference even in the absence of direct measurements between all poses.

A detailed derivation of the Fisher Information Matrix (FIM) is presented using right-invariant perturbations and Lie group geometry. Two equivalent formulations are obtained: a classical sum of per-measurement contributions and a compact factorization separating the topological structure of the measurement graph from pose-dependent kinematics. The FIM is shown to be singular, with a nullspace dimension equal to $c\,\dim(\mathrm{SE}(d))$ for a graph with c connected components, reflecting the intrinsic gauge freedom.

To derive meaningful Cramér–Rao lower bounds, two approaches are developed. The first introduces anchor poses, one per connected component, restricting the problem to a Riemannian submanifold. The second formulates estimation intrinsically on the quotient manifold $\mathrm{SE}(d)^n/\mathrm{SE}(d)^c$. In both cases, useful bounds for unbiased estimators are derived.

Numerical simulations on sparse directed pose graphs in SE(2) with heteroscedastic noise validate the theoretical results, confirming the predicted rank deficiency and convergence of the estimator to the intrinsic bounds.

Sommario

Questa tesi studia il problema della stima di pose sul gruppo di Lie $\mathrm{SE}(d)^n$ a partire da misure relative rumorose definite su un grafo delle misure, possibilmente disconnesso, che rappresenta le osservazioni relative di pose disponibili. L'obiettivo è stimare un insieme di pose assolute incognite a partire da misure relative modellate come $H_{ij} = \mathrm{Exp}(\widehat{\xi_{ij}})\, T_i T_j^{-1}$, dove $T_i \in \mathrm{SE}(d)$ sono le pose assolute e ξ_{ij} è un rumore gaussiano eteroschedastico a media nulla agente nell'algebra di Lie $\mathfrak{se}(d)$. Questo modello di misura preserva la consistenza di gruppo, garantisce l'esattezza in assenza di rumore ed è invariante rispetto ad azioni globali a destra sulle pose.

A causa di tale invarianza, la stima di massima verosimiglianza non ammette una soluzione unica, ma produce un insieme globalmente consistente di trasformazioni relative sul grafo delle misure, consentendo un'inferenza coerente anche in assenza di misure dirette tra tutte le pose.

Viene fornita una derivazione dettagliata della Matrice di Informazione di Fisher (FIM) utilizzando perturbazioni destre e la geometria intrinseca dei gruppi di Lie. Si ottengono due formulazioni equivalenti: una classica come somma dei contributi di ciascuna misura e una fattorizzazione compatta che separa la struttura topologica del grafo delle misure dalla dipendenza cinematica delle pose. Si dimostra che la FIM è singolare, con un nucleo di dimensione pari a $c\,\dim(\mathrm{SE}(d))$ per un grafo con c componenti connesse, riflettendo l'ambiguità di gauge intrinseca.

Per derivare limiti inferiori di Cramér–Rao significativi, vengono sviluppati due approcci. Il primo introduce pose ancorate, cioè fissate al valore reale, una per ciascuna componente connessa, restringendo il problema di stima a una sottovarietà riemanniana. Il secondo formula il problema in modo intrinseco sulla varietà quoziente $\mathrm{SE}(d)^n/\mathrm{SE}(d)^c$. In entrambi i casi vengono derivati limiti inferiori utili per stimatori non distorti.

Simulazioni numeriche su un grafo di pose sparso in SE(2) con rumore eteroschedastico validano i risultati teorici, confermando la deficienza di rango prevista e la convergenza dello stimatore efficiente ai limiti ottenuti.

Contents

List of Figures

List of Graphs

Chapter 1

Introduction

The problem of estimating the absolute configuration of a collection of rigid bodies from noisy relative measurements arises naturally in a wide range of applications, including camera network calibration, robotics, and sensor network localization. In these settings, the unknown parameters are poses, combinations of rotations and translations, that belong to the Lie group $\mathrm{SE}(d)$, while the available data consist solely of relative transformations between pairs of poses. Such measurements can be naturally represented as arcs of a directed graph, whose nodes correspond to the unknown absolute poses.

A key challenge in pose-graph estimation lies in the intrinsic geometric nature of the parameter space. Since $\mathrm{SE}(d)$ is a non-Euclidean manifold with a rich group structure, classical estimation techniques developed in Euclidean spaces are not directly applicable. In particular, both the formulation of realistic noise models and the derivation of meaningful uncertainty bounds as Cramér-Rao bounds require tools from differential geometry and Lie group theory.

The work of N. Boumal [19] constitutes not only a key reference in this direction, but also a genuine source of inspiration for the present study. In that paper, intrinsic Cramér–Rao bounds are developed for estimation problems on Riemannian submanifolds and quotient manifolds. Among the examples considered, Boumal analyzes a synchronization problem on $\mathrm{SO}(3)$, where relative rotation measurements are modeled by perturbing the relative rotation matrices with additive Gaussian noise. This modeling choice leads to a tractable analytical framework and reveals a strong connection between the structure of the Fisher Information Matrix and the topology of the underlying measurement graph. However, it also introduces a conceptual limitation: the resulting noisy measurements do not, in general, lie on $\mathrm{SO}(3)$. As Boumal himself remarks, this model can be disputed, since one would naturally expect measured relative rotations to remain valid elements of the rotation group.

This observation motivates a fundamental question: *is it possible to construct a statistical model for relative measurements on* $\mathrm{SE}(d)^n$ *that both incorporates Gaussian noise and preserves the group structure of the measurements, while remaining consistent in the zero-noise case, in the sense that it exactly recovers the deterministic relative transformation?* Addressing this question is the starting point of the present investigation. We propose a right-invariant measurement model in which each relative transformation is obtained as the left-multiplicative perturbation of the true relative pose by the exponential of a Gaussian random variable in the Lie algebra, namely $H_{ij} = \mathrm{Exp}(\widehat{\xi}_{ij})\, T_i T_j^{-1}$,

with $T_i, T_j \in \mathrm{SE}(d)$ and ξ_{ij} is a zero-mean Gaussian random variable. This construction ensures that noisy measurements remain elements of $\mathrm{SE}(d)$, guarantees perfect consistency in the noise-free case (i.e., $H_{ij} = T_i T_j^{-1}$ when $\xi_{ij} = 0$), and yields a likelihood function that is invariant under global rigid-body transformations.

As in the synchronization problems studied by N. Boumal, the right-invariance of the model induces a *global gauge ambiguity*: any common rigid transformation applied to all poses leaves the likelihood unchanged. A consequence of this is that estimating the absolute poses from relative measurements does not admit a unique maximum likelihood solution. Nevertheless, the maximum likelihood estimation yields a set of relative transformations that are globally consistent with the measurement graph, allowing one to infer coherent relative poses for any pair of nodes, including arcs for which no noisy measurement H_{ij} was originally available.

A central contribution of this work is a detailed analysis of the structure of the Fisher Information Matrix induced by the proposed model. By linearizing the residuals with respect to right-invariant perturbations, we show that the FIM admits a compact factorization that separates the kinematic dependence on the poses from the topological structure of the measurement graph, revealing a block-Laplacian structure. Another consequence of the right-invariance of the model is that the Fisher Information Matrix associated with the estimation problem is singular, with a nullspace whose dimension depends on the number of connected components of the measurement graph. This explicitly highlights the intimate connection between estimation accuracy and graph connectivity already observed in the work of N. Boumal for SO(3), now within a geometrically consistent model on $\mathrm{SE}(d)$.

To derive meaningful lower bounds on the estimation error in the presence of gauge freedom, we adopt and specialize the intrinsic Cramér–Rao bound framework developed by N. Boumal. Two complementary approaches are considered. In the first, the ambiguity is removed by anchoring a non-empty subset of poses, with at least one anchor per connected component, thereby restricting the estimation problem to a Riemannian submanifold of $\mathrm{SE}(d)^n$. In the second, the problem is formulated intrinsically on the quotient manifold obtained by identifying pose configurations that differ only by global gauge transformations. Although these two approaches are conceptually very different, they lead to Cramér–Rao bounds with a remarkably similar algebraic structure, inherited directly from the topology of the measurement graph.

Beyond their theoretical interest, these results have direct practical relevance for problems such as extrinsic calibration of camera networks. In such systems, cameras with different orientations and positions can obtain noisy relative measurements of other cameras using specialized sensors, such as laser-based devices. The framework developed in this work provides both a realistic statistical model for these measurements and rigorous performance limits on the achievable accuracy of any unbiased estimator, thereby offering valuable insight into the fundamental limits of camera network calibration from relative observations.

Chapter 2

Groups, Differential Manifolds and Lie Groups

2.1 Groups

In mathematics, physics, and engineering, one is often interested in describing systems or objects in terms of the transformations they admit and the relations among them. Although the nature of these systems may vary significantly across different contexts, the transformations under consideration frequently exhibit common structural features.

In particular, such transformations are often characterized by the presence of intrinsic symmetries, by the possibility of composing elementary transformations to obtain more complex ones, and by the existence of inverse transformations reversing an other. These properties are largely independent of the specific nature of the system and instead reflect an underlying algebraic structure.

The abstraction of these features leads naturally to the notion of group, which provides a unifying and rigorous framework for their analysis. The definition and basic properties of groups are classical and well established in the literature; in this section we recall them following standard references in abstract algebra and group theory, such as [1, 2, 3, 4].

In group theory, two notational conventions are commonly used to denote the group operation: additive notation and multiplicative notation. In the additive setting, the operation is written as a sum, the identity element is denoted by 0, and the inverse of an element a by $-a$. In multiplicative notation, the operation is written as a product, the identity element is denoted by e, and the inverse of an element g by g^{-1}.

In what follows, we will adopt the multiplicative notation, writing the group operation as $g_1 * g_2$ or simply $g_1 g_2$ when no ambiguity arises.

Definition 2.1.1 (Group)**.** A pair $(G, *)$ is called a **group** if G is a nonempty set and

$$* : G \times G \to G, \qquad (g_1, g_2) \mapsto g_1 * g_2,$$

satisfies the following axioms:

1. **Associativity:** for all $g_1, g_2, g_3 \in G$,
$$(g_1 * g_2) * g_3 = g_1 * (g_2 * g_3).$$
2. **Identity element:** there exists $e \in G$ such that, for every $g \in G$,
$$g * e = e * g = g.$$
3. **Inverse element:** for every $g \in G$, there exists $g^{-1} \in G$ such that
$$g * g^{-1} = g^{-1} * g = e.$$
.

Proposition 2.1.2. In every group $(G, *)$ the following properties hold:

1. **Uniqueness of the identity:** if $e' \in G$ satisfies $g * e' = e' * g = g$ for every $g \in G$, then $e' = e$.
2. **Uniqueness of inverses:** given $g \in G$, if $h \in G$ satisfies $g * h = h * g = e$, then $h = g^{-1}$.
3. **Cancellation law:** given $a, b, c \in G$, if either $a * b = a * c$ or $b * a = c * a$, then $b = c$.

Proof. (1) In G, it holds $e' = e * e' = e$, where the first equality folows from the properties of e, while the second one follows from the properties of e'.

(2) In G, we have $g^{-1} = g^{-1} * e = g^{-1} * (g * h) = (g^{-1} * g) * h = e * h = h$.

(3) It is enough to multiply by the inverse of a, on the left for the first equality, and on the right for the second one. □

Definition 2.1.3 (Abelian group)**.** A group $(G, *)$ is called **abelian** if

$$a * b = b * a \qquad \forall\, a, b \in G.$$

Remark 2.1.4. Although abelian groups form an important class, most groups that appear in algebra, geometry, and physics are non-abelian. In such cases the order of the factors is fundamental, since in general $a * b \neq b * a$.

Groups can be classified according to the cardinality of the underlying set:

- **finite** if the set is finite;
- **countable** if it has the same cardinality as the natural numbers;
- **continuous** if it has the same cardinality as the real numbers.

Example 2.1.5. Examples of groups include:

- S_n, the symmetric group on n elements, consisting of all bijections (permutations) of the set $\{1, \ldots, n\}$ with composition as the group operation; it is finite and non-abelian for $n \geq 3$;
- $(\mathbb{Z}, +)$, a countable abelian group;
- $(\mathbb{R}, +)$, a continuous abelian group;
- $\mathrm{GL}(n, \mathbb{R})$, the group of all invertible $n \times n$ real matrices under matrix multiplication; it is a continuous and non-abelian group for $n \geq 2$.

Definition 2.1.6 (Subgroup)**.** Given a group $(G, *)$, a subset $H \subseteq G$ is a **subgroup** if H becomes a group when endowed with the operation $*$ restricted to H, that is, with the operation inherited from G.

Proposition 2.1.7 (Subgroup criterion)**.** Let $(G, *)$ be a group and let $H \subseteq G$. Then $(H, *)$ is a subgroup of $(G, *)$ if and only if the following conditions hold:

1. H is nonempty;
2. for all $a, b \in H$, the product $a * b^{-1}$ belongs to H.

Example 2.1.8.

1. The set of even integers
$$2\mathbb{Z} = \{2k \mid k \in \mathbb{Z}\}$$
is a subgroup of $(\mathbb{Z}, +)$. Indeed, the sum of two even integers is still even, so $2\mathbb{Z}$ is closed under addition; it contains the identity element 0 (since $0 = 2 \cdot 0$); and for every even integer $2k$, its additive inverse $-2k$ is also even, hence belongs to $2\mathbb{Z}$. Thus $2\mathbb{Z}$ is a subgroup of $(\mathbb{Z}, +)$.

2. A less trivial example involves matrix groups. Consider $\mathrm{GL}(n, \mathbb{R})$, the group of all invertible $n \times n$ real matrices under matrix multiplication. Inside it, take the set
$$\mathrm{SL}(n, \mathbb{R}) = \{A \in \mathrm{GL}(n, \mathbb{R}) \mid \det(A) = 1\}.$$
This is a subgroup of $\mathrm{GL}(n, \mathbb{R})$: the product of two matrices with determinant 1 still has determinant 1 by Binet's theorem; the identity matrix has determinant 1; and for every $A \in \mathrm{SL}(n, \mathbb{R})$, its inverse A^{-1} also satisfies $\det(A^{-1}) = 1$. Thus $\mathrm{SL}(n, \mathbb{R})$ is a subgroup of $\mathrm{GL}(n, \mathbb{R})$ and, as we'll see in next section, it is one of the most fundamental example of a non-abelian Lie groups.

Definition 2.1.9 (Normal subgroup)**.** A subgroup $N \subseteq G$ is called a **normal subgroup** (or **invariant subgroup**) if it is closed under conjugation by elements of G, that is,
$$gNg^{-1} \subseteq N \qquad \text{for all } g \in G.$$
Equivalently, for every $n \in N$ and $g \in G$,
$$gng^{-1} \in N.$$
We denote normality by $N \trianglelefteq G$.

Example 2.1.10. The subgroup

$$Z(G) = \{\, g \in G : gx = xg \text{ for all } x \in G \,\},$$

called the **center** of G, is always a normal subgroup. Indeed, if $z \in Z(G)$ and $g \in G$, then

$$gz = zg \quad \text{ that is equivalent to } \quad gzg^{-1} = z \in Z(G),$$

because z commutes with every element of G. Thus $gzg^{-1} \in Z(G)$ for all g, and so $Z(G) \trianglelefteq G$. When G is abelian, then $Z(G) = G$. When $G = GL(n, \mathbb{R})$, $Z(G) = \{cI \mid c \in \mathbb{R} \setminus \{0\}\}$ the set of scalar matrices.

Definition 2.1.11 (Group homomorphism)**.** Let $(G, *)$ and $(H, \cdot)$ be two groups. A **group homomorphism** is a function $\Phi : G \to H$ such that

$$\Phi(u * v) = \Phi(u) \cdot \Phi(v)$$

for every $u, v \in G$. If Φ is invertible, it is called a **group isomorphism**. In this case, G and H are said to be isomorphic, and we write $G \cong H$.

Example 2.1.12. The determinant map

$$\det : \mathrm{GL}(n, \mathbb{R}) \longrightarrow \mathbb{R}^*$$

is a group homomorphism. Here $\mathrm{GL}(n, \mathbb{R})$ is considered as a group under matrix multiplication, while $\mathbb{R}^*$ (the set of nonzero real numbers) is a group under ordinary multiplication of real numbers $(\cdot)$.
To check that det is a homomorphism, we need to verify that it preserves the group operation. Given any $A, B \in \mathrm{GL}(n, \mathbb{R})$, we have the well-known identity,

$$\det(AB) = \det(A) \cdot \det(B).$$

which shows that det is a group homomorphism.

Definition 2.1.13 (Group action)**.** A (right) action of a group G on a set V is a map

$$\varphi : V \times G \to V, \qquad (p, g) \mapsto p \cdot g,$$

that satisfies:

1. **Identity:** $p \cdot e = p$ for every $p \in V$, where e is the identity of G;
2. **Compatibility:** $(p \cdot g) \cdot h = p \cdot (gh)$ for every $g, h \in G$ and $p \in V$.

Example 2.1.14. The general linear group $\mathrm{GL}(n, \mathbb{R})$ acts naturally on $\mathbb{R}^n$ as a right action. The action is given by matrix multiplication:

$$\mathbb{R}^n \times \mathrm{GL}(n, \mathbb{R}) \longrightarrow \mathbb{R}^n, \qquad (v, A) \longmapsto vA,$$

where vectors in $\mathbb{R}^n$ are regarded as row vectors. This is a group action of $\mathrm{GL}(n, \mathbb{R})$ on $\mathbb{R}^n$ because:

- for every $v \in \mathbb{R}^n$, the identity matrix satisfies

$$v \cdot I = vI = v;$$

- for all $A, B \in \mathrm{GL}(n, \mathbb{R})$ and $v \in \mathbb{R}^n$,

$$v \cdot (AB) = v(AB) = (vA)B = (v \cdot A) \cdot B.$$

2.2 Differentiable Manifolds

Many geometric objects encountered in mathematics and physics exhibit a complicated global behavior while locally resembling open subsets of some Euclidean space. This local–to–global intuition leads naturally to the concept of *manifold*, the basic object studied in differential geometry. The exposition follows the classical references of Lee [5] and Tu [7].

Definition 2.2.1 (Topological manifold)**.** A **topological manifold of dimension** n is a topological space M that satisfies:

- M is *Hausdorff* (any two distinct points admit disjoint neighbourhoods);
- M is *second-countable* (its topology has a countable base);
- every point $p \in M$ has a neighbourhood $U \subseteq M$ and a homeomorphism

$$\varphi : U \to \varphi(U) \subseteq \mathbb{R}^n,$$

  where $\varphi(U)$ is an open subset of $\mathbb{R}^n$ with respect to the Euclidean topology.

The pair (U, φ) is called a **chart**, and a collection of charts covering M is called an **atlas**.

Remark 2.2.2. If two charts (U, φ) and (V, ψ) overlap, the transition map

$$\psi \circ \varphi^{-1} : \varphi(U \cap V) \longrightarrow \psi(U \cap V)$$

is a homeomorphism between open subsets of $\mathbb{R}^n$.

Definition 2.2.3 (Differentiable compatibility)**.** Two charts (U, φ) and (V, ψ) on a topological manifold M are **C^k-compatible** if, whenever $U \cap V \neq \emptyset$, the transition map

$$\psi \circ \varphi^{-1}$$

is a C^k diffeomorphism.

Definition 2.2.4 (Differentiable atlas and smooth manifold)**.** A **C^k atlas** on M is a collection of pairwise C^k-compatible charts. A C^k atlas is **maximal** if it is not properly contained in any larger C^k atlas. A maximal C^k atlas defines a **C^k differentiable structure** on M. When $k = \infty$, M is called a **smooth manifold**.

Remark 2.2.5. Every C^k atlas is contained in a unique maximal one. Thus smooth structures correspond uniquely to maximal atlases.

Example 2.2.6 (The circle S^1)**.** The circle

$$S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$$

is a smooth manifold. In order to verify this according to the definitions, we explicitly construct an atlas and check that the transition maps are smooth. We use stereographic projections, which are standard tools for producing smooth charts on spheres.

Let $N = (0, 1)$ and $S = (0, -1)$ be the north and south poles. Consider the stereographic projection charts

$$\varphi_N : S^1 \setminus \{N\} \to \mathbb{R}, \qquad \varphi_N(x, y) = \frac{x}{1 - y},$$
$$\varphi_S : S^1 \setminus \{S\} \to \mathbb{R}, \qquad \varphi_S(x, y) = \frac{x}{1 + y}.$$

Step 1: Inverse maps (explicit computation). Let $u = \varphi_N(x, y) = \frac{x}{1-y}$. Then $x = u(1 - y)$, hence

$$y = 1 - \frac{x}{u}.$$

Impose the circle equation:

$$x^2 + y^2 = 1 \implies x^2 + \Big(1 - \frac{x}{u}\Big)^2 = 1.$$

Expanding,

$$x^2 + 1 - \frac{2x}{u} + \frac{x^2}{u^2} = 1 \implies x^2 + \frac{x^2}{u^2} - \frac{2x}{u} = 0.$$

Factor:

$$x\Big(x\Big(1 + \frac{1}{u^2}\Big) - \frac{2}{u}\Big) = 0.$$

Since $x = 0$ would force $y = 1$ (the excluded north pole), we use the second factor:

$$x\Big(1 + \frac{1}{u^2}\Big) = \frac{2}{u} \implies x = \frac{2u}{1 + u^2}.$$

Then

$$y = 1 - \frac{x}{u} = 1 - \frac{2}{1 + u^2} = \frac{u^2 - 1}{1 + u^2}.$$

Thus the inverse map is

$$\psi_N : \mathbb{R} \to S^1 \setminus \{N\}, \qquad \psi_N(u) = \Big(\frac{2u}{1 + u^2}, \frac{u^2 - 1}{1 + u^2}\Big).$$

A similar computation for $v = \varphi_S(x, y) = \frac{x}{1+y}$ gives

$$\psi_S(v) = \Big(\frac{2v}{1 + v^2}, \frac{1 - v^2}{1 + v^2}\Big).$$

Step 2: Transition functions. On the overlap $S^1 \setminus \{N, S\}$,

$$\varphi_S \circ \psi_N(u) = \frac{\frac{2u}{1+u^2}}{1 + \frac{u^2-1}{1+u^2}} = \frac{\frac{2u}{1+u^2}}{\frac{2u^2}{1+u^2}} = \frac{1}{u}.$$

Similarly,

$$\varphi_N \circ \psi_S(v) = \frac{1}{v}.$$

These functions are smooth on $\mathbb{R} \setminus \{0\}$. Thus the two stereographic charts are smoothly compatible.

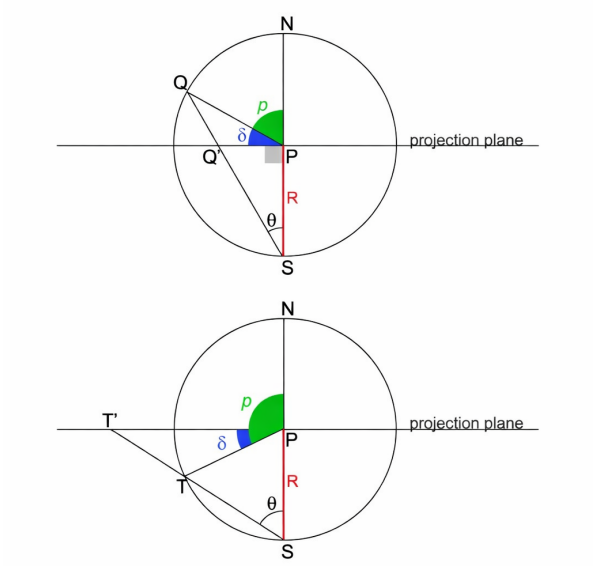


Figure 2.1: Stereographic projection in S^1, Q' (resp. T') denotes the stereographic projection of Q (resp. T) onto the projection plane.

Step 3: An atlas from projections onto the coordinate axes. Consider the open subsets

$$U_+ = \{(x,y) \in S^1 : y > 0\}, \qquad U_- = \{(x,y) \in S^1 : y < 0\},$$
$$V_+ = \{(x,y) \in S^1 : x > 0\}, \qquad V_- = \{(x,y) \in S^1 : x < 0\}.$$

These four sets cover S^1.
Define charts by projection onto the coordinate axes:

$$\chi_+ : U_+ \to (-1,1), \quad \chi_+(x,y) = x,$$
$$\chi_- : U_- \to (-1,1), \quad \chi_-(x,y) = x,$$
$$\eta_+ : V_+ \to (-1,1), \quad \eta_+(x,y) = y,$$
$$\eta_- : V_- \to (-1,1), \quad \eta_-(x,y) = y.$$

Their inverses are

$$\chi_+^{-1}(x) = \big(x, \sqrt{1-x^2}\big), \qquad \chi_-^{-1}(x) = \big(x, -\sqrt{1-x^2}\big),$$
$$\eta_+^{-1}(y) = \big(\sqrt{1-y^2}, y\big), \qquad \eta_-^{-1}(y) = \big(-\sqrt{1-y^2}, y\big),$$

which are smooth on their respective domains.

Compatibility (example). We check compatibility only for the chart χ_+ with φ_N. On $U_+ \cap (S^1 \setminus \{N\})$,

$$(\varphi_N \circ \chi_+^{-1})(x) = \frac{x}{1-\sqrt{1-x^2}}.$$

Rationalizing,

$$\frac{x}{1-\sqrt{1-x^2}} = \frac{1+\sqrt{1-x^2}}{x},$$

which is smooth on $(-1,1)\setminus\{0\}$.

Thus the stereographic charts (φ_N, φ_S) and the atlas given by projections onto the coordinate axes are mutually compatible. Therefore they define the standard smooth structure on S^1.

Definition 2.2.7 (Derivation at a point)**.** Let M be a smooth manifold and let $p \in M$. A **derivation at** p is a linear map

$$X : C^\infty(M) \to \mathbb{R}$$

satisfying the Leibniz rule

$$X(fg) = f(p)X(g) + g(p)X(f) \quad \text{for all } f, g \in C^\infty(M).$$

Definition 2.2.8 (Tangent space)**.** Let M be a smooth manifold and let $p \in M$. The **tangent space** at p, denoted T_pM, is the vector space of all derivations at p. Elements of T_pM are called **tangent vectors at** p.

Remark 2.2.9. Note that this definition does not depend on the choice of a chart. Indeed, if (U, φ) and (V, ψ) are two smooth charts around p, the coordinate expressions of a tangent vector differ by the Jacobian of the transition map, but the derivation itself—namely a linear map $X : C^\infty(M) \to \mathbb{R}$ satisfying the Leibniz rule—is intrinsic and does not change when passing from one chart to another. Thus the tangent space T_pM is defined independently of local coordinates.

Remark 2.2.10. If (U, φ) is a chart with coordinates $(x^1, \dots, x^n)$, then

$$\left\{ \left.\frac{\partial}{\partial x^1}\right|_p, \dots, \left.\frac{\partial}{\partial x^n}\right|_p \right\}$$

is a basis of T_pM.

Definition 2.2.11 (Smooth map and differential)**.** Let M and N be smooth manifolds. A map $f : M \to N$ is **smooth at** p if for charts (U, φ) at p and (V, ψ) at $f(p)$, the map

$$\psi \circ f \circ \varphi^{-1}$$

is smooth at $\varphi(p)$, that is, it admits partial derivatives of all orders in an open neighbourhood of $\varphi(p) \subset \mathbb{R}^m$. If this holds for all $p \in M$, then f is a **smooth map**. For such an f, the **differential** of f at p is the linear map

$$df_p : T_pM \to T_{f(p)}N$$

defined by

$$(df_pX)(g) = X(g \circ f), \qquad g \in C^\infty(N).$$

A bijective smooth map with smooth inverse is a **diffeomorphism**.

Definition 2.2.12 (Submanifold)**.** Let M be a smooth manifold of dimension n and let $S \subseteq M$. We say that S is a **smooth submanifold of dimension** k if for every point $p \in S$ there exist

- an open set $U \subseteq \mathbb{R}^k$,
- a smooth map $\Phi : U \to M$

such that

1. Φ is injective and its differential $d\Phi_q : T_qU \to T_{\Phi(q)}M$ is injective for every $q \in U$ which makes Φ an injective immersion,
2. Φ is a homeomorphism onto its image,
3. $p \in \Phi(U)$ and $\Phi(U) \subseteq S$.

The smooth structure on S is the unique one for which all such maps Φ are smooth embeddings.

Example 2.2.13. Since S^1 is a manifold we can see that an example of submanifold is any nonempty open arc. For instance,

$$U = \{(\cos\theta, \sin\theta) \in S^1 : \theta \in (\alpha, \beta)\}$$

is a 1-dimensional submanifold of S^1. Indeed, the map

$$(\alpha, \beta) \to U, \qquad \theta \mapsto (\cos\theta, \sin\theta)$$

is a smooth parametrization and a homeomorphism onto its image with smooth inverse, so U inherits the structure of a smooth 1-dimensional manifold embedded in S^1.
Thus open arcs are submanifolds of S^1, while closed arcs or single points are not submanifolds of dimension 1.

Definition 2.2.14 (Quotient manifold)**.** Let M be a smooth manifold and $\sim$ an equivalence relation on M. The quotient space $M/\sim$ carries the quotient topology induced by the projection

$$\pi : M \to M/\sim, \qquad p \mapsto [p].$$

If there exists a unique smooth structure on $M/\sim$ such that π is a smooth submersion, then $M/\sim$ is called a **quotient manifold**.

Example 2.2.15. The real projective space

$$\mathbb{RP}^n = S^n/(x \sim -x)$$

is a quotient manifold of the sphere. Indeed, consider the free smooth action of the group $\mathbb{Z}_2 = \{1, -1\}$ on S^n defined by

$$\varepsilon \cdot x = \varepsilon x, \qquad \text{that is, } 1 \cdot x = x,\ -1 \cdot x = -x.$$

This action identifies each point $x \in S^n$ with its antipodal point $-x$. Since the action is smooth, free (no nontrivial element fixes a point), and proper, the quotient

$$S^n/\mathbb{Z}_2$$

inherits a unique smooth structure making the projection

$$\pi : S^n \longrightarrow S^n/\mathbb{Z}_2$$

a smooth local diffeomorphism.
Moreover, two points $x, y \in S^n$ lie in the same orbit under this action if and only if $y = x$ or $y = -x$. Thus the orbits are precisely the lines through the origin in $\mathbb{R}^{n+1}$ intersected with the sphere, which is the usual definition of real projective space:

$$\mathbb{RP}^n = \big\{\text{all 1–dimensional linear subspaces of } \mathbb{R}^{n+1}\big\}.$$

Since each orbit consists of exactly two antipodal points, the quotient map

$$\pi : S^n \to \mathbb{RP}^n$$

is a smooth two-to-one covering map. Thus $\mathbb{RP}^n$ is a smooth manifold obtained as a smooth quotient of S^n by the antipodal identification.

Definition 2.2.16 (Tangent bundle)**.** The **tangent bundle** of M is

$$TM = \bigsqcup_{p \in M} T_pM,$$

with projection $\pi(v_p) = p$. Locally, in a chart (U, φ),

$$\pi^{-1}(U) \simeq \varphi(U) \times \mathbb{R}^n.$$

As in the case of the tangent space, the description of the tangent bundle does not depend on the choice of a chart. A chart provides a local trivialization, but different charts give different coordinate expressions which are related by smooth transition maps.

In fact, the various local representations of the tangent bundle can be viewed as different representatives of the same equivalence class. Each chart (U, φ) induces a local trivialization

$$\pi^{-1}(U) \to \varphi(U) \times \mathbb{R}^n,$$

and on overlaps these trivializations are related by smooth transition maps involving the Jacobian of the coordinate change. Thus TM itself is the intrinsic object, while any particular trivialization is only a coordinate-dependent representative of its equivalence class.

Definition 2.2.17 (Vector field)**.** A **smooth vector field** on M is a smooth section of the tangent bundle,

$$X : M \to TM, \qquad \pi \circ X = \mathrm{id}_M.$$

The space of smooth vector fields is denoted $\mathfrak{X}(M)$.

The concepts introduced above constitute the analytic and geometric backbone of differential geometry. They will be essential in the study of Lie groups: the tangent space at the identity of a Lie group will naturally acquire the structure of a Lie algebra, encoding the infinitesimal symmetries of the group.

2.3 Lie Groups

A Lie group is a mathematical structure that combines the algebraic properties of a group with the geometric structure of a smooth manifold. So, not only we can multiply the elements in such very interesting mathematical objects, but we can also perform calculus on them. The compatibility between these two structures makes Lie groups the natural framework for the study of continuous symmetries in geometry, physics and engineering (see for instance Helgason and Hall [8, 10]).

Definition 2.3.1 (Lie group)**.** A **Lie group** is a smooth manifold G endowed with a group structure $(G, *)$ such that the maps

$$m : G \times G \to G, \qquad m(g, h) = g * h,$$

called the **multiplication map**, and

$$\text{inv} : G \to G, \qquad \text{inv}(g) = g^{-1}$$

called the **inversion map**, are smooth. In other words, the group operation and the inversion are compatible with the smooth structure of G.

Remark 2.3.2. Lie groups combine algebra and geometry in a highly rigid way: the manifold structure allows one to perform differential calculus, while the group structure specifies how local motions compose.

A particularly concrete class of Lie groups is given by the **matrix Lie groups**, that is, subsets of the space of real matrices $M_n(\mathbb{R})$ which are closed under matrix multiplication and inversion, and which are also smooth submanifolds of $M_n(\mathbb{R})$.
Typical examples include:

- the **general linear group**

$$\text{GL}(n, \mathbb{R}) = \{A \in M_n(\mathbb{R}) \mid \det A \neq 0\};$$

- the **special linear group**

$$\text{SL}(n, \mathbb{R}) = \{A \in M_n(\mathbb{R}) : \det A = 1\};$$

- the **orthogonal group**

$$\text{O}(n) = \{A \in M_n(\mathbb{R}) : AA^T = I_n\};$$

- the **special orthogonal group**

$$\text{SO}(n) = \text{O}(n) \cap \text{SL}(n, \mathbb{R});$$

- the **special Euclidean group** (orientation-preserving isometries of the d-dimensional Euclidean space $\mathbb{E}^d \simeq \mathbb{R}^d$)

$$\text{SE}(d) = \left\{ \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \,\middle|\, R \in \text{SO}(d),\ t \in \mathbb{R}^d \right\}.$$

Remark 2.3.3. Each of the groups above is a Lie group: they inherit their smooth structure from the Euclidean space $M_n(\mathbb{R})$, and the group operations are smooth because matrix multiplication and inversion are smooth maps. Matrix Lie groups therefore provide explicit and important examples of Lie groups, and will play a central role in the study of rigid motions and rotations.

Definition 2.3.4 (Translations, Invariant vector fields)**.** Let G be a Lie group. For any $g \in G$, the **right translation** and **left translation** by g are the diffeomorphisms

$$R_g : G \to G, \qquad R_g(h) = h * g, \tag{2.1}$$

$$L_g : G \to G, \qquad L_g(h) = g * h. \tag{2.2}$$

A vector field $X \in \mathfrak{X}(G)$ is called **right-invariant** if

$$dR_g(X_h) = X_{h*g}, \qquad \forall\, g, h \in G, \tag{2.3}$$

and **left-invariant** if

$$dL_g(X_h) = X_{g*h}, \qquad \forall\, g, h \in G. \tag{2.4}$$

Remark 2.3.5. Right-invariant (and Left-invariant) vector fields are uniquely determined by their value at the identity, and this identification allows one to define the Lie algebra of G as the tangent space T_eG.

2.4 Lie Algebras and the Exponential Map

In this section we recall the notion of Lie algebra and its relation with Lie groups. Standard references for Lie groups and Lie algebras include the textbooks by Warner and Hall [9, 10].

Definition 2.4.1 (Lie Algebra, Lie bracket)**.** A **Lie algebra** over $\mathbb{R}$ is a real vector space $\mathfrak{g}$ equipped with a bilinear map

$$[\,\cdot\,,\,\cdot\,] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g},$$

called the **Lie bracket**, satisfying the following properties for all $X, Y, Z \in \mathfrak{g}$:

1. **Antisymmetry:**
$$[X, Y] = -[Y, X].$$

2. **Jacobi Identity:**
$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

Intuitively, a Lie algebra is expected to encode the infinitesimal structure of a Lie group, describing how elements "infinitesimally close" to the identity combine and interact.

Let G be a Lie group. The tangent space at the identity

$$\mathfrak{g} = T_e G$$

is a real vector space. Right translation allows one to transport tangent vectors at e to all other points of G.

Theorem 2.4.2. Let G be a Lie group and $\mathfrak{g} = T_e G$ its tangent space at the identity. For each $X \in \mathfrak{g}$ there exists a unique right-invariant vector field $\widetilde{X} \in \mathfrak{X}(G)$ such that $\widetilde{X}_e = X$. If $X, Y \in \mathfrak{g}$ and $\widetilde{X}, \widetilde{Y}$ are the corresponding right-invariant vector fields, then

$$[X, Y] := [\widetilde{X}, \widetilde{Y}]_e$$

defines a Lie bracket on $\mathfrak{g}$ which makes $\mathfrak{g}$ a Lie algebra.

Idea of the proof. The existence and uniqueness of right-invariant vector fields follow from the properties of the differential of right translations. The commutator $[\widetilde{X}, \widetilde{Y}]$ of two right-invariant vector fields is again right-invariant, and its value at e depends bilinearly and antisymmetrically on X and Y and satisfies the Jacobi identity. For a detailed treatment we refer to [5, 10]. □

Definition 2.4.3 (Lie algebra of G)**.** The vector space $\mathfrak{g} = T_e G$ with the Lie bracket $[X, Y] := [\widetilde{X}, \widetilde{Y}]_e$ of Theorem (2.4.2) is called the **Lie algebra of G**.

Definition 2.4.4 (Exponential map)**.** Let G be a Lie group and $\mathfrak{g}$ its Lie algebra. For $X \in \mathfrak{g}$, consider the unique integral curve $\gamma_X(t)$ of the right-invariant vector field associated with X, satisfying

$$\gamma_X(0) = e, \qquad \gamma_X'(0) = X.$$

The **exponential map** is defined as

$$\mathrm{Exp} : \mathfrak{g} \to G, \qquad \mathrm{Exp}(X) = \gamma_X(1).$$

It is a smooth map satisfying $\mathrm{d}_0\mathrm{Exp} = \mathrm{id}_{\mathfrak{g}}$; hence it is a local diffeomorphism around $0 \in \mathfrak{g}$ and $e \in G$.

Remark 2.4.5. In the case of a matrix Lie group $G \subset \mathrm{GL}(n, \mathbb{R})$, the exponential map coincides with the usual exponential matrix:

$$\mathrm{Exp}(X) = \sum_{k=0}^{\infty} \frac{X^k}{k!}, \qquad X \in \mathfrak{g}.$$

Definition 2.4.6 (Logarithm map)**.** One can define locally the inverse of the exponential map, called the **logarithm map**. In the case of a matrix Lie group $G \subset \mathrm{GL}(n, \mathbb{R})$, the logarithm map coincides with the classical logarithm matrix.
For any element $Y \in G$ sufficiently close to the identity I, the logarithm is given by the convergent series

$$\mathrm{Log}(Y) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} (Y - I)^k,$$

which converges whenever $\|Y - I\| < 1$ and yields an element of the Lie algebra $\mathfrak{g}$.

More generally, there exists an open neighborhood $U \subset G$ of the identity element I such that the map

$$\mathrm{Log} : U \to \mathfrak{g}$$

is well defined and smooth, and satisfies

$$\mathrm{Exp}(\mathrm{Log}(Y)) = Y, \qquad \forall\, Y \in U,$$

as well as

$$\mathrm{Log}(\mathrm{Exp}(X)) = X, \qquad \forall\, X \in \mathfrak{g} \text{ with } \|X\| \text{ sufficiently small.}$$

Example 2.4.7 (The Lie algebra of the Special Euclidean Group $\mathrm{SE}(d)$)**.** The **special Euclidean group** $\mathrm{SE}(d)$ consists of all orientation-preserving rigid motions of the Euclidean space $\mathbb{R}^d$. Each such motion is uniquely determined by a rotation $R \in \mathrm{SO}(d)$ and a translation $t \in \mathbb{R}^d$. Indeed, once we fix coordinates in $\mathbb{E}^d$, a rigid motion is described by the equation $x' = Rx + t$.
Using homogeneous coordinates $X = \begin{pmatrix} x \\ 1 \end{pmatrix}$, the equation above can be rewritten as

$$X' = \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix} X,$$

which is the notation we adopt throughout this thesis. This yields the description

$$\mathrm{SE}(d) = \left\{ \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix} \;\middle|\; R \in \mathrm{SO}(d),\ t \in \mathbb{R}^d \right\}.$$

Semidirect product structure. The multiplication in $\mathrm{SE}(d)$ is performed as follows

$$\begin{pmatrix} R_1 & t_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} R_2 & t_2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} R_1 R_2 & t_1 + R_1 t_2 \\ 0 & 1 \end{pmatrix},$$

that proves the well-known fact that $\mathrm{SE}(d)$ is the semidirect product of $\mathrm{SO}(d)$ and $\mathbb{R}^d$ associated to the group homomorphism $\alpha : \mathrm{SO}(d) \to \mathrm{Aut}(\mathbb{R}^d)$ defined as $\alpha(R) : \mathbb{R}^d \to \mathbb{R}^d, \alpha(R)t = Rt$.
Thus $\mathrm{SO}(d)$ acts on $\mathbb{R}^d$ by the standard linear action $t \mapsto Rt$, and

$$\mathrm{SE}(d) \cong \mathrm{SO}(d) \ltimes \mathbb{R}^d,$$

with the semidirect product law

$$(R_1, t_1) \cdot (R_2, t_2) = (R_1 R_2,\ t_1 + R_1 t_2).$$

The dimension is therefore

$$\dim \mathrm{SE}(d) = \dim \mathrm{SO}(d) + \dim \mathbb{R}^d = \frac{d(d-1)}{2} + d = \frac{d(d+1)}{2}. \tag{2.5}$$

The inverse of an element is the following

$$\begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} R^T & -R^T t \\ 0 & 1 \end{pmatrix}$$

because $R \in \mathrm{SO}(d)$. It is clear that both the group multiplication and inversion are defined through matrix operations, and are therefore smooth. Moreover, since the order of multiplication is essential, the group operation is not commutative, and hence $\mathrm{SE}(d)$ is a non-abelian Lie group.

Translations.

$$T(d) = \left\{ \begin{pmatrix} I & t \\ 0 & 1 \end{pmatrix} : t \in \mathbb{R}^d \right\} \subset \mathrm{SE}(d).$$

(1) $T(d)$ **is a subgroup of** $\mathrm{SE}(d)$ **(subgroup criterion).** First, $T(d) \neq \varnothing$ since

$$e = \begin{pmatrix} I & 0 \\ 0 & 1 \end{pmatrix} \in T(d).$$

Let

$$s_1 = \begin{pmatrix} I & t_1 \\ 0 & 1 \end{pmatrix}, \qquad s_2 = \begin{pmatrix} I & t_2 \\ 0 & 1 \end{pmatrix} \in T(d).$$

Then

$$s_2^{-1} = \begin{pmatrix} I & -t_2 \\ 0 & 1 \end{pmatrix},$$

and thus

$$s_1 s_2^{-1} = \begin{pmatrix} I & t_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} I & -t_2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} I & t_1 - t_2 \\ 0 & 1 \end{pmatrix} \in T(d).$$

By the subgroup criterion, $T(d)$ is a subgroup of $\mathrm{SE}(d)$.
(2) $T(d)$ **is normal in** $\mathrm{SE}(d)$**.** Let

$$g = \begin{pmatrix} R & q \\ 0 & 1 \end{pmatrix} \in \mathrm{SE}(d), \qquad s = \begin{pmatrix} I & t \\ 0 & 1 \end{pmatrix} \in T(d).$$

Since

$$g^{-1} = \begin{pmatrix} R^{-1} & -R^{-1}q \\ 0 & 1 \end{pmatrix},$$

we compute

$$gsg^{-1} = \begin{pmatrix} R & q \\ 0 & 1 \end{pmatrix} \begin{pmatrix} I & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} R^{-1} & -R^{-1}q \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} I & Rt \\ 0 & 1 \end{pmatrix} \in T(d),$$

because $Rt \in \mathbb{R}^d$. Hence $gTg^{-1} \subseteq T$ for all $g \in \mathrm{SE}(d)$, so $T \trianglelefteq \mathrm{SE}(d)$.

Lie algebra. Since $\mathrm{SE}(d)$ is the semidirect product of $\mathrm{SO}(d)$ and $\mathbb{R}^d$, its tangent space at the identity decomposes as the direct sum of the tangent space at the identity of $\mathrm{SO}(d)$ and the tangent space at the origin of $\mathbb{R}^d$. The second one is immediate:

$$T_0\mathbb{R}^d \cong \mathbb{R}^d.$$

To compute the tangent space of $\mathrm{SO}(d)$ at the identity, consider the curve

$$\gamma : (-\delta, \delta) \to \mathrm{SO}(d), \qquad \gamma(t) = e^{tX},$$

for some $X \in M_d(\mathbb{R})$ and small $\delta > 0$. Since

$$(e^{tX})^T = e^{tX^T},$$

the condition $\gamma(t) \in \mathrm{SO}(d)$ for all t implies

$$e^{tX^T} e^{tX} = I,$$

which reduces (differentiating at $t = 0$) to

$$X + X^T = 0.$$

Thus every tangent vector X must be skew–symmetric, and conversely every skew–symmetric X gives such a curve. Therefore

$$\mathfrak{so}(d) = \{\, \Omega \in M_d(\mathbb{R}) \mid \Omega^T = -\Omega \,\},$$

consists of the vector space of skew–symmetric $d \times d$ matrices.
Combining the two components, the Lie algebra of $\mathrm{SE}(d)$ is

$$\mathfrak{se}(d) = \left\{ \begin{pmatrix} \Omega & v \\ 0 & 0 \end{pmatrix} \;\middle|\; \Omega^T = -\Omega,\ v \in \mathbb{R}^d \right\}.$$

Hence, as a vector space one has the decomposition

$$\mathfrak{se}(d) \;\cong\; \mathfrak{so}(d) \oplus \mathbb{R}^d,$$

and it is natural to build a basis by combining a basis of $\mathfrak{so}(d)$ (infinitesimal rotations) with the canonical basis of $\mathbb{R}^d$ (infinitesimal translations).
Let $\{e_1, \ldots, e_d\}$ denote the standard basis of $\mathbb{R}^d$, and let $e_i e_j^\top$ be the rank-one matrix having a 1 in position (i, j) and zeros elsewhere. For each pair of indices $1 \le i < j \le d$, define

$$E_{ij} := e_i e_j^\top - e_j e_i^\top.$$

By construction, $E_{ij}^\top = -E_{ij}$, hence $E_{ij} \in \mathfrak{so}(d)$. Geometrically, E_{ij} represents the infinitesimal generator of a rotation in the (i, j)-plane. The family $\{E_{ij}\}_{1 \le i < j \le d}$ is linearly independent and has cardinality $\frac{d(d-1)}{2}$, which coincides with $\dim \mathfrak{so}(d)$; therefore, it forms a basis of $\mathfrak{so}(d)$.
Translations correspond to the $\mathbb{R}^d$ component of the Lie algebra. A natural and canonical choice of generators is given by the standard vectors $\{e_k\}_{k=1}^d$. Combining rotational and translational generators, one obtains a canonical basis of $\mathfrak{se}(d)$ through the homogeneous matrix representation. For $1 \le i < j \le d$, define the rotational basis elements

$$\widehat{R}_{ij} := \begin{pmatrix} E_{ij} & 0 \\ 0 & 0 \end{pmatrix},$$

and for $k = 1, \dots, d$, define the translational basis elements

$$\widehat{T}_k := \begin{pmatrix} 0 & e_k \\ 0 & 0 \end{pmatrix}.$$

The collection

$$\mathcal{B} = \left\{ \widehat{R}_{ij} \ : \ 1 \le i < j \le d \right\} \ \cup \ \left\{ \widehat{T}_k \ : \ k = 1, \dots, d \right\} \tag{2.6}$$

forms a basis of $\mathfrak{se}(d)$. Indeed, any element $\xi \in \mathfrak{se}(d)$ can be uniquely expressed as

$$\xi = \sum_{1 \le i < j \le d} \omega_{ij} \, \widehat{R}_{ij} \ + \ \sum_{k=1}^{d} v_k \, \widehat{T}_k,$$

where the coefficients ω_{ij} describe the skew-symmetric component $\Omega \in \mathfrak{so}(d)$, while the vector $v = (v_1, \dots, v_d)^\top$ encodes the translational component. Finally, the total number of basis elements is

$$\frac{d(d-1)}{2} + d = \frac{d(d+1)}{2} = \dim \mathfrak{se}(d),$$

which confirms that $\mathcal{B}$ is a canonical and minimal generating set for the Lie algebra of $\mathrm{SE}(d)$.

Lie bracket.
The Lie bracket on $\mathfrak{se}(d)$ is defined as the matrix commutator. For two elements

$$X_i = \begin{pmatrix} \Omega_i & v_i \\ 0 & 0 \end{pmatrix}, \qquad i = 1, 2,$$

a direct multiplication gives

$$X_1 X_2 = \begin{pmatrix} \Omega_1 \Omega_2 & \Omega_1 v_2 \\ 0 & 0 \end{pmatrix}, \qquad X_2 X_1 = \begin{pmatrix} \Omega_2 \Omega_1 & \Omega_2 v_1 \\ 0 & 0 \end{pmatrix}.$$

Subtracting the two matrices yields

$$[X_1, X_2] = \begin{pmatrix} \Omega_1 \Omega_2 - \Omega_2 \Omega_1 & \Omega_1 v_2 - \Omega_2 v_1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} [\Omega_1, \Omega_2] & \Omega_1 v_2 - \Omega_2 v_1 \\ 0 & 0 \end{pmatrix},$$

where $[\Omega_1, \Omega_2]$ denotes the ordinary commutator in $\mathfrak{so}(d)$.

The form of this bracket reflects directly the semidirect-product structure of $\mathrm{SE}(d)$. The first component behaves exactly like the bracket in $\mathfrak{so}(d)$, while the second component couples the linear part Ω_i with the translational part v_j in the expected way: the infinitesimal rotation Ω_1 acts on the vector v_2, and symmetrically for the other term.

Vector representation and hat–vee operators

Every finite-dimensional Lie algebra $\mathfrak{g}$ over $\mathbb{R}$ is, as a vector space, isomorphic to $\mathbb{R}^n$, where $n = \dim \mathfrak{g}$. This means that we can choose a basis $\{E_1, \dots, E_n\}$ of $\mathfrak{g}$ and represent any element $X \in \mathfrak{g}$ uniquely as

$$X = \sum_{i=1}^{n} x_i E_i, \qquad \text{with } x = (x_1, \dots, x_n)^\top \in \mathbb{R}^n.$$

The isomorphism between the algebra and its coordinate representation is then

$$\hat{(\cdot)} : \mathbb{R}^n \to \mathfrak{g}, \quad x \mapsto \hat{x} = \sum_{i=1}^{n} x_i E_i,$$
$$(\cdot)^\vee : \mathfrak{g} \to \mathbb{R}^n, \quad X \mapsto X^\vee = (x_1, \ldots, x_n)^\top,$$

where $\hat{(\cdot)}$ and $(\cdot)^\vee$ are called the **hat** and **vee** operators, respectively.

Example: The case of $\mathfrak{so}(3)$

The Lie algebra $\mathfrak{so}(3)$ consists of all 3×3 skew-symmetric matrices:

$$\mathfrak{so}(3) = \left\{ \Omega \in \mathbb{R}^{3\times 3} \;\middle|\; \Omega^\top = -\Omega \right\}.$$

A standard basis of $\mathfrak{so}(3)$ is given by the three matrices

$$E_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad E_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \qquad E_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Every element $\Omega \in \mathfrak{so}(3)$ can be written uniquely as

$$\Omega = \omega_1 E_1 + \omega_2 E_2 + \omega_3 E_3, \qquad \omega = (\omega_1, \omega_2, \omega_3)^\top \in \mathbb{R}^3,$$

This induces the standard linear isomorphism

$$\widehat{(\cdot)} : \mathbb{R}^3 \xrightarrow{\cong} \mathfrak{so}(3), \qquad \omega \longmapsto \hat{\omega},$$

defined by

$$\hat{\omega} = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}.$$

The inverse map, called the *vee operator*, is

$$(\cdot)^\vee : \mathfrak{so}(3) \xrightarrow{\cong} \mathbb{R}^3, \qquad \Omega^\vee = \begin{pmatrix} \Omega_{32} \\ \Omega_{13} \\ \Omega_{21} \end{pmatrix}.$$

Special identity in dimension 3. In dimension 3 the hat operator encodes the cross product, via

$$\hat{\omega}\, v = \omega \times v, \qquad \forall\, \omega, v \in \mathbb{R}^3.$$

This correspondence is specific to $\mathbb{R}^3$; for $\mathfrak{so}(d)$ with $d \neq 3$ no such cross-product representation exists.

Extension to $\mathfrak{se}(3)$

The Lie algebra $\mathfrak{se}(3)$ of the special Euclidean group SE(3) consists of all matrices of the form

$$\mathfrak{se}(3) = \left\{ \begin{pmatrix} \hat{\omega} & v \\ 0 & 0 \end{pmatrix} \;\middle|\; \omega, v \in \mathbb{R}^3 \right\}.$$

Given $\xi = (\omega, v)^T \in \mathbb{R}^6$, we define the hat and vee maps

$$\widehat{\xi} := \begin{pmatrix} \hat{\omega} & v \\ 0 & 0 \end{pmatrix}, \qquad (\widehat{\xi})^{\vee} = \begin{pmatrix} \omega \\ v \end{pmatrix}.$$

The hat and vee operators provide a linear bijection between $\mathbb{R}^6$ and $\mathfrak{se}(3)$, so that, as vector spaces,

$$\mathfrak{se}(3) \simeq \mathbb{R}^6.$$

It is important to emphasize that this correspondence is purely linear: the Lie bracket on $\mathfrak{se}(3)$ induces a non-trivial bracket on $\mathbb{R}^6$ via the identification above. Therefore, $\mathfrak{se}(3)$ and $\mathbb{R}^6$ share the same vector space structure, but not the same Lie algebra structure unless the latter is endowed with the transported bracket.
This representation is nevertheless extremely useful in applications, as it enables a compact vector form for rigid-body motions and their infinitesimal generators, a key tool in robotics, computer vision, and geometric mechanics.

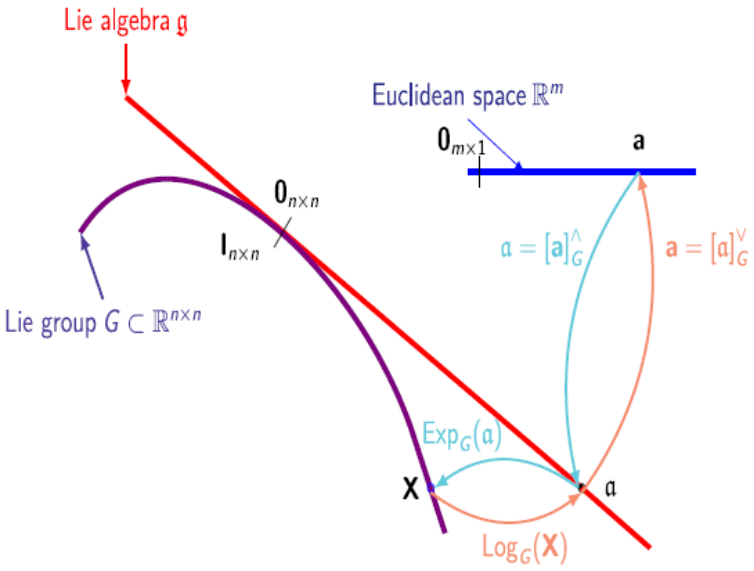


Figure 2.2: Relation between Lie Groups, Lie Algebras and $\mathbb{R}^m$. Image reproduced from [22].

2.5 Adjoint Operator and Jacobians on Lie Groups

We now introduce the Adjoint representation and the Jacobians associated with the exponential map on a Lie group. The presentation is inspired by standard treatments in geometric control and robotics, see e.g. [12, 30].

Definition 2.5.1 (Adjoint operator)**.** Let G be a Lie group with Lie algebra $\mathfrak{g}$. For each $g \in (G, *)$, the Adjoint operator associated with g is the linear map

$$\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}$$

defined by

$$\mathrm{Ad}_g(X) = \frac{d}{dt}\bigg|_{t=0} g * \mathrm{Exp}(tX) * g^{-1}, \qquad X \in \mathfrak{g}.$$

Intuitively, $\mathrm{Ad}_g(X)$ represents how an infinitesimal element X of the algebra is transformed when it is transported through the group element g via conjugation.

Proposition 2.5.2. The map

$$\mathrm{Ad} : G \longrightarrow \mathrm{Aut}(\mathfrak{g}), \qquad g \longmapsto \mathrm{Ad}_g,$$

is a group homomorphism, that is,

$$\mathrm{Ad}_{g*h} = \mathrm{Ad}_g \circ \mathrm{Ad}_h, \qquad \mathrm{Ad}_e = \mathrm{id}_{\mathfrak{g}}.$$

Sketch of the proof. The identity $\mathrm{Ad}_e = \mathrm{id}_{\mathfrak{g}}$ follows immediately from the definition. For $g, h \in G$, using the definition of Ad_g and the group law,

$$g * h * \mathrm{Exp}(tX) * h^{-1} * g^{-1} = g * \left(h * \mathrm{Exp}(tX) * h^{-1}\right) * g^{-1},$$

and differentiating at $t = 0$ yields $\mathrm{Ad}_{g*h}(X) = \mathrm{Ad}_g(\mathrm{Ad}_h(X))$. □

Proposition 2.5.3. For every $g \in G$, the Adjoint operator

$$\mathrm{Ad}_g : \mathfrak{g} \longrightarrow \mathfrak{g}$$

is a linear isomorphism that preserves the Lie bracket, namely

$$[\mathrm{Ad}_g X, \mathrm{Ad}_g Y] = \mathrm{Ad}_g[X, Y], \qquad \forall X, Y \in \mathfrak{g}.$$

Sketch of the proof. Linearity of Ad_g follows from the linearity of the derivative. Invertibility is obtained by observing that $\mathrm{Ad}_{g^{-1}}$ is the inverse of Ad_g.
The preservation of the Lie bracket follows from the definition of the bracket in terms of the commutator of right-invariant vector fields. Indeed, conjugation by g maps right-invariant vector fields into right-invariant vector fields, and the pushforward induced by conjugation preserves the Lie bracket. Evaluating at the identity yields

$$[\mathrm{Ad}_g X, \mathrm{Ad}_g Y] = \mathrm{Ad}_g[X, Y], \qquad \forall X, Y \in \mathfrak{g}.$$

□

Proposition 2.5.4. Let $G \subset \mathrm{GL}(n)$ be a matrix Lie group with Lie algebra $\mathfrak{g} \subset \mathfrak{gl}(n)$. Then, for every $T \in G$ and every $X \in \mathfrak{g}$, the adjoint operator acts as

$$\mathrm{Ad}_T(X) = TXT^{-1}. \tag{2.7}$$

Proof. By definition of the adjoint operator,

$$\mathrm{Ad}_T(X) = \left.\frac{d}{dt}\right|_{t=0} T\,\mathrm{Exp}(tX)\,T^{-1}.$$

Since matrix multiplication is bilinear and independent of t, we may pull the derivative inside the conjugation, obtaining

$$\mathrm{Ad}_T(X) = T\left(\left.\frac{d}{dt}\right|_{t=0} \mathrm{Exp}(tX)\right)T^{-1}.$$

Using the fact that $\left.\frac{d}{dt}\right|_{t=0}\mathrm{Exp}(tX) = X$, we conclude that

$$\mathrm{Ad}_T(X) = TXT^{-1}.$$

□

Proposition 2.5.5 (Conjugation and exponential for matrix Lie groups)**.** Let $G \subseteq \mathrm{GL}(n)$ be a matrix Lie group with Lie algebra $\mathfrak{g} \subseteq \mathbb{R}^{n\times n}$, and let $T \in G$. Then for every $X \in \mathfrak{g}$,

$$T\,\mathrm{Exp}(X)\,T^{-1} = \mathrm{Exp}\big(TXT^{-1}\big). \tag{2.8}$$

In particular the adjoint operator for matrix groups becomes

$$\mathrm{Ad}_T(X) = TXT^{-1},$$

then

$$T\,\mathrm{Exp}(X)\,T^{-1} = \mathrm{Exp}\big(\mathrm{Ad}_T X\big). \tag{2.9}$$

Moreover, the same identity holds for any (not necessarily matrix) Lie group.

Proof. (For matrix groups) Since G is a matrix Lie group, its exponential map coincides with the matrix exponential restricted to $\mathfrak{g}$. Recall that for any matrix A,

$$\mathrm{Exp}(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

Let $T \in \mathrm{GL}(n)$ and $X \in \mathbb{R}^{n\times n}$. Using absolute convergence of the series and the identity $(TXT^{-1})^k = TX^kT^{-1}$ (proved by induction), we compute:

$$\begin{aligned}\mathrm{Exp}(TXT^{-1}) &= \sum_{k=0}^{\infty}\frac{(TXT^{-1})^k}{k!} = \sum_{k=0}^{\infty}\frac{TX^kT^{-1}}{k!} = T\left(\sum_{k=0}^{\infty}\frac{X^k}{k!}\right)T^{-1}\\ &= T\,\mathrm{Exp}(X)\,T^{-1}.\end{aligned}$$

Rearranging yields

$$T\,\mathrm{Exp}(X)\,T^{-1} = \mathrm{Exp}(TXT^{-1}).$$

If G is a matrix Lie group and $T \in G$, then $TXT^{-1} \in \mathfrak{g}$ for all $X \in \mathfrak{g}$ (because conjugation by T is a Lie algebra automorphism), hence the above identity applies with $X \in \mathfrak{g}$ and Exp the group exponential. One can obtain the identity (2.9) for a generic Lie group, not necessarily a matrix Lie group. □

The infinitesimal counterpart of the Adjoint representation is obtained by differentiating Ad at the identity.

Definition 2.5.6 (Adjoint map)**.** The **adjoint map** of the Lie algebra $\mathfrak{g}$ is the linear map

$$\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g}), \qquad X \mapsto \mathrm{ad}_X,$$

defined by

$$\mathrm{ad}_X(Y) = [X, Y], \qquad X, Y \in \mathfrak{g}.$$

Proposition 2.5.7. For each $X \in \mathfrak{g}$ one has

$$\mathrm{ad}_X = \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}_{\mathrm{Exp}(tX)}.$$

Moreover, for all $X, Y \in \mathfrak{g}$,

$$\mathrm{ad}_{[X,Y]} = [\mathrm{ad}_X, \mathrm{ad}_Y],$$

where the bracket on the right-hand side denotes the commutator of endomorphisms of $\mathfrak{g}$.

Sketch of the proof. Differentiating the identity

$$\mathrm{Ad}_{\mathrm{Exp}(tX)} \circ \mathrm{Ad}_{\mathrm{Exp}(sY)} = \mathrm{Ad}_{\mathrm{Exp}(tX)\mathrm{Exp}(sY)}$$

with respect to t and s at $(0, 0)$ yields the stated relations. See [10] for a detailed argument. □

Left and right Jacobians of the exponential map

The exponential and logarithm maps allow one to move between the Lie algebra $\mathfrak{g}$ and the Lie group G. The derivatives of these maps play a crucial role in linearization, uncertainty propagation, and optimization on manifolds.

Definition 2.5.8 (Left and right Jacobians of the exponential map)**.** Let G be a Lie group with Lie algebra $\mathfrak{g}$ and exponential map $\mathrm{Exp} : \mathfrak{g} \to G$. For $X \in \mathfrak{g}$ we define the **left Jacobian** $J_\ell(X)$ and the **right Jacobian** $J_r(X)$ as the linear maps

$$J_\ell(X) = d_{\mathrm{Exp}(X)} L_{\mathrm{Exp}(-X)} \circ d_X \mathrm{Exp}, \qquad J_r(X) = d_{\mathrm{Exp}(X)} R_{\mathrm{Exp}(-X)} \circ d_X \mathrm{Exp}, \tag{2.10}$$

where L_g and R_g denote the left and right translations on G (2.1), (2.2). Thus $J_\ell(X), J_r(X) : \mathfrak{g} \to \mathfrak{g}$ are linear endomorphisms of the Lie algebra.

Intuitively, these operators describe how a small perturbation $\delta X \in \mathfrak{g}$ is transported through the exponential map when the perturbation is applied on the left or on the right. More precisely, for δX sufficiently small one has the first-order approximations

$$\mathrm{Exp}(X + \delta X) \approx \mathrm{Exp}(X)\,\mathrm{Exp}(J_\ell(X)\,\delta X), \qquad \mathrm{Exp}(X + \delta X) \approx \mathrm{Exp}(J_r(X)\,\delta X)\,\mathrm{Exp}(X),$$

where the approximation is understood in the sense of neglecting higher-order terms in $\|\delta X\|$.

Adjoint for $\mathrm{SE}(d)$

For the special Euclidean group

$$\mathrm{SE}(d) = \left\{ \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix} \;\middle|\; R \in \mathrm{SO}(d),\ t \in \mathbb{R}^d \right\},$$

the Lie algebra is

$$\mathfrak{se}(d) = \left\{ \begin{pmatrix} \Omega & v \\ 0 & 0 \end{pmatrix} \;\middle|\; \Omega \in \mathfrak{so}(d),\ v \in \mathbb{R}^d \right\}.$$

Since $\mathrm{SE}(d) \subset \mathrm{GL}(d+1, \mathbb{R})$ is a matrix Lie group, its adjoint action is given by conjugation, as stated in equation (2.7). Hence, for any $T \in \mathrm{SE}(d)$ and $X \in \mathfrak{se}(d)$, one has $\mathrm{Ad}_T(X) = TXT^{-1}$. In particular, letting

$$T = \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix} \in \mathrm{SE}(d), \qquad X = \begin{pmatrix} \Omega & v \\ 0 & 0 \end{pmatrix} \in \mathfrak{se}(d),$$

with $R \in \mathrm{SO}(d)$, $\Omega \in \mathfrak{so}(d)$ and $t, v \in \mathbb{R}^d$, and using $R^{-1} = R^\top$, we get

$$\mathrm{Ad}_T(X) = \begin{pmatrix} R\Omega R^\top & -R\Omega R^\top t + Rv \\ 0 & 0 \end{pmatrix}. \tag{2.11}$$

Equivalently, identifying $\mathfrak{se}(d) \simeq \mathfrak{so}(d) \oplus \mathbb{R}^d$, the adjoint action can be written as

$$\mathrm{Ad}_{(R,t)}(\Omega, v) = \big(R\Omega R^\top,\ -R\Omega R^\top t + Rv\big).$$

Specialization to $\mathrm{SE}(2)$ (hat–vee representation)

In dimension 2, the Lie algebra $\mathfrak{so}(2)$ is one-dimensional and can be identified with $\mathbb{R}$ via the hat–vee isomorphism

$$\omega \;\longleftrightarrow\; \widehat{\omega} = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}.$$

For any rotation $R \in \mathrm{SO}(2)$, one has

$$R\,\widehat{\omega}\,R^\top = \widehat{\omega},$$

moreover, for $p \in \mathbb{R}^2$,

$$-\,\widehat{\omega}\,p = \omega \begin{pmatrix} -p_2 \\ p_1 \end{pmatrix} = \widehat{p}\,\omega, \qquad \widehat{p} = \begin{pmatrix} 0 & -p_2 \\ p_1 & 0 \end{pmatrix}.$$

Hence, for planar twists $\xi = (\omega, v) \in \mathbb{R} \times \mathbb{R}^2$, the adjoint action of $\mathrm{SE}(2)$ reads

$$\mathrm{Ad}_{(R,p)}(\omega, v) = \begin{pmatrix} 1 & 0 \\ \widehat{p} & R \end{pmatrix} \begin{pmatrix} \omega \\ v \end{pmatrix} = \begin{pmatrix} \omega \\ \widehat{p}\,\omega + Rv \end{pmatrix}.$$

Left and right Jacobians of SE(2) For $d = 2$, write $\xi = (\omega, v) \in \mathbb{R}^3$ and

$$\hat{\xi} = \begin{pmatrix} \hat{\omega} & v \\ 0 & 0 \end{pmatrix} \in \mathfrak{se}(2).$$

The exponential map on $\mathfrak{se}(2)$ admits the closed-form expression

$$\mathrm{Exp}(\hat{\xi}) = \begin{pmatrix} \mathrm{Exp}(\hat{\omega}) & J_\ell(\omega)\, v \\ 0 & 1 \end{pmatrix},$$

where $\mathrm{Exp}(\hat{\omega})$ is the planar rotation matrix of angle ω and $J_\ell(\omega)$ is the left Jacobian of $\mathfrak{so}(2)$, given by

$$J_\ell(\omega) = \begin{pmatrix} \dfrac{\sin\omega}{\omega} & -\dfrac{1-\cos\omega}{\omega} \\ \dfrac{1-\cos\omega}{\omega} & \dfrac{\sin\omega}{\omega} \end{pmatrix}, \qquad \omega \neq 0,$$

with $J_\ell(0) = I_2$. The right Jacobian satisfies $J_r(\omega) = J_\ell(-\omega)$.

Specialization to SE(3) (hat–vee representation)

In dimension 3, using the hat–vee isomorphism $\Omega = \widehat{\omega}$ between $\mathfrak{so}(3)$ and $\mathbb{R}^3$, we have

$$R\Omega R^\top = \widehat{R\omega}, \qquad -R\Omega R^\top p = -\widehat{R\omega}\, p = p \times (R\omega) = \widehat{p}\, R\omega.$$

Thus for twists $\xi = (\omega, v) \in \mathbb{R}^3 \times \mathbb{R}^3$,

$$\mathrm{Ad}_{(R,p)}(\omega, v) = \begin{pmatrix} R & 0 \\ \widehat{p}\, R & R \end{pmatrix} \begin{pmatrix} \omega \\ v \end{pmatrix} = \begin{pmatrix} R\omega \\ \widehat{p}\, R\omega + Rv \end{pmatrix}.$$

Left and right Jacobians of SE(3) For $d = 3$, write $\xi = (\omega, v) \in \mathbb{R}^6$ and

$$\hat{\xi} = \begin{pmatrix} \hat{\omega} & v \\ 0 & 0 \end{pmatrix} \in \mathfrak{se}(3).$$

The exponential map on $\mathfrak{se}(3)$ can be written as

$$\mathrm{Exp}(\hat{\xi}) = \begin{pmatrix} \mathrm{Exp}(\hat{\omega}) & J_\ell(\omega)\, v \\ 0 & 1 \end{pmatrix},$$

where $J_\ell(\omega)$ is the left Jacobian of $\mathfrak{so}(3)$. A classical computation based on the series expansion of the matrix exponential and on Rodrigues' formula for rotations shows that

$$J_\ell(\omega) = I - \frac{1-\cos\|\omega\|}{\|\omega\|^2}\hat{\omega} + \frac{\|\omega\| - \sin\|\omega\|}{\|\omega\|^3}\hat{\omega}^2,$$

for $\omega \neq 0$, while $J_\ell(0) = I$. The right Jacobian satisfies $J_r(\omega) = J_\ell(-\omega)$. We refer to [12, 30] for a detailed derivation of these closed-form expressions.
These Jacobians are essential in practical computations involving small perturbations on SE(3), such as pose optimization, uncertainty propagation, and control linearization, because they map between variations in the algebra and variations in the group in a consistent, first-order manner.

Chapter 3

Statistical Estimation and the Cramér–Rao Bound

In this chapter we recall some basic notions of parametric estimation theory and the Cramér–Rao inequality, in both Euclidean spaces and Lie group settings. The exposition follows classical references such as Kay [14] and Amari–Nagaoka [13].

3.1 Statistical Estimators

Let $\theta \in \Theta \subseteq \mathbb{R}^l$ be an unknown parameter vector, and let X be a random vector taking values in $\mathcal{X} \subseteq \mathbb{R}^n$ whose probability density (or mass) function depends on θ, denoted by $p(x|\theta)$.

Definition 3.1.1 (Sample and estimator)**.** Let $(X_1, \ldots, X_n)$ be independent and identically distributed (*i.i.d.*) random vectors with common density $p(\cdot \mid \theta)$. A realization

$$\mathcal{D} = \{x_1, x_2, \ldots, x_n\},$$

with $x_i \in \mathcal{X}$, is called a sample of size n.
An **estimator** of θ is a measurable function

$$T : \mathcal{X}^n \to \Theta,$$

and

$$\hat{\theta} = T(x_1, \ldots, x_n)$$

is called an estimate of θ.

Definition 3.1.2 (Unbiased estimator)**.** An estimator $\hat{\theta}$ is said to be **unbiased** if

$$\mathbb{E}_\theta[\hat{\theta}] = \theta \qquad \forall\, \theta \in \Theta.$$

Otherwise, it is **biased**, and its bias is defined as

$$\mathrm{Bias}(\hat{\theta}) = \mathbb{E}_\theta[\hat{\theta}] - \theta.$$

Definition 3.1.3 (Covariance matrix)**.** The **covariance matrix** of an estimator $\hat{\theta}$ is

$$\mathrm{Cov}_\theta(\hat{\theta}) = \mathbb{E}_\theta\Big[(\hat{\theta} - \mathbb{E}_\theta[\hat{\theta}])(\hat{\theta} - \mathbb{E}_\theta[\hat{\theta}])^\top\Big] \in \mathbb{R}^{l \times l}.$$

For unbiased estimators, this simplifies to

$$\mathrm{Cov}_\theta(\hat{\theta}) = \mathbb{E}_\theta[(\hat{\theta} - \theta)(\hat{\theta} - \theta)^\top].$$

3.2 Maximum Likelihood Estimation (MLE)

Definition 3.2.1 (Likelihood and log-likelihood)**.** Let $\mathcal{D} = \{x_1, \ldots, x_n\}$ be observed data generated from a statistical model with joint density (or mass function)

$$p(x_1, \ldots, x_n \mid \theta).$$

The **likelihood function** and the **log-likelihood function** of the parameter θ given the data are defined as

$$L(\theta \mid \mathcal{D}) := p(x_1, \ldots, x_n \mid \theta), \qquad \mathcal{L}(\theta \mid \mathcal{D}) := \log L(\theta \mid \mathcal{D}).$$

If the observations $x_1, \ldots, x_n$ are independent, the joint density factorizes as

$$p(x_1, \ldots, x_n \mid \theta) = \prod_{i=1}^{n} p_i(x_i \mid \theta),$$

where the functions $p_i(\cdot \mid \theta)$ may in general differ from one another. In this case, the likelihood and log-likelihood become

$$L(\theta \mid \mathcal{D}) = \prod_{i=1}^{n} p_i(x_i \mid \theta), \qquad \mathcal{L}(\theta \mid \mathcal{D}) = \sum_{i=1}^{n} \log p_i(x_i \mid \theta).$$

In the special case where the observations are *i.i.d.* (independent and identically distributed), all densities coincide, $p_i(x \mid \theta) = p(x \mid \theta)$, leading to the simplified form

$$\mathcal{L}(\theta \mid \mathcal{D}) = \sum_{i=1}^{n} \log p(x_i \mid \theta)$$

Definition 3.2.2 (Maximum Likelihood Estimator)**.** The **Maximum Likelihood Estimator (MLE)** of θ is defined as

$$\hat{\theta}_{\text{MLE}} = \arg\max_{\theta \in \Theta} \mathcal{L}(\theta \mid \mathcal{D}).$$

If the loglikelihood is differentiable it satisfies

$$\nabla_\theta \mathcal{L}(\hat{\theta}_{\text{MLE}}) = 0.$$

3.3 The Fisher Information Matrix

Definition 3.3.1 (Fisher Information Matrix)**.** Let X be a random vector with probability density $p(x; \theta)$. The **Fisher Information Matrix** (FIM) is defined as

$$\mathbf{F}(\theta) = \mathbb{E}_\theta\left[(\nabla_\theta \log p(x \mid \theta))\,(\nabla_\theta \log p(x \mid \theta))^\top\right] \in \mathbb{R}^{l \times l}.$$

When differentiation and expectation can be interchanged,

$$\mathbf{F}(\theta) = -\mathbb{E}_\theta\left[\nabla^2_\theta \log p(X \mid \theta)\right].$$

Remark 3.3.2 (Interpretation)**.** $\mathbf{F}(\theta)$ measures the curvature of the log-likelihood around its maximum. A Fisher information with larger values implies sharper curvature and therefore greater precision in estimating θ.

Proposition 3.3.3 (Additivity)**.** For *i.i.d.* samples $X_1, \ldots, X_n$ with common density $p(\cdot \mid \theta)$, the Fisher information satisfies

$$\mathbf{F}_n(\theta) = n\,\mathbf{F}_1(\theta),$$

where $\mathbf{F}_1(\theta)$ denotes the Fisher information contributed by a single observation.

Proposition 3.3.4 (Asymptotic properties of the MLE)**.** As the number of samples n goes to infinity, and under standard regularity conditions, the Maximum Likelihood Estimator $\hat{\theta}_{\mathrm{MLE}}$ satisfies:

- **Consistency:** $\hat{\theta}_{\mathrm{MLE}} \to \theta$ in probability;
- **Asymptotic normality:**

  $$\sqrt{n}\,(\hat{\theta}_{\mathrm{MLE}} - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}\big(0,\, \mathbf{F}(\theta)^{-1}\big)\,,$$

  with $\mathbf{F}(\theta)$ denoting the Fisher information matrix when it is invertible.
- **Asymptotic efficiency:** it achieves the Cramér–Rao lower bound, whose meaning will be discussed later in next section.

3.4 The Cramér–Rao Inequality

Theorem 3.4.1 (Cramér–Rao Inequality)**.** Let $\hat{\theta}$ be an unbiased estimator of θ, and assume that the Fisher Information Matrix $\mathbf{F}(\theta)$ is invertible. Then, under standard regularity conditions,

$$\mathrm{Cov}_\theta(\hat{\theta}) \succeq \mathbf{F}(\theta)^{-1}, \tag{3.1}$$

that is,

$$\mathrm{Cov}_\theta(\hat{\theta}) - \mathbf{F}(\theta)^{-1} \text{ is positive semidefinite.}$$

Proof sketch. Let $s(X;\theta) = \nabla_\theta \log p(X;\theta)$ denote the score function, which satisfies $\mathbb{E}_\theta[s(X;\theta)] = 0$. Differentiating $\mathbb{E}_\theta[\hat{\theta}] = \theta$ and applying the multivariate Cauchy–Schwarz inequality yields

$$\mathrm{Cov}_\theta(\hat{\theta}) \succeq \mathbf{F}(\theta)^{-1}.$$

We refer to [13, 14] for a complete proof. □

The Cramér–Rao inequality provides a fundamental lower bound on the covariance of any unbiased estimator, thus quantifying the best achievable estimation accuracy given a statistical model. It serves as a benchmark to assess the efficiency of practical estimators and to evaluate how much information the observations carry about the unknown parameters.

3.5 Estimation on Lie Groups

In many applications, the unknown parameter the unknown parameter belongs to a Lie group, not to a Euclidean space. For instance, in our inference problem the parameter to be estimated will belong to $\mathrm{SE}(d)^n$.
A detailed treatment of estimation theory on Lie groups and manifolds can be found in [13, 32, 33].

Let G be a Lie group of dimension l, with identity element e, Lie algebra $\mathfrak{g}$, and exponential and logarithm maps

$$\mathrm{Exp} : \mathfrak{g} \to G, \qquad \mathrm{Log} : G \supset U \to \mathfrak{g},$$

defined locally around the identity. Let X be a random variable with density $p(x|\theta)$ depending on an unknown parameter $\theta \in G$.

Definition 3.5.1 (Estimator on a Lie group)**.** An **estimator** of $\theta \in G$ is a measurable map

$$\hat{\theta} : \mathcal{X}^n \to G,$$

which maps the collection of observed data to a single element of the group.

Unlike the Euclidean case, the expression

$$\mathbb{E}[\hat{\theta}] - \theta$$

is not well defined on a Lie group, since subtraction is not well defined in the setting of Lie group theory. As a consequence, the classical definition of unbiasedness must be replaced by a more general one.

Definition 3.5.2 (Generalized bias)**.** Let $\hat{\theta}$ be an estimator of $\theta \in G$. The **bias** of $\hat{\theta}$ is defined as

$$\mathrm{Bias}(\hat{\theta}) = \mathbb{E}_\theta\Big[\mathrm{Log}\big(\theta^{-1}\hat{\theta}\big)\Big] \ \in \ \mathfrak{g},$$

where $\mathrm{Log}(\cdot)$ denotes the logarithm map of the group.

An estimator is said to be **unbiased** if

$$\mathbb{E}_\theta\Big[\mathrm{Log}\big(\theta^{-1}\hat{\theta}\big)\Big] = 0.$$

Remark 3.5.3. This definition measures the estimation error in the tangent space $T_\theta G$, identified with the Lie algebra $\mathfrak{g}$ via left translation. In the Euclidean case $G = \mathbb{R}^d$, the group operation is addition and $\mathrm{Log}(\theta^{-1}\hat{\theta}) = \hat{\theta} - \theta$, so unbiasedness reduces to the classical notion.

Definition 3.5.4 (Covariance matrix on a Lie group)**.** The **covariance matrix** of an estimator $\hat{\theta}$ of an element $\theta \in G$ is defined as

$$\mathrm{Cov}_\theta(\hat{\theta}) = \mathbb{E}_\theta\Big[\mathrm{Log}\big(\theta^{-1}\hat{\theta}\big)\mathrm{Log}\big(\theta^{-1}\hat{\theta}\big)^\top\Big] \ \in \ \mathbb{R}^{l\times l},$$

where $\mathrm{Log}(\cdot)$ denotes the logarithm map of the group.

Maximum Likelihood Estimation on Lie Groups

In the Euclidean setting, the MLE is defined as the maximizer of the log-likelihood over $\Theta \subseteq \mathbb{R}^l$ and satisfies a first-order optimality condition $\nabla_\theta \mathcal{L}(\hat{\theta}_{\mathrm{MLE}}) = 0$. When the parameter space is a Lie group G, the definition remains the same in spirit, but all derivatives must be understood intrinsically on G.

Definition 3.5.5 (Log-likelihood on a Lie group)**.** Let $\theta \in G$ and let $X_1, \ldots, X_n$ be *i.i.d.* with density $p(\cdot \mid \theta)$. The log-likelihood is the smooth function

$$\mathcal{L}(\theta|\mathcal{D}) = \sum_{i=1}^{n} \log p(x_i \mid \theta), \qquad \theta \in G.$$

Definition 3.5.6 (Maximum Likelihood Estimator on G)**.** An **MLE on a Lie group** is any maximizer

$$\hat{\theta}_{\mathrm{MLE}} = \arg\max_{\theta \in G} \mathcal{L}(\theta|\mathcal{D}).$$

To write first-order optimality conditions, we consider variations of the parameter θ along smooth curves on the group. Since the tangent space $T_\theta G$ can be identified with the Lie algebra $\mathfrak{g}$ via left translation, any tangent direction at θ can be written as $(dL_\theta)_e(\xi)$ for some $\xi \in \mathfrak{g}$. The exponential map provides a natural way to generate such variations: for any $\xi \in \mathfrak{g}$, the curve

$$t \longmapsto \theta \operatorname{Exp}(t\xi)$$

is obtained by left-translating the one-parameter subgroup $\operatorname{Exp}(t\xi)$. This curve passes through θ at $t = 0$ and has tangent vector

$$\left.\frac{d}{dt}\right|_{t=0} \theta \operatorname{Exp}(t\xi) = (dL_\theta)_e(\xi) \in T_\theta G,$$

thus representing an arbitrary infinitesimal variation of θ.

Definition 3.5.7 (Riemannian gradient and score)**.** Fix an inner product $\langle \cdot, \cdot \rangle$ on $\mathfrak{g}$, inducing a left-invariant Riemannian metric on G. The **directional derivative** of $\mathcal{L}$ at θ in direction $\xi \in \mathfrak{g}$ is

$$\mathrm{D}\mathcal{L}(\theta)[\xi] := \left.\frac{d}{dt} \mathcal{L}(\theta \operatorname{Exp}(t\xi)|\mathcal{D})\right|_{t=0}.$$

The **Riemannian gradient** $\operatorname{grad} \mathcal{L}(\theta) \in T_\theta G$ is the unique tangent vector such that

$$\left\langle \theta^{-1} \operatorname{grad} \mathcal{L}(\theta),\, \xi \right\rangle = \mathrm{D}\mathcal{L}(\theta)[\xi], \qquad \forall\, \xi \in \mathfrak{g}.$$

Equivalently, defining the **score** $S_\theta(X) \in \mathfrak{g}$ by

$$\left\langle S_\theta(X), \xi \right\rangle = \left.\frac{d}{dt} \log p(X \mid \theta \operatorname{Exp}(t\xi))\right|_{t=0}, \qquad \forall\, \xi \in \mathfrak{g}, \tag{3.2}$$

where $\langle \cdot, \cdot \rangle$ denotes an inner product on the Lie algebra $\mathfrak{g}$ and $\mathrm{Exp} : \mathfrak{g} \to G$ is the exponential map we have that the directional derivative of the log-likelihood can be written as

$$\mathrm{D}\mathcal{L}(\theta)[\xi] = \sum_{i=1}^{N} \big\langle S_\theta(X_i), \xi \big\rangle = \Big\langle \sum_{i=1}^{N} S_\theta(X_i),\, \xi \Big\rangle. \tag{3.3}$$

By the definition of the Riemannian gradient associated with the chosen inner product, this identity implies that

$$\theta^{-1}\mathrm{grad}\, \mathcal{L}(\theta) = \sum_{i=1}^{N} S_\theta(X_i) \;\in\; \mathfrak{g}. \tag{3.4}$$

Remark 3.5.8 (Euclidean case)**.** When the parameter space is the Euclidean space $G = \mathbb{R}^l$ endowed with the additive group structure, the intrinsic definition of the score introduced above reduces to the classical one. Indeed, in this case the Lie algebra is $\mathfrak{g} = \mathbb{R}^l$ and the exponential map coincides with the identity, so that

$$\theta \,\mathrm{Exp}(t\xi) = \theta + t\xi.$$

Therefore, for any $\xi \in \mathbb{R}^l$,

$$\frac{d}{dt}\bigg|_{t=0} \log p(X \mid \theta + t\xi) = \big\langle \mathrm{grad}_\theta \log p(X;\theta),\, \xi \big\rangle,$$

by the definition of the Euclidean gradient. Since this identity holds for all ξ, it follows that the intrinsic score $S_\theta(X)$ coincides with the classical score

$$S_\theta(X) = s(X;\theta) := \nabla_\theta \log p(X;\theta).$$

Remark 3.5.9 (Optimization on manifolds)**.** In practice, computing $\hat{\theta}_{\mathrm{MLE}}$ on G is an optimization problem on a manifold. Typical algorithms are Riemannian gradient descent and Riemannian trust-region / Newton methods, which iteratively update $\theta_{k+1} = \theta_k \mathrm{Exp}(\alpha_k \eta_k)$ where $\eta_k \in \mathfrak{g}$ is (a function of) the Riemannian gradient at θ_k.

Fisher Information Matrix on a Lie group

Let $\{X_i\}_{i=1}^n$ be *i.i.d.* random variables with likelihood $p(x \mid \theta)$, where $\theta \in G$ and G is a Lie group. Assume that $p(x \mid \theta)$ is smooth with respect to θ.

Definition 3.5.10 (Fisher Information Matrix on a Lie group)**.** The **Fisher Information Matrix** (FIM) at $\theta \in G$ is the bilinear form on $\mathfrak{g}$ defined by

$$\mathbf{F}(\theta) = \mathbb{E}_\theta\big[S_\theta(X)\, S_\theta(X)^\top\big] \;\in\; \mathbb{R}^{l \times l}, \tag{3.5}$$

where a basis of $\mathfrak{g}$ has been fixed and the score is defined as in (3.2) .

Remark 3.5.11 (Euclidean case)**.** When $G = \mathbb{R}^l$, the Lie algebra coincides with the parameter space and the exponential map reduces to addition. In this case, the intrinsic Fisher information matrix reduces exactly to the classical Fisher information matrix.

Consequences for estimation theory

The lack of a global vector space structure leads to several important differences with respect to the Euclidean theory:

- Unbiasedness is *local* and defined via the logarithmic map;
- Covariance is computed in the tangent space rather than on the group itself;
- The Fisher Information Matrix becomes a bilinear form on $\mathfrak{g}$;
- The Cramér–Rao bound involves curvature terms when the manifold is non-flat.

In particular, for Lie groups endowed with a Riemannian metric, one can derive an **intrinsic Cramér–Rao bound** of the form

$$\mathrm{Cov}_\theta(\hat{\theta}) \succeq \mathbf{F}(\theta)^{-1} + \mathcal{O}(\|\mathrm{Cov}_\theta(\hat{\theta})\|^2), \tag{3.6}$$

where $\mathbf{F}(\theta)$ denotes the intrinsic Fisher information matrix. The additional higher-order terms reflect the curvature of the parameter space.

Remark 3.5.12 (Relation with Euclidean CRB)**.** When $G = \mathbb{R}^l$, the Lie algebra coincides with the parameter space, the curvature vanishes, and the intrinsic Cramér–Rao bound reduces exactly to the classical bound.

3.6 Singular or Ill-Conditioned Fisher Information Matrices

In the general setting, the Fisher information matrix may be singular (or even identically zero along some directions) whenever the data do not excite all degrees of freedom of the parameter, or when the model presents intrinsic indeterminacies and constraints. In such situations, the lack of invertibility is not merely a technical inconvenience: it signals that only a subspace of directions in the tangent space is statistically identifiable.

The standard practice to handle these degeneracies is to reformulate the estimation problem on a parameter space that explicitly encodes the available information. On manifolds, and in partciular on Lie groups, two broad geometric settings naturally capture this idea:

- **Riemannian submanifolds** (constraints): the true parameter is known *a priori* to lie in a submanifold $\bar{\mathcal{P}} \subset \mathcal{P}$ of some parent manifold $\mathcal{P}$, so that only tangent directions in $T_\theta\mathcal{P}$ are meaningful. This situation typically arises by *anchoring* (i.e., fixing) some of the unknown parameters, thereby restricting the admissible configurations to a lower-dimensional submanifold.
- **Riemannian quotient manifolds** (non-identifiability): the parameter is only identifiable up to an equivalence relation $\sim$, so the effective parameter space is the quotient $\bar{\mathcal{P}} = \mathcal{P}/\sim$, and directions tangent to equivalence classes carry no information.

In both cases, the Fisher information computed in the parent space may be singular because it contains directions that are either forbidden (submanifold constraints) or unidentifiable (quotient directions). Intrinsic Cramér–Rao bounds for submanifolds and quotient manifolds, due to N. Boumal [19], provide a unified framework valid even in the presence of singular Fisher information matrices under the hyphotesis of high SNR. In this work we will focus on the simplest forms (neglecting curvature) and we report the two main results.

Theorem 3.6.1 (Cramér–Rao bound on submanifolds)**.** Let $\bar{\mathcal{P}} \subset \mathcal{P}$ be a Riemannian submanifold of a parent manifold $\mathcal{P}$. Denote by

$$\Pi : T_\theta \mathcal{P} \to T_\theta \bar{\mathcal{P}}$$

the orthogonal projector onto the tangent space of the submanifold at θ. The Fisher Information Matrix $\mathbf{F}(\theta)$ is computed in the ambient space $T_\theta \mathcal{P}$ and may be singular.

For any unbiased estimator $\hat{\theta}$ defined on $\bar{\mathcal{P}}$ and in the small-error regime (neglecting curvature terms), the intrinsic Cramér–Rao bound takes the form

$$\mathrm{Cov}_\theta(\hat{\theta}) \succeq \big(\Pi\, \mathbf{F}(\theta)\, \Pi\big)^\dagger. \tag{3.7}$$

In particular, taking the trace on both sides yields

$$\mathrm{tr}\big(\mathrm{Cov}_\theta(\hat{\theta})\big) \geq \mathrm{tr}\Big(\big(\Pi\, \mathbf{F}(\theta)\, \Pi\big)^\dagger\Big), \tag{3.8}$$

which provides a scalar lower bound on the total estimation variance.

Here $A^\dagger$ denotes the **Moore–Penrose pseudoinverse**, i.e., the unique matrix satisfying the four Penrose conditions

$$AA^\dagger A = A, \qquad A^\dagger A A^\dagger = A^\dagger, \qquad (AA^\dagger)^\top = AA^\dagger, \qquad (A^\dagger A)^\top = A^\dagger A.$$

The pseudoinverse exists for any matrix and coincides with the usual inverse whenever A is invertible.

Theorem 3.6.2 (Cramér–Rao bound on quotient manifolds)**.** Let $\bar{\mathcal{P}} = \mathcal{P}/\!\sim$ be a Riemannian quotient manifold, where $\sim$ is an equivalence relation on the parent manifold $\mathcal{P}$. In this setting, the Fisher Information Matrix $\mathbf{F}(\theta)$ computed on the parent space is singular, since directions tangent to the equivalence classes are unidentifiable.

For any intrinsically unbiased estimator $\hat{\theta}$ defined on the quotient space and in the small-error regime (neglecting curvature terms), the intrinsic Cramér–Rao bound is expressed as

$$\mathrm{Cov}_\theta(\hat{\theta}) \succeq \mathbf{F}(\theta)^\dagger. \tag{3.9}$$

In particular, taking the trace on both sides yields the scalar bound

$$\mathrm{tr}\big(\mathrm{Cov}_\theta(\hat{\theta})\big) \geq \mathrm{tr}\big(\mathbf{F}(\theta)^\dagger\big), \tag{3.10}$$

which lower-bounds the total intrinsic estimation variance on the quotient manifold.

Remark 3.6.3. Since a Lie group G is in general not flat, exact formulations of the bound involve higher-order curvature correction terms. These terms arise from the non-Euclidean geometry of the parameter space and contribute only at second order in the estimation error. As a result, they are usually negligible in practical applications.

Chapter 4

A right-invariant model for relative measurements in $\mathrm{SE}(d)^n$

4.1 Graph Visualization of a Set of Relative Measurements on $\mathrm{SE}(d)^n$

We consider a directed graph $\mathcal{G} = (\mathcal{T}, \mathcal{E})$ composed of n vertices, whose underlying undirected graph is assumed to be connected. Each node is associated with a rigid-body transformation in $\mathrm{SE}(d)$, referred to as an absolute pose:

$$T_i \in \mathrm{SE}(d) \qquad i = 1, \ldots, n.$$

As described in detail in Chapter 2, an element of the Special Euclidean group $\mathrm{SE}(d)$ is given by a pair (R, t), where

$$R \in \mathrm{SO}(d), \qquad t \in \mathbb{R}^d,$$

representing a rotation and a translation in $\mathbb{R}^d$, respectively. In matrix form, such a transformation can be represented as a $(d+1) \times (d+1)$ homogeneous matrix:

$$T_i = \begin{pmatrix} R_i & t_i \\ 0_{1\times d} & 1 \end{pmatrix} \qquad \in \mathrm{GL}(d+1, \mathbb{R}).$$

Arcs of the graph encode relative pose measurements between pairs of nodes. It means that each arc $(i, j) \in \mathcal{E}$ is associated with a relative transformation

$$T_{ij} \in \mathrm{SE}(d),$$

which represents the transformation mapping coordinates from the reference frame of T_j to that of T_i. Since the composition of rigid body transformations belongs to $\mathrm{SE}(d)$, all such relative measurements are elements of $\mathrm{SE}(d)$ as well.

The relative transformations associated with opposite arc directions are related by inversion. In particular, reversing the orientation of an arc corresponds to replacing the relative transformation with its group inverse:

$$T_{ij} = T_{ji}^{-1}. \tag{4.1}$$

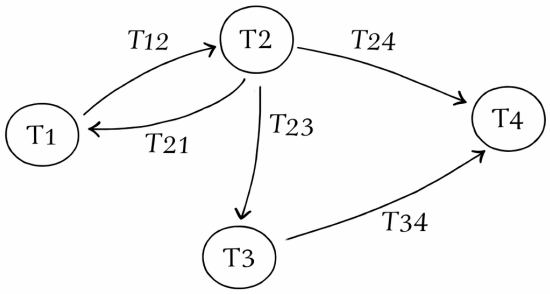


Figure 4.1: Example of a directed graph with nodes $T_i \in \mathrm{SE(d)}$ with their available arcs

4.2 Deterministic case: reconstruction of absolute poses from relative measurements

In the deterministic, noise-free case, we can evaluate the relative measurement as:

$$T_{ij} = T_i T_j^{-1},$$

which represents the pose of node i as seen from node j.

Pose reconstruction in a complete graph

In a complete graph, where all relative transformations T_{ij} are available, the reconstruction of the absolute poses starting from the relative ones is a solved problem. By choosing an hamiltonian directed path

$$T_{i_1} \to T_{i_2} \to \cdots \to T_{i_k},$$

and selecting $T_{i_1} = G$ as the absolute reference frame, it is possible to express all other poses recursively as

$$T_{i_k} = T^{-1}_{i_{k-1}i_k} T^{-1}_{i_{k-2}i_{k-1}} \cdots T^{-1}_{i_1 i_2} G = T_{i_k i_{k-1}} T_{i_{k-1} i_{k-2}} \cdots T_{i_2 i_1} G. \tag{4.2}$$

This formula shows that all absolute poses are uniquely determined once a global reference G is fixed, demonstrating algebraically the dependence of the entire configuration on a single global coordinate frame.

Extension to connected graphs

The same reasoning applies even if the graph is not complete, as long as the underlying undirected graph is connected. In this case one can always find a path between any

two nodes also in the directed case. Indeed, for every available relative transformation T_{ij},we can use the inverse relation (4.1). Therefore, it is always possible to find a path and apply the same reconstruction procedure along that path. Note that in this case it is not ensured that there exists an hamiltonian path in the graph, but one can use an arbitrary walk.

Disconnected graphs

If the graph has more connected components, the previous formula (4.2) holds independently for each component. Each connected component contributes one independent global reference frame, hence adding one degree of freedom to the global configuration. Consequently, the number of connected components in the graph coincides with the number of degrees of freedom of the system in $\mathrm{SE}(d)$.

4.3 The model in the stochastic case

Let us now consider the case where the relative measurements are affected by heteroskedastic noise and the graph is initially connected. We denote by $\mathcal{E}$ the set of all directed arcs of the graph. Let $\mathcal{T} = \{T_1, T_2, \dots, T_n\}$ denote the set of poses of interest, where each T_i represents a node of the graph. The measurement graph is therefore defined as the ordered pair

$$(\mathcal{T}, \mathcal{E}).$$

We define for every arc $(i, j) \in \mathcal{E}$ the noise ξ_{ij} as a random variable taking values in $\mathbb{R}^q$, where

$$q = \frac{d(d+1)}{2} = \dim(\mathrm{SE}(d)).$$

The noise is assumed to be normally distributed as

$$\xi_{ij} \sim \mathcal{N}(0, \Sigma_{ij}),$$

where $\Sigma_{ij} \in \mathbb{R}^{q \times q}$ is a *known* positive definite covariance matrix, possibly different for any directed arc $(i, j) \in \mathcal{E}$. We denote by $\Omega_{ij} = \Sigma_{ij}^{-1}$ the inverse of Σ_{ij}.
The errors ξ_{ij} are assumed to be independent for each directed arc of the graph. This independence assumption implies that the expectation of the product between two noise terms satisfies:

$$\mathbb{E}\big[\, \xi_{ij}\, \xi_{kl}^{\top} \,\big] = \begin{cases} \Sigma_{ij}, & \text{if } (i,j) = (k,l), \\ 0, & \text{if } (i,j) \neq (k,l). \end{cases}$$

In the model we are going to study, for every arc (i, j) in $\mathcal{E}$, the error perturbs the noise–free measurement according to the equation:

$$H_{ij} = \mathrm{Exp}(\widehat{\xi_{ij}})\, T_i T_j^{-1}, \tag{4.3}$$

where:

- the hat operator $\widehat{(\cdot)}$ maps a vector in $\mathbb{R}^q$ to the corresponding element of the Lie algebra $\mathfrak{se}(d)$ with the standard base (2.6) ;

- the operator $\mathrm{Exp}(\cdot)$ is the exponential map projecting an element of the Lie algebra $\mathfrak{se}(d)$ onto the group $\mathrm{SE}(d)$. As said, in the case of a matrix Lie group, this operation corresponds to the exponential of a matrix.

Since $\mathrm{Exp}(\widehat{\xi}_{ij}) \in \mathrm{SE}(d)$ and $T_i, T_j \in \mathrm{SE(d)}$, it follows that $H_{ij} \in \mathrm{SE(d)}$ as well. In the absence of error ($\xi_{ij} = 0$), we have:

$$H_{ij} = T_i T_j^{-1} = T_{ij}.$$

coherently with the deterministic case.

The perturbation model adopted in (4.3) assumes that measurement noise acts through left multiplication. This choice reflects a conceptual distinction between left- and right-multiplicative noise models.
When noise is introduced via left multiplication, the perturbation acts directly on the relative transformation being measured. In this case, the error is naturally expressed in the local reference frame of the sensing device and affects the relative pose independently of the unknown absolute poses. This modeling choice is well suited to applications such as camera-based measurements, where relative poses are estimated from image data and uncertainty is inherently associated with the observed relative motion.
Conversely, introducing noise through right multiplication corresponds to expressing the perturbation in the reference frame of the absolute poses. As a result, the effect of the noise becomes dependent on the (unknown) global configuration, causing the measurement uncertainty to be entangled with the absolute poses. This interpretation is less consistent with the nature of relative measurements in pose-graph problems.
For these reasons, left-multiplicative noise is adopted throughout this work in order to provide a more faithful model of camera-generated relative pose measurements. Nevertheless, there is no fundamental obstacle to developing an analogous formulation based on right-multiplicative noise, and a similar analysis could be carried out under this alternative assumption.

Proposition 4.3.1 (Right-invariance of the measurement model)**.** Consider the perturbed relative measurement

$$H_{ij} = \mathrm{Exp}(\widehat{\xi}_{ij})\, T_i T_j^{-1}, \qquad T_i, T_j \in \mathrm{SE}(d).$$

Then H_{ij} is invariant under a global right action, in the sense that replacing each pose T_k with $T_k G$ leaves the relative measurement unchanged.

Proof. For any $G \in \mathrm{SE}(d)$,

$$\mathrm{Exp}(\widehat{\xi}_{ij})(T_i G)(T_j G)^{-1} = \mathrm{Exp}(\widehat{\xi}_{ij})\, T_i G G^{-1} T_j^{-1} = \mathrm{Exp}(\widehat{\xi}_{ij})\, T_i T_j^{-1} = H_{ij},$$

which proves that the perturbed relative measurement is invariant under a global right action. □

The aim of the estimation problem is to recover the absolute poses $T_1, T_2, \ldots, T_n$ from the noisy relative measurements H_{ij} and from the probabilistic model of the errors ξ_{ij}. However, the solution is determined only up to a global transformation in $\mathrm{SE}(d)$, since relative measurements do not provide an absolute reference frame. This inherent gauge freedom introduces a structural ambiguity in the estimation problem, which will lead to a singular Fisher Information Matrix.

Proposition 4.3.2. In the same notation as before, if the noise $\widehat{\xi}_{ij}$ is small enough, we have

$$\xi_{ij} = \left[\mathrm{Log}\big(H_{ij}\, T_j\, T_i^{-1}\big)\right]^{\vee}, \qquad \forall\, (i,j) \in \mathcal{E}.$$

Proof. Starting from the measurement model

$$H_{ij} = \mathrm{Exp}(\widehat{\xi}_{ij})\, T_i\, T_j^{-1}, \qquad \forall\, (i,j) \in \mathcal{E},$$

we multiply both sides on the right by $T_j T_i^{-1}$, obtaining

$$H_{ij}\, T_j\, T_i^{-1} = \mathrm{Exp}(\widehat{\xi}_{ij})\, T_i\, T_j^{-1}\, T_j\, T_i^{-1}.$$

Since $T_j^{-1} T_j = I$ and $T_i T_i^{-1} = I$, this reduces to

$$H_{ij}\, T_j\, T_i^{-1} = \mathrm{Exp}(\widehat{\xi}_{ij}).$$

Assuming that the noise $\widehat{\xi}_{ij}$ is sufficiently small, the group element $\mathrm{Exp}(\widehat{\xi}_{ij})$ lies in a neighborhood of the identity. Therefore, the logarithmic map is well defined and can be applied to both sides, yielding

$$\mathrm{Log}\big(H_{ij}\, T_j\, T_i^{-1}\big) = \widehat{\xi}_{ij}.$$

Finally, applying the $\vee$ operator, which is the inverse of the $\widehat{(\cdot)}$ operator, we obtain

$$\xi_{ij} = \left[\mathrm{Log}\big(H_{ij}\, T_j\, T_i^{-1}\big)\right]^{\vee}.$$

□

Remark. Since $\xi_{ij} \sim \mathcal{N}(0, \Sigma_{ij})$ by assumption, it follows that

$$\left[\mathrm{Log}\big(H_{ij}\, T_j\, T_i^{-1}\big)\right]^{\vee} \sim \mathcal{N}(0, \Sigma_{ij}).$$

Accordingly, we denote by

$$r_{ij} := \left[\mathrm{Log}\big(H_{ij}\, T_j\, T_i^{-1}\big)\right]^{\vee} \in \mathbb{R}^q$$

the residual associated with the arc (i,j), which is distributed as a Gaussian random vector with zero mean and covariance matrix Σ_{ij}.
It is important to distinguish between the noise variable ξ_{ij} and the residual r_{ij}. The former represents the unobservable random noise appearing in the generative measurement model, whereas the latter is a deterministic quantity once the absolute poses T_i and T_j are fixed and can be directly computed from the observed measurement H_{ij}. When the estimated poses coincide with the true ones, the residual r_{ij} coincides with the corresponding noise realization ξ_{ij}.

4.4 The Log-Likelihood Function

Proposition 4.4.1 (Log-likelihood of the pose-graph model)**.** Let $\Theta = \{T_1, \dots, T_n\}$ denote the collection of unknown absolute poses, with $T_i \in \mathrm{SE}(d)$ for all $i = 1, \dots, n$. For each directed arc $(i, j) \in \mathcal{E}$, let $r_{ij}(\Theta) \in \mathbb{R}^q$ be the residuals associated to Θ. Assume that the residuals are independent across arcs and satisfy

$$r_{ij}(\Theta) \sim \mathcal{N}(0, \Sigma_{ij}), \qquad \Sigma_{ij} \in \mathbb{R}^{q \times q} \text{ positive definite},$$

and denote $\Omega_{ij} = \Sigma_{ij}^{-1}$. Then the log-likelihood of the measurements given Θ is

$$\mathcal{L}(\Theta) = -\frac{1}{2} \sum_{(i,j) \in \mathcal{E}} r_{ij}(\Theta)^\top \, \Omega_{ij} \, r_{ij}(\Theta) - \frac{1}{2} \sum_{(i,j) \in \mathcal{E}} \log \det\bigl(2\pi \Sigma_{ij}\bigr). \tag{4.4}$$

Equivalently,

$$\mathcal{L}(\Theta) = -\frac{1}{2} \sum_{(i,j) \in \mathcal{E}} r_{ij}(\Theta)^\top \, \Omega_{ij} \, r_{ij}(\Theta) + C,$$

where the constant term

$$C = -\frac{1}{2} \sum_{(i,j) \in \mathcal{E}} \log \det\bigl(2\pi \Sigma_{ij}\bigr)$$

does not depend on Θ.

Proof. Recall that if $\mathrm{X} \in \mathbb{R}^q$ is a Gaussian random vector distributed as $\mathrm{X} \sim \mathcal{N}(0, \Sigma)$ with $\Sigma \succ 0$, then its probability density function is

$$p(x) = \frac{1}{\sqrt{(2\pi)^q \det(\Sigma)}} \exp\left(-\frac{1}{2} x^\top \Sigma^{-1} x \right).$$

Taking the logarithm yields the log-density

$$\log p(x) = -\frac{1}{2} x^\top \Sigma^{-1} x - \frac{1}{2} \log\bigl((2\pi)^q \det(\Sigma)\bigr).$$

Since $\det\bigl((2\pi)\Sigma\bigr) = (2\pi)^q \det(\Sigma)$ for any $q \times q$ matrix Σ, the normalization term can equivalently be written as a single determinant. Applying this formula to each residual $r_{ij}(\Theta) \in \mathbb{R}^q$ with covariance Σ_{ij} yields

$$\log p\bigl(r_{ij}(\Theta)\bigr) = -\frac{1}{2} r_{ij}(\Theta)^\top \Sigma_{ij}^{-1} r_{ij}(\Theta) - \frac{1}{2} \log \det\bigl((2\pi)\Sigma_{ij}\bigr).$$

Since the residuals are assumed independent over all arcs $(i, j) \in \mathcal{E}$, the joint likelihood factorizes as

$$p\bigl(\{r_{ij}(\Theta)\}_{(i,j) \in \mathcal{E}}\bigr) = \prod_{(i,j) \in \mathcal{E}} p\bigl(r_{ij}(\Theta)\bigr),$$

and hence the log-likelihood is given by the sum of the individual log-densities:

$$\mathcal{L}(\Theta) = \sum_{(i,j) \in \mathcal{E}} \log p\bigl(r_{ij}(\Theta)\bigr).$$

Substituting the expression above and setting $\Omega_{ij} = \Sigma_{ij}^{-1}$ yields (4.4), concluding the proof.

□

This formulation will serve as the basis for the problem of estimating the absolute poses by maximizing the log-likelihood, or equivalently, by minimizing the weighted quadratic term associated with the residuals.

Maximum Likelihood Estimation

The Maximum Likelihood Estimate (MLE) of the parameters is defined as:

$$\Theta^{\mathrm{MLE}} = \arg\max_{\Theta} \mathcal{L}(\Theta).$$

However, also the log-likelihood remains unchanged under any common right multiplication of all poses by an arbitrary transformation $G \in \mathrm{SE}(d)$. As a consequence, if $\{T_i^{\mathrm{MLE}}\}$ is an optimal configuration, since the relative measurement graph is considered connected, then every set of poses of the form

$$\{T_i^{\mathrm{MLE}} G \,|\, G \in \mathrm{SE}(d)\}$$

is also optimal. Therefore, the MLE of the absolute poses is not unique, since the solution is defined only up to a global rigid-body transformation G.
Nevertheless, this ambiguity does not affect the estimation of the *relative poses*, because all configurations related by a global transformation G produce identical relative transformations $T_i T_j^{-1}$. In particular, we can define the MLE of the relative transformations as

$$\left(T_i T_j^{-1}\right)^{\mathrm{MLE}} := T_i^{\mathrm{MLE}} \left(T_j^{\mathrm{MLE}}\right)^{-1}.$$

This definition is well-posed, since for any global transformation $G \in \mathrm{SE}(d)$ we have

$$\left(T_i^{\mathrm{MLE}} G\right) \left(T_j^{\mathrm{MLE}} G\right)^{-1} = T_i^{\mathrm{MLE}} G G^{-1} \left(T_j^{\mathrm{MLE}}\right)^{-1} = T_i^{\mathrm{MLE}} \left(T_j^{\mathrm{MLE}}\right)^{-1}.$$

Therefore, the estimates of the relative poses are invariant with respect to the choice of the global reference frame, and are thus defined in a consistent and unique manner, despite the global ambiguity in the absolute poses. Furthermore, once an optimal configuration of the absolute poses has been obtained, this formulation enables the reconstruction of consistent relative transformations between any pair of nodes in the graph, including those for which no direct measurement H_{ij} was originally available.

4.5 Gradient of the Log-Likelihood

From score to local coordinates. In Section (3.5), the Fisher Information Matrix has been defined on a Lie group in terms of the score $S_\theta(X) \in \mathfrak{g}$.
In order to derive an explicit and computable expression of the Fisher Information Matrix for the considered estimation problem, we express the Riemannian gradient and the associated score in local coordinates of the tangent space induced by right-invariant exponential perturbations, in accordance with the right-invariance of the model. More precisely, for each parameter $\theta \in G$, variations are represented as

$$\theta \;\mapsto\; \theta \,\mathrm{Exp}(\widehat{\delta}), \qquad \widehat{\delta} \in \mathfrak{g}.$$

With this choice, the intrinsic score $S_\theta(X) \in \mathfrak{g}$ admits a coordinate representation in a fixed basis of the Lie algebra, which we denote by $\nabla_\Theta \mathcal{L}(\Theta)$.
Accordingly, the Fisher Information Matrix can be written in matrix form as

$$\mathbf{F}(\Theta) = \mathbb{E}_\Theta\big[\nabla_\Theta \mathcal{L}(\Theta)\, \nabla_\Theta \mathcal{L}(\Theta)^\top\big],$$

where $\nabla_\Theta \mathcal{L}(\Theta)$ denotes the coordinate representation of the Riemannian gradient of the log-likelihood in the chosen tangent basis.
We remind that in our case:

$$\mathcal{L}(\Theta) = -\frac{1}{2} \sum_{(i,j)\in\mathcal{E}} r_{ij}(\Theta)^\top\, \Omega_{ij}\, r_{ij}(\Theta) + C,$$

where $r_{ij}(\Theta) = \big[\mathrm{Log}\big(H_{ij} T_j T_i^{-1}\big)\big]^\vee \in \mathbb{R}^q$, $\Omega_{ij} = \Sigma_{ij}^{-1}$, and C is constant with respect to the set of parameters Θ.

Gradient of the Log-Likelihood

For convenience, for each arc $(i,j) \in \mathcal{E}$, define the term:

$$\mathcal{L}_{ij}(\Theta) := r_{ij}(\Theta)^\top\, \Omega_{ij}\, r_{ij}(\Theta),$$

so that:

$$\mathcal{L}(\Theta) = -\frac{1}{2} \sum_{(i,j)\in\mathcal{E}} \mathcal{L}_{ij}(\Theta) + C.$$

To compute $\nabla_\Theta \mathcal{L}(\Theta)$ we exploit the linearity of the gradient:

$$\nabla_\Theta \mathcal{L}(\Theta) = -\frac{1}{2} \sum_{(i,j)\in\mathcal{E}} \nabla_\Theta \mathcal{L}_{ij}(\Theta).$$

Proposition 4.5.1 (Gradient of a weighted residual on $\mathrm{SE}(d)$)**.** Let $r_{ij} : \mathrm{SE}(d) \times \mathrm{SE}(d) \to \mathbb{R}^q$ be a differentiable residual and let $\Omega_{ij} \in \mathbb{R}^{q\times q}$ be symmetric positive definite. Consider right-invariant perturbations $T_k(\delta_k) = T_k \,\mathrm{Exp}(\widehat{\delta_k})$, $k \in \{i,j\}$, and define

$$\phi(\delta_i, \delta_j) = \tilde{r}_{ij}(\delta_i, \delta_j)^\top \Omega_{ij}\, \tilde{r}_{ij}(\delta_i, \delta_j), \qquad \delta_i, \delta_j \in \mathbb{R}^q.$$

The gradient of ϕ at $(\delta_i, \delta_j) = (0,0)$ has components

$$\frac{\partial \phi}{\partial \delta_i}(0,0) = 2\, J_i^\top\, \Omega_{ij}\, r_{ij}, \qquad \frac{\partial \phi}{\partial \delta_j}(0,0) = 2\, J_j^\top\, \Omega_{ij}\, r_{ij},$$

where

$$J_i := \left.\frac{\partial \tilde{r}_{ij}}{\partial \delta_i}\right|_{(0,0)}, \qquad J_j := \left.\frac{\partial \tilde{r}_{ij}}{\partial \delta_j}\right|_{(0,0)}.$$

Equivalently, letting $\delta = [\delta_i^\top\ \delta_j^\top]^\top$ and $J_{r_{ij}} := [\, J_i \ \ J_j\,]$, the Euclidean gradient of ϕ at $(0,0)$ is

$$\nabla \phi(0,0) = 2\, J_{r_{ij}}^\top\, \Omega_{ij}\, r_{ij}.$$

The gradient of the log-likelihood with respect to the parameter vector Θ is given by

$$\nabla_\Theta \mathcal{L}(\Theta) = -\sum_{(i,j)\in\mathcal{E}} J_{r_{ij}}(\Theta)^\top\, \Omega_{ij}\, r_{ij}(\Theta). \tag{4.5}$$

Proof. Right-invariant exponential perturbations provide local Euclidean coordinates on the tangent space, so that $\phi(\delta_i, \delta_j)$ is a smooth real-valued function of the variables $(\delta_i, \delta_j) \in \mathbb{R}^{2q}$.

Applying the chain rule to the quadratic form

$$\phi(\delta_i, \delta_j) = \tilde{r}_{ij}(\delta_i, \delta_j)^\top \Omega_{ij}\, \tilde{r}_{ij}(\delta_i, \delta_j),$$

one obtains

$$\frac{\partial \phi}{\partial \delta_i} = 2 \left(\frac{\partial \tilde{r}_{ij}}{\partial \delta_i} \right)^\top \Omega_{ij}\, \tilde{r}_{ij}, \qquad \frac{\partial \phi}{\partial \delta_j} = 2 \left(\frac{\partial \tilde{r}_{ij}}{\partial \delta_j} \right)^\top \Omega_{ij}\, \tilde{r}_{ij}.$$

Evaluating the above expressions at $(\delta_i, \delta_j) = (0, 0)$ yields the desired result. □

4.6 Two expressions for the Fisher Information Matrix

Theorem 4.6.1 (Fisher Information for Pose Parameters in SE(d))**.** Let Θ denote the vector of parameters $\mathrm{T}_i \in \mathrm{SE}(d)$, where each one represents the pose of node i. Assume that the residuals $\{r_{ij}\}_{(i,j)\in\varepsilon}$ are independent, zero-mean random vectors with covariance matrices Σ_{ij}, and define $\Omega_{ij} = \Sigma_{ij}^{-1}$. Then, the Fisher Information Matrix associated with Θ is given by

$$\mathbf{F}(\Theta) = \sum_{(i,j)\in\varepsilon} J_{r_{ij}}(\Theta)^\top\, \Omega_{ij}\, J_{r_{ij}}(\Theta) \;\in\; \mathbb{R}^{nq\times nq}. \tag{4.6}$$

Proof. By definition, the Fisher Information Matrix is given by

$$\mathbf{F}(\Theta) = \mathbb{E}\big[\, \nabla_\Theta \mathcal{L}(\Theta) \nabla_\Theta \mathcal{L}(\Theta)^\top \,\big],$$

Substituting the explicit expression of the score $\nabla_\Theta \mathcal{L}(\Theta)$ yields

$$\mathbf{F}(\Theta) = \mathbb{E}\left[\Big(\sum_{(i,j)\in\varepsilon} J_{r_{ij}}^\top \Omega_{ij}\, r_{ij} \Big) \Big(\sum_{(k,\ell)\in\varepsilon} J_{r_{k\ell}}^\top \Omega_{k\ell}\, r_{k\ell} \Big)^\top \right].$$

Expanding the product, we obtain the double summation

$$\mathbf{F}(\Theta) = \sum_{(i,j)\in\varepsilon} \sum_{(k,\ell)\in\varepsilon} \mathbb{E}\Big[J_{r_{ij}}^\top\, \Omega_{ij}\, r_{ij}\, r_{k\ell}^\top\, \Omega_{k\ell}^\top\, J_{r_{k\ell}} \Big].$$

Since the information matrices $\Omega_{k\ell}$ are symmetric, $\Omega_{k\ell}^\top = \Omega_{k\ell}$, and the Jacobians $J_{r_{ij}}$ are deterministic once Θ is fixed, they can be taken outside the expectation, yielding

$$\mathbf{F}(\Theta) = \sum_{(i,j)\in\varepsilon} \sum_{(k,\ell)\in\varepsilon} J_{r_{ij}}^\top\, \Omega_{ij}\, \mathbb{E}\big[r_{ij}\, r_{k\ell}^\top \big]\; \Omega_{k\ell}\, J_{r_{k\ell}}.$$

By assumption, the residuals are independent and satisfy

$$\mathbb{E}[r_{ij}] = 0, \qquad \mathbb{E}\big[r_{ij}\, r_{k\ell}^\top \big] = \begin{cases} \Sigma_{ij}, & \text{if } (i,j) = (k,\ell), \\ 0, & \text{otherwise.} \end{cases}$$

As a consequence, all cross terms in the double summation vanish, and only the diagonal terms with $(i,j) = (k,\ell)$ remain. Therefore,

$$\mathbf{F}(\Theta) = \sum_{(i,j)\in\varepsilon} J_{r_{ij}}^{\top}\, \Omega_{ij}\, \Sigma_{ij}\, \Omega_{ij}\, J_{r_{ij}}.$$

Recalling that $\Omega_{ij} = \Sigma_{ij}^{-1}$, we have

$$\Omega_{ij}\, \Sigma_{ij}\, \Omega_{ij} = \Omega_{ij},$$

which yields the final expression

$$\mathbf{F}(\Theta) = \sum_{(i,j)\in\varepsilon} J_{r_{ij}}^{\top}\, \Omega_{ij}\, J_{r_{ij}}.$$

This concludes the proof. □

This expression coincides with the definition of the Fisher Information Matrix as the covariance of the Riemannian gradient, when the latter is represented in local tangent coordinates through exponential perturbations.

Sparse Structure and Closed Form of the Jacobians $J_{r_{ij}}(\Theta)$ by right perturbation

Recall that the residuals are defined as

$$r_{ij}(\Theta) \;=\; \big[\mathrm{Log}\big(H_{ij}\, T_j\, T_i^{-1}\big)\big]^{\vee} \in \mathbb{R}^q,$$

and depend only on the parameters of the two poses involved, T_i and T_j.
Since r_{ij} depends only on T_i and T_j, the Jacobian $J_{r_{ij}}(\Theta)$ is *sparse* and has nonzero contributions only in the blocks associated with T_i and T_j. If we arrange it in blocks with respect to the component $\{T_i\}_{i=1}^n$, we obtain:

$$J_{r_{ij}}(\Theta) = [\,0\ \cdots\ 0\ \underbrace{J_i}_{\text{(block } i\text{)}}\ 0\ \cdots\ 0\ \underbrace{J_j}_{\text{(block } j\text{)}}\ 0\ \cdots\ 0\,],$$

where each block is a $q \times q$ matrix.
We now derive closed-form expressions for the non-zero blocks of the Jacobians $J_{r_{ij}}(\Theta)$. To this end, we adopt *right-invariant* perturbations, consistent with the chosen pose parametrization:

$$T_k(\delta_k) \;=\; T_k\, \mathrm{Exp}(\widehat{\delta_k}), \qquad \delta_k \in \mathbb{R}^q.$$

Given the residual

$$r_{ij} = \mathrm{Log}(Z)^{\vee}, \qquad Z := H_{ij}\, T_j\, T_i^{-1} \in \mathrm{SE}(d),$$

we analyze how right-invariant infinitesimal perturbations on T_j and T_i affect Z.

Step 1: Apply right-invariant perturbations.

$$T_i(\delta_i) = T_i\, \mathrm{Exp}(\widehat{\delta_i}), \qquad T_j(\delta_j) = T_j\, \mathrm{Exp}(\widehat{\delta_j}).$$

Substituting into the definition of Z yields

$$Z(\delta) = H_{ij}\, T_j\, \mathrm{Exp}(\widehat{\delta_j})\, \big(T_i\, \mathrm{Exp}(\widehat{\delta_i})\big)^{-1}.$$

Remark on notation. Throughout this section, the vectors $\delta_k \in \mathbb{R}^q$ denote the coordinate representations of elements of the Lie algebra $\mathfrak{se}(d)$ with respect to a fixed basis, so that $\widehat{\delta_k} \in \mathfrak{se}(d)$. The adjoint operator $\mathrm{Ad}_T : \mathfrak{se}(d) \to \mathfrak{se}(d)$ is represented in coordinates with respect to a fixed basis of the Lie algebra $\mathfrak{se}(d)$ by the matrix $\mathbf{Ad}_T \in \mathbb{R}^{q\times q}$, defined by

$$\mathbf{Ad}_T\, \delta \;=\; \big(\mathrm{Ad}_T(\widehat{\delta})\big)^{\vee}.$$

Step 2: Invert the perturbed right factor. Using the identity

$$(T\,\mathrm{Exp}(\widehat{\delta}))^{-1} = \mathrm{Exp}(-\widehat{\delta})\,T^{-1},$$

we obtain

$$Z(\delta) = H_{ij}\,T_j\,\mathrm{Exp}(\widehat{\delta_j})\,\mathrm{Exp}(-\widehat{\delta_i})\,T_i^{-1}.$$

Step 3: Factor out the nominal value Z. Recall that $Z = H_{ij}\,T_j\,T_i^{-1}$. By inserting the identity $T_iT_i^{-1} = I$ between the two exponential terms, we can rewrite $Z(\delta)$ as

$$\begin{aligned} Z(\delta) &= H_{ij}\,T_j\,T_i^{-1}\,\big(T_i\,\mathrm{Exp}(\widehat{\delta_j})\,T_i^{-1}\big)\,\big(T_i\,\mathrm{Exp}(-\widehat{\delta_i})\,T_i^{-1}\big) \\ &= Z\;\cdot\;\big(T_i\,\mathrm{Exp}(\widehat{\delta_j})\,T_i^{-1}\big)\,\big(T_i\,\mathrm{Exp}(-\widehat{\delta_i})\,T_i^{-1}\big). \end{aligned}$$

Step 4: Use the adjoint action. For any $T \in \mathrm{SE}(d)$ and any $X \in \mathfrak{se}(d)$, the identity (2.9) holds:

$$T\,\mathrm{Exp}(X)\,T^{-1} = \mathrm{Exp}\big(\mathrm{Ad}_T\,X\big).$$

Applying this identity and using the fact that conjugation is a group automorphism, we obtain

$$T_i\,\mathrm{Exp}(\widehat{\delta_j})\,\mathrm{Exp}(-\widehat{\delta_i})\,T_i^{-1} = \mathrm{Exp}\big(\widehat{\mathbf{Ad}_{T_i}\delta_j}\big)\,\mathrm{Exp}\big(-\widehat{\mathbf{Ad}_{T_i}\delta_i}\big).$$

Step 5: First-order approximation (BCH). Using the first-order Baker–Campbell–Hausdorff approximation

$$\mathrm{Exp}(\widehat{a})\,\mathrm{Exp}(\widehat{b}) \approx \mathrm{Exp}(\widehat{a+b}),$$

we obtain

$$\mathrm{Exp}\big(\widehat{\mathbf{Ad}_{T_i}\delta_j}\big)\,\mathrm{Exp}\big(-\widehat{\mathbf{Ad}_{T_i}\delta_i}\big) \approx \mathrm{Exp}\big(\widehat{\mathbf{Ad}_{T_i}(\delta_j-\delta_i)}\big).$$

Resulting perturbation on Z. Defining the combined perturbation

$$\eta := \mathbf{Ad}_{T_i}\delta_j - \mathbf{Ad}_{T_i}\delta_i \in \mathbb{R}^q,$$

we obtain the first-order approximation

$$\boxed{Z(\delta) \;\approx\; Z\,\mathrm{Exp}(\widehat{\eta}), \qquad \eta = \mathbf{Ad}_{T_i}(\delta_j-\delta_i).}$$

Linearization of the residual. Right-trivializing the differential of the logarithm yields

$$\mathrm{d}Z \;\approx\; Z\,\widehat{\eta}, \qquad \mathrm{d}\big(\mathrm{Log}(Z)\big) \;\approx\; \mathbf{J}_r^{-1}\big(\mathrm{Log}(Z)\big)\,\eta.$$

This follows from the first-order expansion of the logarithm map under a right-invariant perturbation, namely:

$$\mathrm{Log}(Z\,\mathrm{Exp}(\widehat{\eta})) = \mathrm{Log}(Z) + \mathbf{J}_r^{-1}(\mathrm{Log}(Z))\,\eta + O(\|\eta\|^2),$$

where $\mathbf{J}_r$ denotes the right Jacobian of the exponential map in SE(d).
Since the residual is defined as

$$r_{ij} = \mathrm{Log}(Z)^{\vee},$$

and the vee operator is a fixed linear isomorphism $\vee : \mathfrak{se}(d) \to \mathbb{R}^q$, its differential is the identity. Therefore,

$$\mathrm{d}r_{ij} \;\approx\; \mathbf{J}_r^{-1}\big(r_{ij}\big)\,\mathbf{Ad}_{T_i}(\delta_j - \delta_i).$$

We have then proven the following result:

Proposition 4.6.2 (Closed-form Jacobian blocks of the relative pose residual)**.** Let $r_{ij}(\Theta) \in \mathbb{R}^q$ be the residual associated with the relative pose measurement between poses T_i and T_j. Under the right-invariant perturbation model introduced above, the nonzero blocks of the Jacobian of r_{ij} with respect to the state vector Θ are given by

$$J_i = -\,\mathbf{J}_r^{-1}\big(r_{ij}(\Theta)\big)\,\mathbf{Ad}_{T_i}, \qquad J_j = \mathbf{J}_r^{-1}\big(r_{ij}(\Theta)\big)\,\mathbf{Ad}_{T_i}. \tag{4.7}$$

Remark 4.6.3 (Choice of right-invariant perturbations)**.** The adoption of right-invariant perturbations in the derivation above is consistent with the right-invariant nature of the measurement model, in which the relative pose error is expressed as $H_{ij}T_jT_i^{-1}$. This choice leads to a linearization where perturbations are naturally composed on the right and transported through the adjoint representation associated with T_i. It is worth noting, however, that this choice only affects the coordinate representation of the associated Jacobians and of the resulting Fisher Information Matrix. The intrinsic Fisher information, interpreted as a bilinear form on the tangent space of the Lie group, is invariant to the choice between right- or left-invariant perturbation models and would remain unchanged under an alternative parametrization.

We now derive an alternative expression for the Fisher Information Matrix that explicitly exploits the sparsity of the measurements and makes both the topology of the measurement graph and the geometry of the Lie group SE(d) transparent. In particular, the Fisher Information Matrix can be factorized into a purely *topological* component and a *kinematic* component depending on the state estimate Θ. This decomposition highlights the internal structure of $\mathbf{F}(\Theta)$, clearly separating the contribution induced by the measurement graph from that arising from the geometry of SE(d).

Factorization via arc-node selection.

For each arc $(i, j) \in \mathcal{E}$, we introduce the kinematic matrix

$$L_{ij}(\Theta) := \begin{bmatrix} -A_{ij} & A_{ij} \end{bmatrix} \in \mathbb{R}^{q\times 2q}, \qquad A_{ij} := \mathbf{J}_r^{-1}(r_{ij}(\Theta))\mathbf{Ad}_{T_i},$$

and the arc-node selection operator

$$S_{ij} \doteq \begin{bmatrix} e_i^\top \otimes I_q \\ e_j^\top \otimes I_q \end{bmatrix} \in \mathbb{R}^{2q \times nq},$$

where $e_i, e_j \in \mathbb{R}^n$ denote the i-th and j-th canonical basis vectors. These definitions allow the residual Jacobian to be written compactly as

$$J_{r_{ij}}(\Theta) = L_{ij}(\Theta)\, S_{ij}.$$

From the per-arc contribution

$$Q_{ij}(\Theta) \;:=\; L_{ij}(\Theta)^\top\, \Omega_{ij}\, L_{ij}(\Theta) \in \mathbb{R}^{2q \times 2q},$$

we construct the block-diagonal matrix

$$\mathcal{Q}(\Theta) \;:=\; \text{blkdiag}\left(Q_{ij}(\Theta)\right) \in \mathbb{R}^{(2q|\mathcal{E}|)\times(2q|\mathcal{E}|)}.$$

Likewise, collecting all arc-node selection matrices yields

$$\mathcal{S} \;:=\; \begin{bmatrix} S_{i_1 j_1} \\ S_{i_2 j_2} \\ \vdots \\ S_{i_{|\mathcal{E}|} j_{|\mathcal{E}|}} \end{bmatrix} \in \mathbb{R}^{(2q|\mathcal{E}|)\times(nq)}.$$

With this notation, the Fisher Information Matrix admits the following compact and structured expression.

Theorem 4.6.4 (Equivalent Formulation of the Fisher Information Matrix for Pose Parameters in $\mathrm{SE}(d)$)**.** Let Θ denote the parameter vector, with each $T_i \in \mathrm{SE}(d)$ representing the pose of node i. Then the Fisher Information Matrix associated with Θ can be written as

$$\boxed{\mathbf{F}(\Theta) = \mathcal{S}^\top\, \mathcal{Q}(\Theta)\, \mathcal{S}} \;\in\; \mathbb{R}^{nq \times nq}. \tag{4.8}$$

In this factorization, the dependence on the measurement graph is encoded exclusively in the matrix $\mathcal{S}$, which is fixed for a given topology, while all kinematic dependence on the poses including right Jacobians, adjoint matrices, and residuals enters solely through $\mathcal{Q}(\Theta)$.

This separation is useful both analytically and numerically: the topological structure of the FIM is fully determined by $\mathcal{S}$, whereas its state dependence is captured entirely by $\mathcal{Q}(\Theta)$.

4.7 Singularity of the Fisher Information Matrix

Theorem 4.7.1 (Nullspace dimension and zero-eigenvalue multiplicity for a connected measurement graph)**.** Consider the Fisher Information Matrix $\mathbf{F}(\Theta)$ associated with the pose parameters Θ, where $T_i \in \mathrm{SE}(d)$ for $i = 1, \ldots, n$ and $q = \dim(\mathrm{SE}(d))$. Assume that the undirected measurement graph of $(\mathcal{T}, \mathcal{E})$ is connected and that $\Omega_{ij} \succ 0$ for all $(i, j) \in \mathcal{E}$. Then the Fisher Information Matrix is singular and satisfies

$$\dim(\ker(\mathbf{F}(\Theta))) = q.$$

Moreover, since $\mathbf{F}(\Theta)$ is real and symmetric, the eigenvalue 0 has algebraic multiplicity equal to q.

Step 1: Gauge Symmetry Directions Belong to the Nullspace Consider a global right action $G(\delta) = \mathrm{Exp}(\widehat{\delta}) \in \mathrm{SE}(d)$ applied uniformly to all poses:

$$T_k(\delta) = T_k\, G(\delta), \qquad k = 1, \dots, n,$$

with $\delta \in \mathbb{R}^q$. Because the measurement model is right-invariant, for every edge $(i, j) \in \mathcal{E}$ the corresponding residual remains unchanged:

$$r_{ij}(\Theta(\delta)) = r_{ij}(\Theta).$$

Therefore,

$$\frac{d}{d\delta}\, r(\Theta(\delta))\Big|_{\delta=0} = 0.$$

The associated infinitesimal variation of the parameter vector is

$$\Delta\Theta_G = [\, \delta^\top,\ \delta^\top,\ \dots,\ \delta^\top \,]^\top \in \mathbb{R}^{nq},$$

and satisfies

$$\mathbf{F}(\Theta)\, \Delta\Theta_G = 0.$$

Since δ has q independent components, this shows that at least q linearly independent directions belong to the nullspace of $\mathbf{F}(\Theta)$:

$$\dim(\ker(\mathbf{F}(\Theta))) \geq q.$$

Step 2: No Additional Null Directions for a Connected Graph We now show that no further independent directions belong to the nullspace. Using the compact factorization

$$\mathbf{F}(\Theta) = \mathcal{S}^\top \mathcal{Q}(\Theta)\, \mathcal{S},$$

it holds for any $v \in \mathbb{R}^{nq}$ that

$$v^\top \mathbf{F}(\Theta)\, v = (\mathcal{S}v)^\top \mathcal{Q}(\Theta)\, (\mathcal{S}v),$$

where $\mathcal{Q}(\Theta) = \mathrm{blkdiag}(Q_{ij})$ and each block $Q_{ij} = L_{ij}^\top \Omega_{ij} L_{ij} \succeq 0$ with $\Omega_{ij} \succ 0$. Hence,

$$v \in \mathrm{null}\left(\mathbf{F}(\Theta)\right) \quad \Longleftrightarrow \quad \mathcal{S}v \in \bigcap_{(i,j)\in\varepsilon} \mathrm{null}(Q_{ij}).$$

For a fixed edge (i, j), one has $Q_{ij}u = 0$ if and only if $L_{ij}u = 0$. Writing $u = \begin{bmatrix} u_i \\ u_j \end{bmatrix}$, this condition becomes

$$-A_{ij}u_i + A_{ij}u_j = 0.$$

Since the matrix

$$A_{ij} = \mathbf{J}_r^{-1}\big(r_{ij}(\Theta)\big)\, \mathbf{Ad}_{T_i}$$

is the product of two invertible matrices, it is invertible too. Indeed, the right Jacobian $\mathbf{J}_r\big(r_{ij}(\Theta)\big)$ is invertible in the domain of interest, hence its inverse exists. Moreover, the adjoint representation $\mathbf{Ad}_{T_i}$ is nonsingular, since it represents an automorphism of the Lie algebra. Consequently, the above condition is equivalent to

$$u_i = u_j.$$

This relation enforces a linear consistency constraint between the perturbations associated with the two nodes connected by the edge (i, j). In particular, once the perturbation $u_j \in \mathbb{R}^q$ is fixed, the corresponding perturbation u_i is uniquely determined. Because the measurement graph is connected, such constraints propagate along any path in the graph. Fixing an arbitrary reference node, all other node perturbations are uniquely determined by its value through successive applications of the above relation. Hence, the set of vectors u satisfying $L_{ij}u = 0$ for all $(i, j) \in \mathcal{E}$ forms a linear subspace of dimension at most q. Therefore,

$$\dim(\ker(\mathbf{F}(\Theta))) = q.$$

Finally, since the Fisher Information Matrix $\mathbf{F}(\Theta)$ is real and symmetric by construction, it admits an orthonormal basis of eigenvectors and is therefore diagonalizable by the spectral theorem. For real symmetric matrices, the algebraic and geometric multiplicities of each eigenvalue coincide. Since the nullspace of $\mathbf{F}(\Theta)$ has dimension q, it follows that the eigenvalue 0 has algebraic multiplicity equal to q.

Remark 4.7.2 (Structure-induced singularity)**.** Even by inspecting the structure of $\mathbf{F}(\Theta)$, one can intuit its singularity. Each row and column associated with a pose is coupled to the others only through relative quantities of the form $T_i T_j^{-1}$, and never to an absolute reference. As a consequence, $\mathbf{F}(\Theta)$ has the structure of a *block Laplacian* of the measurement graph. As for any Laplacian matrix, this structure implies the existence of a nontrivial nullspace associated with a global rigid-body transformation of the reference frame.

4.8 Extension to Multiple Connected Components

The results derived in the case of a single connected component extend naturally to the general scenario in which the measurement graph contains multiple connected components. The two formulations for the Fisher Information Matrix, (4.6) and (4.8), remain the same. In particular, all invariance properties of the log-likelihood remain unchanged: within each connected component, the measurement model is still right-invariant with respect to a common transformation in $\mathrm{SE}(d)$, and therefore the likelihood is unaffected by a global change of reference frame applied uniformly to all poses belonging to the same component.
Consequently, if the graph consists of C connected components, then the maximum-likelihood estimate of the absolute poses is not unique, since each component may undergo an independent rigid-body transformation. If we denote by $\mathcal{T}_1, \ldots, \mathcal{T}_C$ the node sets associated with the C components, any optimal configuration

$$\{T_i^{\mathrm{MLE}} \mid i \in \mathcal{T}_1 \cup \cdots \cup \mathcal{T}_C\}$$

generates an entire family of equivalent MLE solutions of the form

$$\left\{ T_i^{\mathrm{MLE}} G_k \;\middle|\; T_i \in \mathcal{T}_k,\ G_k \in \mathrm{SE}(d),\ k = 1, \ldots, C \right\},$$

where each component is transformed independently. As in the single-component case, this ambiguity does not affect the estimation of the relative poses within each component, which remain uniquely defined since all transformations G_k cancel out in expressions of the form $T_i T_j^{-1}$.

The presence of multiple connected components also influences the structure of the Fisher Information Matrix. Since no measurements couple poses belonging to different components, the FIM becomes block-diagonal, with one independent block for each connected component. As a consequence, one may either treat the overall FIM as a single block-diagonal matrix, or equivalently consider the C smaller FIMs associated with each individual component, which can be analyzed and inverted independently. Both viewpoints are completely equivalent, and the choice between them is purely notational or computational, depending on the application at hand.

Dimension of the Kernel of the FIM with c Connected Components

When the measurement graph contains multiple connected components, the Fisher Information Matrix (FIM) associated with the pose parameters exhibits a nontrivial nullspace whose dimension reflects the independent gauge freedoms present in each component. The following theorem characterizes this dimension exactly.

Theorem 4.8.1 (Dimension of the Fisher Information Nullspace for a non-connected graph of measurements)**.** Let the undirected graph induced by $\mathcal{E}$ have $c \geq 1$ connected components $\mathcal{T}_1, \ldots, \mathcal{T}_c$, with node sets $\{\mathcal{T}_\alpha\}_{\alpha=1}^c$ forming a partition of $\{1, \ldots, n\}$. Then the Fisher Information Matrix $\mathbf{F}(\Theta)$ associated with the parameters Θ, with $\dim(\mathrm{SE}(d)) = q$, satisfies

$$\dim(\ker(\mathbf{F}(\Theta))) = qc. \tag{4.9}$$

Moreover, the algebraic multiplicity of the eigenvalue 0 is also equal to qc.

Proof. Let the measurement graph induced by $\mathcal{E}$ have $c \geq 1$ connected components $\mathcal{T}_1, \ldots, \mathcal{T}_c$ with node sets $\{\mathcal{T}_\alpha\}_{\alpha=1}^c$.

Step 1: Block-diagonal structure. Since there are no edges between distinct connected components, measurements only couple poses within the same $\mathcal{T}_\alpha$. Therefore, by permuting the ordering of the parameters so that the poses of each component are grouped together, there exists a permutation matrix $\mathbf{P}$ such that

$$\mathbf{P}\,\mathbf{F}(\Theta)\,\mathbf{P}^\top = \mathrm{blkdiag}\big(\mathbf{F}_1(\Theta), \ldots, \mathbf{F}_c(\Theta)\big).$$

Equivalently, up to a permutation similarity, the Fisher Information Matrix decomposes as the direct sum of the Fisher Information Matrices associated with each connected component,

$$\mathbf{F}(\Theta) \simeq \mathbf{F}_1(\Theta) \,\oplus\, \cdots \,\oplus\, \mathbf{F}_c(\Theta).$$

Step 2: Nullspace dimension. By the result for the connected case, each block satisfies

$$\dim(\ker(\mathbf{F}_\alpha(\Theta))) = q, \qquad \alpha = 1, \ldots, c.$$

Moreover, the nullspace of a block-diagonal matrix is the direct sum of the nullspaces of its diagonal blocks. Hence,

$$\dim(\ker(\mathbf{P}\,\mathbf{F}(\Theta)\,\mathbf{P}^\top)) = \sum_{\alpha=1}^{c} \dim(\ker(\mathbf{F}_\alpha(\Theta))) = \sum_{\alpha=1}^{c} q = qc.$$

Finally, permutations preserve rank and nullity, so

$$\dim(\ker(\mathbf{F}(\Theta))) = qc.$$

Step 3: Zero eigenvalue multiplicity. Since $\mathbf{F}(\Theta)$ is real symmetric, it is diagonalizable and the algebraic and geometric multiplicities coincide. Therefore, the eigenvalue 0 has algebraic multiplicity equal to $\dim(\ker(\mathbf{F}(\Theta))) = qc$. □

4.9 M independent measures per arc

Up to this point, we have assumed that a single measurement is available for each arc (i, j), denoted by H_{ij}, and we have derived $\mathbf{F}^{(1)}(\Theta) = \mathbf{F}(\Theta)$ as the Fisher Information Matrix (FIM) obtained by considering one observation per arc.
We now extend this formulation by assuming that, for each arc (i, j), m measurements $\{H_{ij}^l\}_{l=1}^m$ are available. Each independent measurement follows the generative model

$$H_{ij}^l = \mathrm{Exp}(\widehat{\xi}_{ij}^l)\, T_i T_j^{-1}, \qquad l = 1, \dots, m, \tag{4.10}$$

where $\widehat{\xi}_{ij}^l \in \mathbb{R}^q$ is a random perturbation modeled as a zero-mean Gaussian vector,

$$\widehat{\xi}_{ij}^l \sim \mathcal{N}(0, \Sigma_{ij})\,. \tag{4.11}$$

and T_i and T_j denote the unknown absolute poses associated with nodes i and j of the relative measurement graph, which belong to the Lie group $\mathrm{SE}(d)$ and are the quantities to be estimated.
Under this assumption, all measurements collected on the same arc (i, j) are independent and identically distributed, since they share the same noise distribution characterized by the covariance matrix Σ_{ij}. Moreover, measurements associated with different arcs are assumed to be mutually independent.
By exploiting the additivity of the Fisher information for independent observations, the FIM in the case of m measurements per arc is given by

$$\mathbf{F}(\Theta)_m = \sum_{(i,j)} \sum_{l=1}^{m} \mathbf{F}_{ij}^{(1)}(\Theta) = m \sum_{(i,j)} \mathbf{F}_{ij}^{(1)}(\Theta),$$

where $\mathbf{F}_{ij}^{(1)}(\Theta)$ denotes the Fisher information contribution associated with a single measurement on edge (i, j).
Finally, recalling that $\mathbf{F}(\Theta)_1$ is defined as the total Fisher Information Matrix obtained by considering one measurement per arc,

$$\mathbf{F}(\Theta)_1 := \sum_{(i,j)} \mathbf{F}_{ij}^{(1)}(\Theta),$$

we obtain the compact scaling relationship

$$\mathbf{F}(\Theta)_m = m\, \mathbf{F}(\Theta)_1, \tag{4.12}$$

which shows that, under the assumed noise model, the available information grows linearly with the number of independent measurements collected for each edge. Moreover, since we have shown that the Fisher Information Matrix obtained by considering a single measurement per arc, $\mathbf{F}(\Theta)_1$, is singular and satisfies

$$\dim\big(\ker(\mathbf{F}(\Theta)_1)\big) = qc,$$

where c denotes the number of connected components of the graph, the same property holds when m measurements per arc are available. Indeed, the scaling factor does not affect either the rank or the null space of the matrix. As a consequence, $\mathbf{F}(\Theta)_m$ is also singular and

$$\dim\big(\ker(\mathbf{F}(\Theta)_m)\big) = qc.$$

Chapter 5

Handling Singular Fisher Information Matrices: Anchored and Quotient Approaches

5.1 Setting the Stage for Intrinsic Cramér–Rao Bounds

The pose-graph estimation problem, with the statistical model based on (4.3) is naturally defined on the product manifold

$$\mathcal{P} = \mathrm{SE}(d)^n.$$

Within this setting, we have shown that the associated Fisher Information Matrix (FIM) is singular due to the presence of global gauge symmetries. In particular, if the measurement graph has c connected components, then

$$\dim(\ker(\mathbf{F}(\Theta))) = qc \qquad q = \dim \mathrm{SE}(d) = \frac{d(d+1)}{2}.$$

As a consequence, the Cramér–Rao Bound (CRB) cannot be applied in its classical form, which requires an invertible Fisher Information Matrix.
In this chapter, we develop two rigorous geometric approaches to eliminate gauge ambiguities and derive meaningful intrinsic Cramér–Rao bounds. Throughout the chapter, we work under the additional assumption of a *high signal-to-noise ratio (high-SNR) regime*, meaning that estimation errors are sufficiently small for local linearizations on the manifold to be accurate and for the logarithm map to be well defined in a neighbourhood of the true parameters.
The two approaches considered are the following:

1. **Anchored CRB:** a set of anchor poses is fixed to their true values, with at least one anchor selected for each connected component of the relative-measurement graph. The set of anchor indices is denoted by $A \subseteq \{1, \dots, n\}$. This constraint restricts the estimation problem to a Riemannian submanifold of the original parameter space, thereby removing the gauge freedom associated with the global transformations within each connected component.

2. **Quotient-manifold CRB:** the relative measurement model induces an equivalence relation on $\mathrm{SE}(d)^n$ by identifying all absolute pose configurations that give rise to the same likelihood. This equivalence relation is generated by independent global right actions of $\mathrm{SE}(d)$ acting on each connected component of the measurement graph. The resulting equivalence classes define the quotient manifold

$$\bar{\mathcal{P}}_{\mathcal{O}} = \mathrm{SE}(d)^n / \mathrm{SE}(d)^c,$$

which represents the space of intrinsically identifiable parameters. Vertical directions correspond to unobservable gauge variations, while horizontal directions encode observable degrees of freedom, providing the natural setting for the intrinsic Cramér–Rao bound.

The geometric framework developed here follows the intrinsic Cramér–Rao bounds for submanifolds and quotient manifolds introduced by N. Boumal [19], and is therefore directly applicable to the present setting, where the parameter space $\mathrm{SE}(d)^n$ is a Lie group.

5.2 Anchored Estimation on a Submanifold of $\mathrm{SE}(d)^n$

In this section, we present the complete geometric construction and the algebraic derivation of the orthogonal projector required for N. Boumal's submanifold Cramér–Rao bound, under the assumption that at least one pose is anchored (i.e. fixed to the true value) in each connected component.

We consider n poses $T_1, \dots, T_n \in \mathrm{SE}(d)$ and represent them as a single point

$$Q = (T_1, \dots T_n) \in \mathcal{P} := \mathrm{SE}(d)^n.$$

The parameter space $\mathcal{P} = \mathrm{SE}(d)^n$ is a product manifold. As such, its tangent space at a point $Q = (T_1, \dots, T_n) \in \mathcal{P}$ is given by the direct sum of the tangent spaces of the individual factors, namely

$$T_Q\mathcal{P} = T_{T_1}\mathrm{SE}(d) \ \oplus \ \cdots \ \oplus \ T_{T_n}\mathrm{SE}(d).$$

Proposition 5.2.1 (Tangent space of $\mathrm{SE}(d)$ via left translation)**.** For every $T_i \in \mathrm{SE}(d)$, the tangent space at T_i satisfies

$$T_{T_i}\mathrm{SE}(d) = \{\, T_i S \mid S \in \mathfrak{se}(d) \,\}.$$

In particular, since $q = \dim \mathfrak{se}(d) = \frac{d(d+1)}{2}$, the tangent space $T_{T_i}\mathrm{SE}(d)$ is a vector space of dimension q and is canonically isomorphic to $\mathbb{R}^q$.

Proof. Consider the left translation by T_i,

$$L_{T_i} : \mathrm{SE}(d) \to \mathrm{SE}(d), \qquad L_{T_i}(G) = T_i G.$$

Since L_{T_i} is a diffeomorphism, its differential at the identity,

$$(dL_{T_i})_e : T_e\mathrm{SE}(d) \to T_{T_i}\mathrm{SE}(d),$$

is a linear isomorphism. By definition of the Lie algebra, $T_e\mathrm{SE}(d) = \mathfrak{se}(d)$. Moreover, for any $S \in \mathfrak{se}(d)$, take a smooth curve $\gamma(t) \subset \mathrm{SE}(d)$ such that $\gamma(0) = e$ and $\dot{\gamma}(0) = S$. Then the curve $L_{T_i}(\gamma(t)) = T_i\gamma(t)$ satisfies

$$L_{T_i}(\gamma(0)) = T_i, \qquad \frac{d}{dt}\Big|_{t=0} T_i\gamma(t) = (dL_{T_i})_e(S).$$

In the matrix-group setting, the differential of left translation acts by left multiplication, hence $(dL_{T_i})_e(S) = T_iS$. Therefore every tangent vector at T_i is of the form T_iS for some $S \in \mathfrak{se}(d)$, and conversely every T_iS arises as the velocity at $t = 0$ of the curve $t \mapsto T_i \operatorname{Exp}(tS)$. This proves

$$T_{T_i}\mathrm{SE}(d) = \{\, T_iS \mid S \in \mathfrak{se}(d) \,\}.$$

□

Via a fixed linear isomorphism $T_{T_i}\mathrm{SE}(d) \simeq \mathbb{R}^q$, each tangent space is represented in local coordinates. Accordingly, a tangent vector to the product manifold $\mathcal{P} = \mathrm{SE}(d)^n$ at a point $Q = (T_1, \ldots, T_n)$ can be written in stacked form as

$$v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}, \qquad v_i \in \mathbb{R}^q,$$

where v_i represents the infinitesimal variation associated with the pose T_i.
We now select a set of indices

$$A \subseteq \{1, \ldots, n\},$$

referred to as the anchor set, such that at least one index is selected for each connected component of the measurement graph. The elements of A correspond to poses that are assumed to be known a priori.
More precisely, for each $i \in A$, the pose T_i is fixed to the real value $\bar{T}_i \in \mathrm{SE}(d)$ and is therefore not subject to estimation.
The resulting anchored submanifold is the subset of the product manifold defined by

$$\bar{\mathcal{P}}_A := \big\{(T_1, \ldots, T_n) \in \mathrm{SE}(d)^n \;\big|\; T_i = \bar{T}_i \;\forall\, i \in A\big\}.$$

This construction restricts the admissible configurations to those consistent with the anchoring constraints, thereby removing the corresponding gauge degrees of freedom and yielding a lower-dimensional submanifold on which the estimation problem is formulated.

Tangent Space of the Anchored Submanifold Since the poses indexed by A are fixed, they cannot vary, and therefore the corresponding tangent components must vanish. As a result, the tangent space of $\bar{\mathcal{P}}_A$ at a point $Q = (T_1, \ldots, T_n)$ is given by

$$T_Q\bar{\mathcal{P}}_A = \left(\bigoplus_{i \notin A} T_{T_i}\mathrm{SE}(d)\right) \;\oplus\; \left(\bigoplus_{i \in A} \{0\}\right).$$

In stacked coordinates, a tangent vector to $\bar{\mathcal{P}}_A$ takes the form

$$v = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}, \qquad v_i = 0 \ \ \forall\, i \in A,$$

where each component $v_i \in \mathbb{R}^q$. The nonzero components v_i represent admissible infinitesimal variations of the non-anchored poses.

Orthogonal Projection onto the Anchored Tangent Space The projector $\Pi : T_Q\mathcal{P} \to T_Q\bar{\mathcal{P}}_A$ acts by preserving the tangent components associated with non-anchored poses and annihilating those corresponding to anchored poses. To this end, introduce a diagonal matrix $M \in \mathbb{R}^{n\times n}$ defined by

$$M_{ii} = \begin{cases} 1, & i \notin A, \\ 0, & i \in A, \end{cases} \qquad M_{ij} = 0 \ \text{ for } i \neq j.$$

The orthogonal projector onto $T_Q\bar{\mathcal{P}}_A$ is then given by

$$\Pi = M \otimes I_q \ \in \ \mathbb{R}^{nq\times nq},$$

where I_q denotes the identity matrix in $\mathbb{R}^q$.
Its action on a stacked tangent vector $v \in T_Q\mathcal{P}$ is

$$(\Pi v)_i = \begin{cases} v_i, & i \notin A, \\ 0, & i \in A. \end{cases}$$

By construction, the matrix Π is symmetric and idempotent.

Theorem 5.2.2 (Anchored intrinsic Cramér–Rao bound neglecting curvature)**.** Consider an estimation problem on the manifold $\mathcal{P} = \mathrm{SE}(d)^n$ with measurement model (4.3), and let $\mathbf{F}$ denote the Fisher Information Matrix. Assume that curvature terms in the intrinsic Cramér–Rao bound are negligible.
Let $\Pi : T_\Theta\mathcal{P} \to T_\Theta\mathcal{P}$ be the orthogonal projector onto the tangent space, at Θ, of the anchored submanifold induced by a given anchor set.
Then, for any unbiased estimator defined on the anchored submanifold $\bar{\mathcal{P}}_A$, the covariance matrix $\mathrm{Cov}_\Theta(\hat{\Theta})$ of the intrinsic estimation error satisfies

$$\mathrm{Cov}_\Theta(\hat{\Theta}) \succeq \big(\Pi\,\mathbf{F}\,\Pi\big)^\dagger = \big((M \otimes I_q)\,\mathbf{F}\,(M \otimes I_q)\big)^\dagger. \tag{5.1}$$

Furthermore, the total variance admits the bound

$$\mathrm{tr}\big(\mathrm{Cov}_\Theta(\hat{\Theta})\big) \ \geq \ \mathrm{tr}\big(\Pi\,\mathbf{F}\,\Pi\big)^\dagger. \tag{5.2}$$

Proof. The result follows directly from Theorem 2 in [19], specialized to the submanifold setting and under the assumption that curvature terms are negligible.
Indeed, the anchored parameter space $\bar{\mathcal{P}}_A \subset \mathcal{P}$ is a Riemannian submanifold of $\mathcal{P} = \mathrm{SE}(d)^n$, and the matrix $\Pi = M \otimes I_q$ is precisely the orthogonal projector from $T_Q\mathcal{P}$ onto $T_Q\bar{\mathcal{P}}_A$. Applying the aforementioned result yields the bound

$$\mathrm{Cov}_\Theta(\hat{\Theta}) \succeq \big(\Pi\,\mathbf{F}\,\Pi\big)^\dagger,$$

which proves (5.1). The trace inequality (5.2) follows immediately as a consequence of this bound and is included in the same result. □

Component-wise interpretation of the anchored bound Under the assumption that the measurement graph has c connected components and that the anchor set A contains at least one pose per component, the matrix $\Pi\,\mathbf{F}\,\Pi$ inherits a block structure aligned with the decomposition of the parameter space induced by the graph. Indeed, by reordering the variables so that poses belonging to the same connected component are grouped together, there exists a permutation matrix $\mathbf{P}$ such that

$$\mathbf{P}\,\mathbf{F}\,\mathbf{P}^\top = \mathrm{blkdiag}\big(\mathbf{F}_1, \dots, \mathbf{F}_c\big), \qquad \mathbf{P}\,\Pi\,\mathbf{P}^\top = \mathrm{blkdiag}\big(\Pi_1, \dots, \Pi_c\big),$$

where each $\mathbf{F}_\alpha$ is the Fisher Information Matrix associated with the subgraph induced by $\mathcal{T}_\alpha$, and Π_α is the corresponding anchoring projector on that component. Consequently,

$$\mathbf{P}\,(\Pi\mathbf{F}\Pi)\,\mathbf{P}^\top = (\mathbf{P}\Pi\mathbf{P}^\top)(\mathbf{P}\mathbf{F}\mathbf{P}^\top)(\mathbf{P}\Pi\mathbf{P}^\top) = \mathrm{blkdiag}\big(\Pi_1\mathbf{F}_1\Pi_1, \dots, \Pi_c\mathbf{F}_c\Pi_c\big).$$

As a consequence, the bound in (5.1) decomposes into c independent intrinsic Cramér–Rao bounds, each acting on the parameters associated with a single connected component. This reflects the fact that anchoring removes the gauge freedom within each component independently, while no information is exchanged across distinct components.

From an algebraic point of view, the projector $\Pi = M \otimes I_q$ enforces the anchoring constraints by restricting the Fisher Information Matrix to the subspace of admissible tangent directions. The resulting bound therefore characterizes the estimation accuracy attainable under fixed-anchor constraints, component by component.

5.3 Quotient-Manifold Estimation on $\mathrm{SE}(d)^n$

Let $\mathrm{SE}(d)$ denote the special Euclidean group, and consider the configuration space $\mathrm{SE}(d)^n$ of n absolute poses. Assume that the undirected measurement graph consists of $c \geq 1$ connected components, denoted by $\mathcal{T}_1, \dots, \mathcal{T}_c$, which form a partition of $\{1, \dots, n\}$.

In this setting, relative measurements within each connected component are invariant under an independent global rigid motion acting on that component. This invariance can be formalized by introducing a right action of the group $\mathrm{SE}(d)^c$ on $\mathrm{SE}(d)^n$.

Specifically, for $\Theta = (T_1, \dots, T_n) \in \mathrm{SE}(d)^n$ and $(G_1, \dots, G_c) \in \mathrm{SE}(d)^c$, we define the *right action*

$$\Phi : \mathrm{SE}(d)^n \times \mathrm{SE}(d)^c \to \mathrm{SE}(d)^n, \qquad \Phi(\Theta, (G_1, \dots, G_c)) = (T_1 G_{\alpha(1)}, \dots, T_n G_{\alpha(n)}), \tag{5.3}$$

where $\alpha(i)$ denotes the index of the connected component containing node i.

Two configurations $\Theta, \Theta' \in \mathrm{SE}(d)^n$ are said to be *gauge-equivalent* if they belong to the same orbit of this action, that is, if there exists $(G_1, \dots, G_c) \in \mathrm{SE}(d)^c$ such that

$$\Theta' = \Phi(\Theta, (G_1, \dots, G_c)).$$

The orbit through Θ,

$$\mathcal{O}_\Theta = \big\{(T_1 G_{\alpha(1)}, \dots, T_n G_{\alpha(n)}) : \ (G_1, \dots, G_c) \in \mathrm{SE}(d)^c\big\}, \tag{5.4}$$

collects all absolute pose configurations that induce the same set of relative poses and therefore the same likelihood function. In other words, configurations lying in the same orbit are statistically indistinguishable on the basis of the available measurements. The set of all such orbits defines the quotient space

$$\bar{\mathcal{P}}_{\mathcal{O}} := \mathrm{SE}(d)^n / \mathrm{SE}(d)^c,$$

together with the canonical projection

$$\pi : \mathrm{SE}(d)^n \to \bar{\mathcal{P}}_{\mathcal{O}}, \qquad \pi(\Theta) := [\Theta].$$

Since the action (5.3) is smooth and proper, the quotient $\bar{\mathcal{P}}_{\mathcal{O}}$ inherits a smooth manifold structure, and π is a smooth submersion. A dimension count yields

$$\dim(\bar{\mathcal{P}}_{\mathcal{O}}) = n \dim(\mathrm{SE}(d)) - c \dim(\mathrm{SE}(d)) = (n-c) \dim(\mathrm{SE}(d)),$$

corresponding to the removal of c independent gauge degrees of freedom.

From a differential-geometric perspective, the tangent space at $\Theta \in \mathrm{SE}(d)^n$ admits a natural decomposition induced by the quotient structure. The tangent space of the total space is

$$T_\Theta \mathrm{SE}(d)^n \cong T_{T_1}\mathrm{SE}(d) \times \cdots \times T_{T_n}\mathrm{SE}(d).$$

The *vertical subspace* at Θ is defined as the kernel of the differential of the projection,

$$\mathcal{V}_\Theta := \ker(d\pi_\Theta),$$

and coincides with the tangent space to the gauge orbit $\mathcal{O}_\Theta$. It is spanned by infinitesimal gauge transformations acting independently on each connected component. More explicitly, for $\xi_1, \ldots, \xi_c \in \mathfrak{se}(d)$, the curve

$$\Theta(t) = (T_1 \mathrm{Exp}(t\xi_{\alpha(1)}), \ldots, T_n \mathrm{Exp}(t\xi_{\alpha(n)}))$$

lies entirely in $\mathcal{O}_\Theta$, and its velocity at $t = 0$ belongs to $\mathcal{V}_\Theta$. Consequently,

$$\dim(\mathcal{V}_\Theta) = c \dim(\mathrm{SE}(d)).$$

To characterize the directions corresponding to observable variations of the parameters, we introduce the *horizontal subspace* $\mathcal{H}_\Theta$ as a complement of $\mathcal{V}_\Theta$. After choosing an inner product on $T_\Theta \mathrm{SE}(d)^n$, the horizontal space is defined as the orthogonal complement

$$\mathcal{H}_\Theta := \mathcal{V}_\Theta^\perp, \qquad T_\Theta \mathrm{SE}(d)^n = \mathcal{V}_\Theta \oplus \mathcal{H}_\Theta.$$

Geometrically, $\mathcal{H}_\Theta$ contains exactly those infinitesimal variations of the poses that are not induced by global rigid motions and therefore correspond to identifiable, observable directions of the estimation problem.

As a special case, when the measurement graph is connected ($c = 1$), the action (5.3) reduces to the standard right diagonal action of $\mathrm{SE}(d)$ on $\mathrm{SE}(d)^n$, the quotient space becomes $\mathrm{SE}(d)^n / \mathrm{SE}(d)$, and the vertical and horizontal spaces reduce to the classical decomposition associated with a single global gauge freedom.

If we consider curvature terms of the bound negligible, N. Boumal's intrinsic Cramér–Rao bound on quotient manifolds applies directly in this setting and yields the following result.

Theorem 5.3.1 (Quotient Intrinsic Cramér–Rao bound neglecting curvature)**.** Let the undirected measurement graph consist of $c \geq 1$ connected components, and consider the estimation problem on the quotient manifold

$$\bar{\mathcal{P}}_{\mathcal{O}} = \mathrm{SE}(d)^n / \mathrm{SE}(d)^c.$$

Neglecting curvature terms, the covariance matrix $\mathrm{Cov}_{\Theta}(\hat{\Theta})$ of any unbiased estimator defined on $\bar{\mathcal{P}}_{\mathcal{O}}$ satisfies the intrinsic Cramér–Rao bound

$$\mathrm{Cov}_{\Theta}(\hat{\Theta}) \succeq \mathbf{F}^{\dagger}, \tag{5.5}$$

where $\mathbf{F}$ denotes the Fisher Information Matrix associated with the full parameter vector Θ. Moreover, the total variance admits the scalar bound

$$\mathrm{tr}\big(\mathrm{Cov}_{\Theta}(\hat{\Theta})\big) \;\geq\; \mathrm{tr}\big(\mathbf{F}^{\dagger}\big). \tag{5.6}$$

Proof. The result follows as a direct application of Theorem 3 in [19]. Indeed, when the undirected measurement graph has $c \geq 1$ connected components, the gauge equivalence relation induced by the relative measurement model identifies configurations that differ by independent global right actions of $\mathrm{SE}(d)$ on each connected component, as described in (5.4). Quotienting the parameter space $\mathrm{SE}(d)^n$ with respect to these equivalence classes yields the quotient manifold $\bar{\mathcal{P}}_{\mathcal{O}} = \mathrm{SE}(d)^n / \mathrm{SE}(d)^c$.
In this setting, the assumptions of Theorem 3 in [19] are satisfied, and neglecting curvature terms leads directly to the intrinsic Cramér–Rao bound (5.5). The associated trace inequality (5.6) is also covered by the same theorem. □

Component-wise interpretation of the quotient CRB To interpret this result more explicitly, observe that, after a suitable permutation of the variables—performed by the same permutation matrix $\mathbf{P}$ introduced in the anchored case, which groups together the poses belonging to the same connected component—the Fisher Information Matrix satisfies

$$\mathbf{P}\,\mathbf{F}\,\mathbf{P}^{\top} = \mathrm{blkdiag}(\mathbf{F}_1, \ldots, \mathbf{F}_c),$$

where each block $\mathbf{F}_{\alpha}$ coincides with the Fisher Information Matrix of the connected subproblem associated with $\mathcal{T}_{\alpha}$. Accordingly, the Moore–Penrose pseudoinverse satisfies

$$\mathbf{P}\,\mathbf{F}^{\dagger}\,\mathbf{P}^{\top} = \mathrm{blkdiag}(\mathbf{F}_1^{\dagger}, \ldots, \mathbf{F}_c^{\dagger}).$$

This block-diagonal structure shows that the intrinsic Cramér–Rao bound on the quotient manifold decomposes naturally into c independent bounds, each governing the estimation accuracy within a single connected component. From an algebraic standpoint, the global bound is therefore equivalent to the collection of the intrinsic bounds associated with the individual connected subproblems, reflecting the fact that no information is exchanged across distinct components.

5.4 Identifiable maximum likelihood estimators for CRB validation

Constructing explicitly unbiased estimators in pose-graph estimation problems is a nontrivial task, both in the anchored and in the quotient settings. Rather than attempting

to enforce unbiasedness directly, a natural strategy consists in exploiting the asymptotic efficiency of the maximum likelihood estimator, which provides a principled connection between estimation performance and Cramér–Rao bounds in identifiable parameter spaces.
In order to empirically validate the Cramér–Rao bounds derived in this chapter, it is therefore necessary to consider estimation settings in which the MLE is uniquely defined. As discussed earlier, the gauge symmetry induced by relative pose measurements prevents uniqueness of the MLE on the full space $\mathrm{SE}(d)^n$. Two complementary identifiable formulations are considered, corresponding to the anchored space $\bar{\mathcal{P}}_A$ and to the quotient space $\bar{\mathcal{P}}_{\mathcal{O}}$.

In the anchored setting, the MLE is obtained by restricting the likelihood to $\bar{\mathcal{P}}_A$, i.e., by fixing the poses indexed by A to their ground-truth values and optimizing over the remaining variables. This restriction removes the gauge ambiguity and leads to a unique maximizer,

$$\widehat{\Theta}_A^{\mathrm{MLE}} = \arg\max_{\Theta \in \bar{\mathcal{P}}_A} \mathcal{L}(\Theta). \tag{5.7}$$

In this case the anchored MLE satisfies the standard regularity conditions required for asymptotic efficiency. As the number of independent measurements per edge grows, the empirical covariance of the anchored MLE approaches the intrinsic Cramér–Rao bound given in (5.1), in agreement with the asymptotic efficiency of maximum likelihood estimation.

In the quotient setting, the estimation problem is formulated intrinsically on $\bar{\mathcal{P}}_{\mathcal{O}}$, where the likelihood is invariant along gauge orbits and defines a function on equivalence classes. The quotient MLE is then defined as the unique maximizer of the induced cost function on the quotient,

$$[\widehat{\Theta}]_{\mathcal{O}}^{\mathrm{MLE}} = \arg\max_{[\Theta] \in \bar{\mathcal{P}}_{\mathcal{O}}} \mathcal{L}([\Theta]). \tag{5.8}$$

Although individual representatives in $\mathrm{SE}(d)^n$ are not uniquely defined, the maximizing equivalence class is unique.
From a differential-geometric perspective, this optimality condition can be expressed by requiring that the log-likelihood be stationary with respect to all horizontal (i.e., identifiable) variations at a representative $\widehat{\Theta}$ of the optimal equivalence class,

$$\Pi_{\mathcal{H}_{\widehat{\Theta}}}\big(\nabla_\Theta \mathcal{L}([\widehat{\Theta}]_{\mathcal{O}}^{\mathrm{MLE}})\big) = 0,$$

where $\Pi_{\mathcal{H}_{\widehat{\Theta}}}$ denotes the projection operator onto the horizontal subspace at $\widehat{\Theta}$, since variations along vertical directions leave the cost unchanged.

In both the anchored and quotient formulations, the existence of a unique and identifiable MLE provides the essential link between the theoretical bounds and empirical performance. The asymptotic efficiency of $\widehat{\Theta}_A^{\mathrm{MLE}}$ on $\bar{\mathcal{P}}_A$ and of $[\widehat{\Theta}]_{\mathcal{O}}^{\mathrm{MLE}}$ on $\bar{\mathcal{P}}_{\mathcal{O}}$ ensures that the corresponding Cramér–Rao bounds can be meaningfully verified through numerical experiments.

5.5 On the Role of Curvature in Intrinsic CRBs

Although the anchored and quotient-manifold formulations are conceptually distinct, they exhibit a remarkably similar algebraic structure. In both cases, after a suitable permutation of the variables, the Fisher Information Matrix and the corresponding Cramér–Rao bound decompose into independent blocks associated with the connected components of the measurement graph. This reflects the fact that gauge ambiguities—and their removal—act independently within each connected component, with no statistical information exchanged across disconnected subgraphs.

Despite this structural similarity, the two approaches differ in their geometric interpretation and level of abstraction. The anchored formulation operates on a concrete and easily interpretable submanifold of the original parameter space, obtained by fixing a selected subset of poses to their true values. Identifiability is enforced directly by explicitly constraining specific degrees of freedom.
By contrast, the quotient-manifold formulation removes gauge ambiguities by identifying all configurations that induce the same likelihood. The resulting space is not a subset of $\mathrm{SE}(d)^n$, but a space of equivalence classes representing families of statistically indistinguishable configurations. While less concrete geometrically, this construction is conceptually closer to the intrinsic statistical structure of the problem.

From a theoretical perspective, both formulations admit a refined treatment including curvature-dependent correction terms in the Cramér–Rao bound. These arise because $\mathrm{SE}(d)$, and hence $\mathrm{SE}(d)^n$, is a non-flat manifold, so second-order geometric effects contribute to the bound. Although their inclusion yields, in principle, sharper bounds, their explicit evaluation is analytically demanding.

Nevertheless, when m independent measurements are available for each arc, the curvature-dependent correction term vanishes asymptotically as $m \to \infty$. More precisely, it scales as $O(1/m^2)$, decaying faster than the leading-order contribution, which scales as $O(1/m)$. Consequently, for sufficiently large m, the dominant contribution to the estimation error is fully captured by the leading-order term derived in this chapter, while curvature effects become negligible. This behavior is further supported by the numerical simulation presented in the following chapter.

Chapter 6

A Case Study on Sparse Pose-Graph Estimation on SE(2)

6.1 Problem Setup and Objectives

In this chapter, we study a synthetic pose-graph estimation problem on the Lie group SE(2) with five nodes ($n = 5$) and a sparse set of directed relative measurements.
We first consider a controlled baseline scenario in which a single relative measurement is available for each directed arc of the graph. This setting allows us to isolate and analyze the fundamental properties of the estimation problem, focusing on:

- maximum likelihood estimation (MLE) under left-multiplicative noise on SE(2),
- the structure of the Fisher Information Matrix (FIM) and the dimension of its nullspace,
- the role of graph connectivity and gauge freedoms in determining identifiability.

We then extend the analysis to the case in which m independent measurements are available for each arc. In this second phase, we investigate the behavior of the MLE as m increases, and assess whether it exhibits the expected asymptotic efficiency, as predicted by estimation theory.
The results reported in this chapter are obtained from an actual numerical simulation (Python code used to generate the plots provided separately in the appendix).
In the analyzed example, the available directed measurement arcs are:

$$\mathcal{E} = \{(1,2),(2,1),(3,4),(4,5),(5,4),(5,3)\}, \qquad |\mathcal{E}| = 6.$$

This measurement graph is composed of two connected components ($c = 2$):

$$\{1,2\} \quad \text{and} \quad \{3,4,5\}.$$

As a consequence, we expect a *six-dimensional* global gauge freedom (three degrees of freedom per connected component), which will be reflected by the rank deficiency of the Fisher matrix.

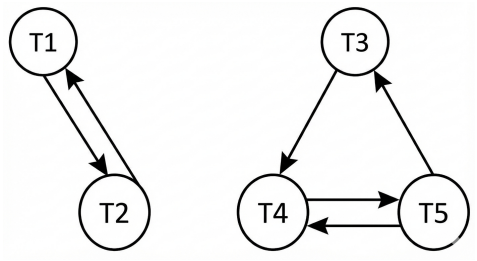


Figure 6.1: Directed graph representation of the measurement topology.

6.2 Lie Group Model and Notation

An element of SE(2) is represented as

$$T = \begin{bmatrix} R(\omega) & v \\ 0 & 1 \end{bmatrix}, \qquad R(\omega) \in \mathrm{SO}(2), \quad v \in \mathbb{R}^2. \tag{6.1}$$

The Lie algebra $\mathfrak{se}(2)$ is parameterized by

$$\xi = [\omega, v_x, v_y]^\top \in \mathbb{R}^3,$$

with the associated hat operator $\widehat{\xi} \in \mathfrak{se}(2)$.

Chosen basis of $\mathfrak{se}(2)$: we use the standard canonical generator basis

$$E_1 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

With this convention, the hat map is the linear identification

$$\widehat{\xi} \;=\; \omega\, E_1 + v_x\, E_2 + v_y\, E_3 \;=\; \begin{bmatrix} 0 & -\omega & v_x \\ \omega & 0 & v_y \\ 0 & 0 & 0 \end{bmatrix} \;=\; \begin{pmatrix} \hat{\omega} & v \\ 0 & 0 \end{pmatrix} \in \mathfrak{se}(2)$$

Exponential and Logarithm maps: The exponential and logarithm maps $\mathrm{Exp}_{\mathrm{SE}(2)}$ and $\mathrm{Log}_{\mathrm{SE}(2)}$ are implemented in closed form and are used consistently throughout the simulation, estimation, and information analysis.

Identifying an element of the Lie algebra $\mathfrak{se}(2)$ with $\xi = (\omega, v)^\top \in \mathbb{R} \times \mathbb{R}^2$, the exponential map $\mathrm{Exp}_{\mathrm{SE}(2)} : \mathfrak{se}(2) \to \mathrm{SE}(2)$ is given by

$$\mathrm{Exp}_{\mathrm{SE}(2)}(\hat{\xi}) = \begin{pmatrix} R(\omega) & V(\omega)\, v \\ 0 & 1 \end{pmatrix},$$

where

$$R(\omega) = \begin{pmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{pmatrix}, \qquad V(\omega) = \begin{pmatrix} \frac{\sin\omega}{\omega} & -\frac{1-\cos\omega}{\omega} \\ \frac{1-\cos\omega}{\omega} & \frac{\sin\omega}{\omega} \end{pmatrix},$$

with the convention $V(0) = I_2$.

Conversely, for an element

$$T = \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix} \in \mathrm{SE}(2),$$

the logarithm map $\mathrm{Log}_{\mathrm{SE}(2)} : \mathrm{SE}(2) \to \mathfrak{se}(2)$ is equal to

$$\mathrm{Log}_{\mathrm{SE}(2)}(T) = \begin{pmatrix} \hat{\omega} & V(\omega)^{-1} t \\ 0 & 0 \end{pmatrix} \in \mathfrak{se}(2), \qquad \omega \in (-\pi, \pi],\ R = R(\omega),$$

Noise-Free Measurements

The (unknown) true absolute poses are chosen as:

$$\bar{T}_1 = \begin{bmatrix} 0.8525 & -0.5227 & -0.6 \\ 0.5227 & 0.8525 & 0.2 \\ 0 & 0 & 1 \end{bmatrix} \qquad \bar{T}_2 = \begin{bmatrix} 0.4976 & -0.8674 & 0.9 \\ 0.8674 & 0.4976 & 0.1 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\bar{T}_3 = \begin{bmatrix} 0.7317 & 0.6816 & -0.2 \\ -0.6816 & 0.7317 & 1.2 \\ 0 & 0 & 1 \end{bmatrix} \qquad \bar{T}_4 = \begin{bmatrix} 0.9553 & -0.2955 & 1.1 \\ 0.2955 & 0.9553 & 1.6 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\bar{T}_5 = \begin{bmatrix} 0.9888 & 0.1494 & 2.0 \\ -0.1494 & 0.9888 & 0.9 \\ 0 & 0 & 1 \end{bmatrix}.$$

The superscript $\bar{(\cdot)}$ denotes the *noise-free* quantities. These values are used to generate synthetic measurements and for the exact values of the Fisher Information Matrix.

6.3 Measurement Noise Model

Measurements are corrupted using (4.3):

$$H_{ij} = \mathrm{Exp}(\widehat{\xi_{ij}})\, \bar{T}_{ij} \qquad \xi_{ij} \sim \mathcal{N}(0, \Sigma_{ij}),$$

where $\xi_{ij} \in \mathbb{R}^3$ represents noise in $\mathfrak{se}(2)$ coordinates.
The noise has been choosen to be *heteroscedastic*: each directed edge has its own covariance matrix, that we choose diagonal, to reduce the computations needed in each simulation.

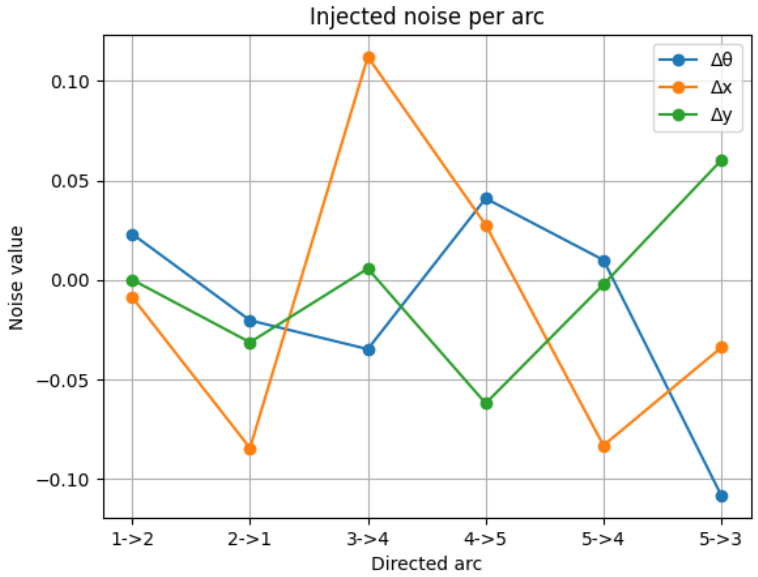
$$\Sigma_{ij} = \mathrm{diag}(\sigma^2_{\omega,ij}, \sigma^2_{x,ij}, \sigma^2_{y,ij}), \qquad \Omega_{ij} = \Sigma_{ij}^{-1}.$$

The six edge-specific covariances used in the simulation are:

$$\Sigma_{12} = \mathrm{diag}(0.0060,\ 0.0010,\ 0.0080), \qquad \Sigma_{21} = \mathrm{diag}(0.0020,\ 0.0090,\ 0.0010),$$

$$\Sigma_{34} = \mathrm{diag}(0.0050,\ 0.0070,\ 0.0090), \qquad \Sigma_{45} = \mathrm{diag}(0.0070,\ 0.0060,\ 0.0100),$$

$$\Sigma_{54} = \mathrm{diag}(0.0090,\ 0.0080,\ 0.0060), \qquad \Sigma_{53} = \mathrm{diag}(0.0065,\ 0.0055,\ 0.0075).$$



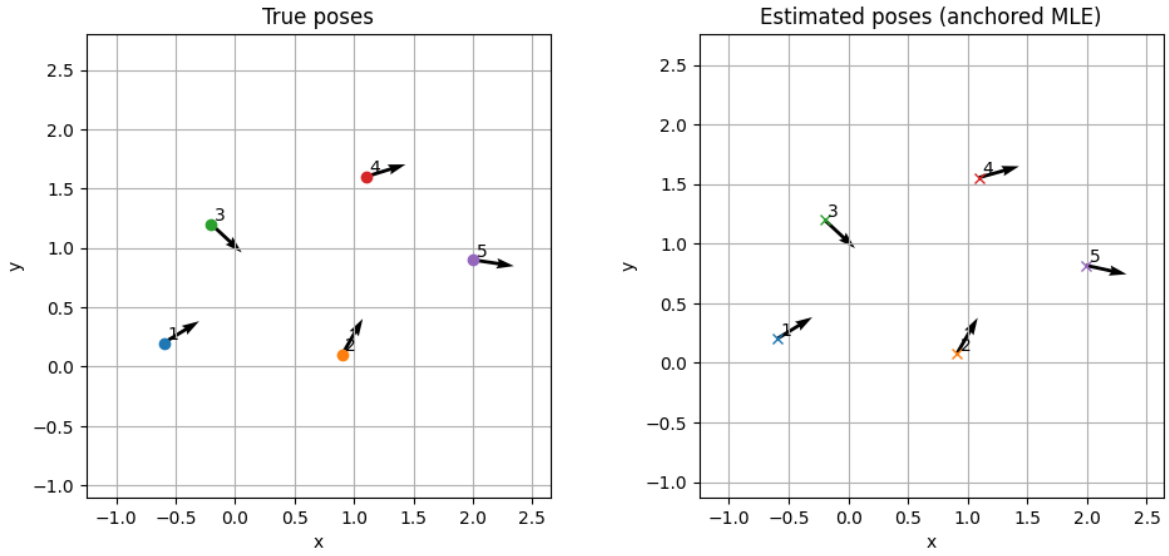
Graph 6.1: Injected noise samples ξ_{ij} in $\mathfrak{se}(2)$ for the directed sparse arc set. The rotational component ω and translational components (v_x, v_y) are shown for each directed arc.

At this stage, each directed arc is associated with a single relative measurement only. Consequently, the noise samples shown in Graph (6.1) represent the *only* random realizations injected into the relative pose measurements for each arc, and they are used to construct the corresponding noisy observations. In the next phase of the analysis, the same noise generation process will be repeated multiple times: for each directed arc, m independent noise realizations will be drawn, resulting in m relative measurements per arc. This will allow us to study the estimation performance in a multi-measurement setting.

6.4 Maximum Likelihood Estimation with Gauge Fixing

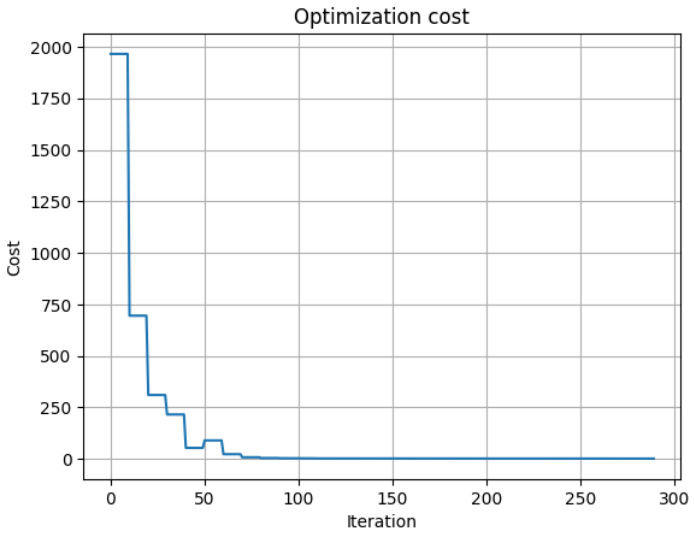
As discussed above, due to the disconnected structure of the directed measurement graph, the pose-graph estimation problem admits a six-dimensional gauge freedom, corresponding to three degrees of freedom for each connected component. As a consequence, the Fisher Information Matrix is rank-deficient and the maximum likelihood estimate is not unique unless an explicit gauge is fixed.
In the present experiment, a unique MLE solution is obtained by anchoring one pose per connected component. Specifically, the poses T_1 and T_3 are fixed to their ground-truth values $\bar{T}_1$ and $\bar{T}_3$, respectively. This choice is made solely for validation and visualization purposes: any other fixed reference (e.g., setting $T_1 = I$ and $T_3 = I$) would lead to an equivalent solution up to a rigid transformation within each component. However, anchoring the poses to their true values provides a clearer comparison between estimated and ground-truth quantities.
Graph (6.2) shows the resulting MLE absolute pose estimates. Within each connected component, the relative configuration of the poses is accurately recovered, confirming

that the optimization successfully exploits the available relative measurements despite the sparsity and directionality of the graph.



Graph 6.2: Comparison between true and estimated absolute poses.

The left image shows the ground-truth poses $\bar{T}_1, \ldots, \bar{T}_5 \in \mathrm{SE}(2)$, while the right image reports the poses estimated via maximum likelihood estimation (MLE), after anchoring T_1 and T_3 to their ground-truth values to remove gauge freedom. The estimated poses closely match the true ones, both in position and orientation, highlighting the accuracy of the proposed estimation approach.
The convergence behavior of the MLE optimization is illustrated in Graph (6.3), which reports the evolution of the negative log-likelihood cost function. After a rapid initial decrease, the cost converges to a small residual value, indicating that the estimated poses are highly consistent with the noisy measurements.



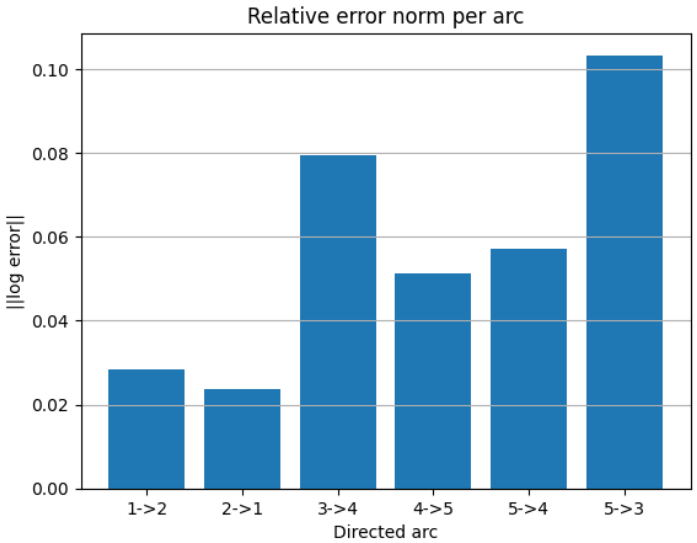
Graph 6.3: Cost function history during the MLE optimization. The rapid convergence indicates a well-conditioned optimization once the gauge is fixed.

6.5 Quality of the Estimation Through Relative Measurements

While absolute pose estimates depend on the chosen gauge, the quality of the solution can be assessed in a fully gauge-invariant manner by analyzing the reconstructed relative transformations. Indeed, for any choice of gauge within a connected component, the relative motions induced by the MLE solution remain unchanged.
For each available directed arc $(i, j) \in \mathcal{E}$, we compare the true relative motion $\bar{T}_{ij}$ with its MLE counterpart T_{ij}^{MLE} through the Lie-group error

$$\delta_{ij} = \mathrm{Log}\Big(\bar{T}_{ij}\,(T_{ij}^{\mathrm{MLE}})^{-1}\Big) \in \mathfrak{se}(2).$$

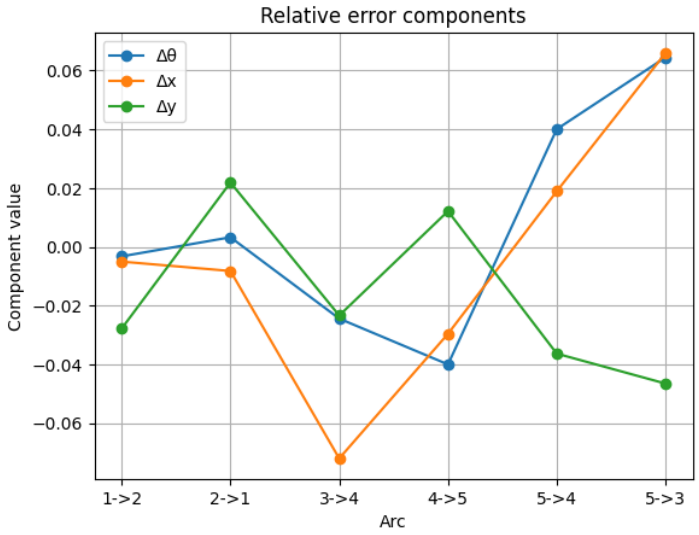
By construction, δ_{ij} is invariant to any global transformation applied to all poses within the same connected component, and therefore provides a meaningful measure of estimation accuracy independently of the adopted gauge.
Graph (6.4) reports the squared ℓ_2 norm of δ_{ij} for all available directed edges. The errors are consistently small across all measurements, indicating that the MLE solution accurately reconstructs the relative poses despite the presence of heteroscedastic noise. It is important to emphasize that, in a realistic pose-graph estimation problem, the ground-truth absolute poses $\bar{T}_i$ are not available. As a consequence, the comparison of relative transformations represents the only physically meaningful and observable validation criterion. The small relative errors observed here therefore constitute the true evidence of a reliable and consistent reconstruction, showing that the MLE solution remains close to the noise-free relative motions and correctly explains the noisy measurements.



Graph 6.4: Squared norm of the relative pose error $\|\mathrm{Log}(T_{ij}^{\star}(T_{ij}^{\mathrm{MLE}})^{-1})\|^2$ for all available directed edges.

A more detailed insight is provided in Graph (6.5), which shows the individual components of δ_{ij}, namely the rotational error $\Delta\theta$ and the translational errors Δx and Δy.

The magnitude of each component remains consistent with the low noise levels selected for the corresponding measurement edges. In particular, edges characterized by smaller noise variances exhibit correspondingly smaller estimation errors.
This behavior confirms that the MLE solution correctly reflects the heteroscedastic noise model: the lower the measurement noise associated with an edge, the higher the accuracy of the reconstructed relative transformation. As a result, the estimator effectively exploits the information content of each measurement and adapts its precision to the uncertainty structure of the problem.

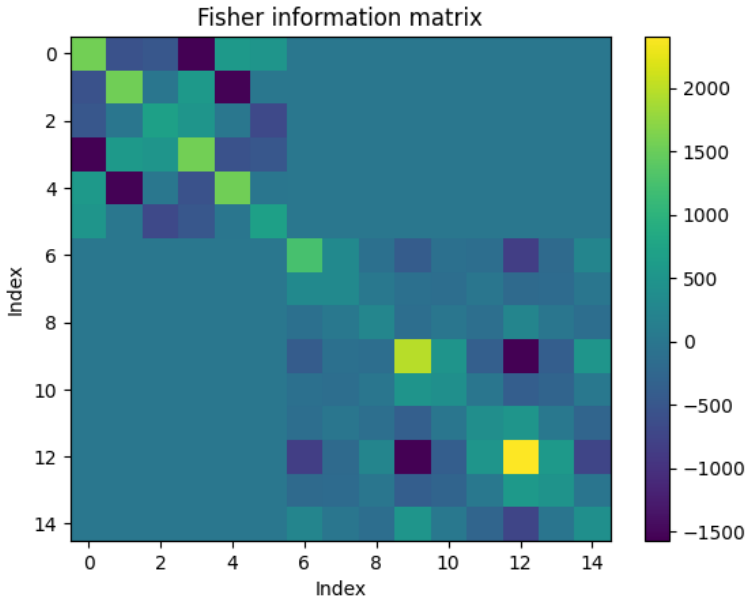


Graph 6.5: Components of the relative pose error $\delta_{ij} = (\Delta\theta, \Delta x, \Delta y)$ for each arc. The errors are small compared to the corresponding noise standard deviations.

Overall, these results confirm that, although the absolute MLE solution is defined only up to a gauge transformation within each connected component, the estimated relative motions are highly accurate. This demonstrates the effectiveness of the MLE formulation on SE(2) even in the presence of sparse, directed, and heteroscedastic measurements.

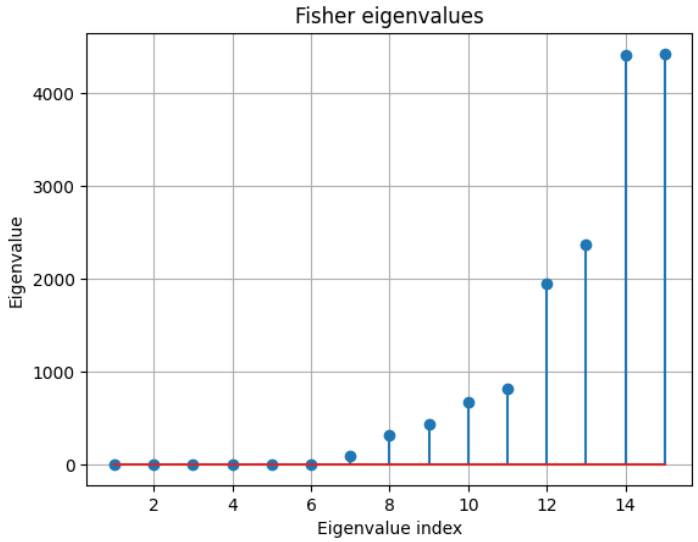
6.6 Structure and Spectrum of the Fisher Information Matrix

The Fisher Information Matrix is computed from (4.6), which provides its explicit expression in terms of the stacked pose parameters. In the sparse and directed measurement setting under consideration, the resulting matrix inherits the structure of the underlying measurement graph.
As predicted by the theoretical results presented in Chapter 3, the Fisher Information Matrix exhibits a block-diagonal structure, with each block corresponding to one connected component of the measurement graph. In the present example, the first block is associated with the component $\{1, 2\}$, while the second block corresponds to the component $\{3, 4, 5\}$. This structure is clearly visible in the heatmap representation shown in Graph (6.6).



Graph 6.6: Heatmap of the Fisher Information Matrix. The block-diagonal structure reflects the two connected components of the directed measurement graph.

The eigenvalue spectrum of the Fisher Information Matrix is reported in Graph (6.7). As expected, exactly six eigenvalues are equal to zero, empirically confirming Theorem (4.9). These zero eigenvalues correspond to the six-dimensional gauge freedom of the problem, namely three unobservable degrees of freedom (rotation and translation) for each connected component.



Graph 6.7: Eigenvalues of the Fisher Information Matrix. The presence of six zero eigenvalues reflects the gauge freedom induced by the two disconnected components.

It is worth emphasizing that the Fisher Information Matrix is a theoretical quantity defined as a function of the true (unknown) parameters of the model. In practical scenarios, these parameters are not available and the FIM is therefore often evaluated

at an estimated solution, such as the MLE. In this numerical experiment, however, the objective is to illustrate a theoretical limit. For this reason, the Fisher Information Matrix is computed using the ground-truth parameter values.

6.7 Empirical Validation of the Cramér–Rao Bounds

We conclude the analysis by empirically validating the two intrinsic Cramér–Rao bounds derived in this thesis, starting from the anchored formulation stated in Theorem (5.2.2) and then considering the quotient formulation given in Theorem (5.3.1). The goal is to assess whether, and under which conditions, the empirical covariance of the MLE approaches the theoretical limits predicted by the Fisher Information Matrix $\mathbf{F}(\Theta)$, under the assumption that curvature terms are negligible.

Monte Carlo design and Fisher scaling with repeated measurements

To confirm the theoretical results in a controlled and interpretable manner, we adopt a Monte Carlo strategy in which the topology of the measurement graph is kept fixed, while the number of independent measurements per directed arc is progressively increased. This choice preserves the sparse and disconnected structure of the original graph and avoids altering its gauge properties, while allowing us to investigate the asymptotic regime predicted by Cramér–Rao theory.
Specifically, for each arc $(i, j) \in \mathcal{E}$ we generate m independent and identically distributed noisy measurements according to the heteroscedastic noise model introduced above. Under the standard *i.i.d.* assumption, the contributions of these measurements to the log-likelihood are additive. As a consequence, the Fisher Information Matrix scales linearly with the number of samples:

$$\mathbf{F}(\Theta)_m \;=\; m\,\mathbf{F}(\Theta)_1,$$

where $\mathbf{F}(\Theta)_1$ denotes the Fisher Information Matrix corresponding to a single measurement per directed arc.
This property allows us to predict *a priori* how the Fisher information evolves as m increases and provides a principled way to drive the estimation problem toward the asymptotic regime in which the intrinsic Cramér–Rao bounds become tight. Rather than increasing the number of nodes—which would change the graph topology and significantly increase computational costs—we enrich the information content of the same graph by increasing the number of measurements per arc. This modeling choice is also more representative of practical pose-graph estimation scenarios, where the same relative transformations are often observed multiple times.

Anchored formulation on $\bar{\mathcal{P}}_A$

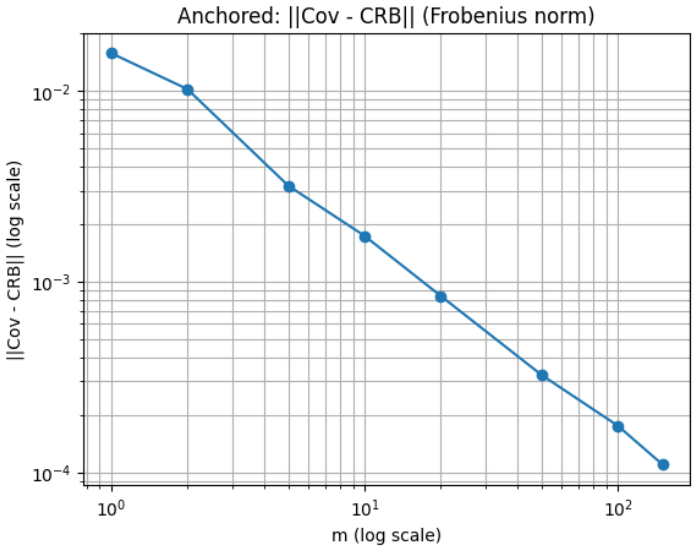
We first consider the anchored formulation of the problem. To remove the gauge freedom and obtain a well-defined estimator, one pose per connected component is fixed, yielding the anchored submanifold

$$\bar{\mathcal{P}}_A \subset \mathrm{SE}(2)^5, \qquad A = \{1, 3\},$$

where the poses T_1 and T_3 are constrained to their ground-truth values. The resulting identifiable parameter space has dimension 9.
In practice, the MLE of Θ is computed by maximizing the log-likelihood while explicitly anchoring the values of T_1 and T_3 in the optimization. For each Monte Carlo trial and each value of m, the remaining poses are estimated, and the intrinsic estimation error is computed in Lie algebra coordinates. From the collection of Monte Carlo samples, the empirical covariance $\widehat{\mathrm{Cov}}_\Theta(\hat{\Theta})$ of the intrinsic MLE error on $\bar{\mathcal{P}}_A$ is estimated.
The empirical covariance is then compared with the anchored intrinsic Cramér–Rao bound given by equation (5.1). To obtain a compact and coordinate-independent validation metric, we analyze the Frobenius norm of the difference between the empirical covariance and the theoretical bound:

$$\left\| \widehat{\mathrm{Cov}}_\Theta(\hat{\Theta}) - \left(\Pi\, \mathbf{F}(\Theta)_m\, \Pi\right)^\dagger \right\|_F,$$

where Π denotes the orthogonal projector onto the tangent space of the anchored submanifold.
Graph (6.8) reports this quantity as a function of the number of measurements per directed arc m, using logarithmic scales on both axes. The norm of the difference decreases rapidly with m, exhibiting an approximately linear decay in log-log scale, which indicates an exponential convergence toward zero. This behavior provides clear empirical evidence of the asymptotic efficiency of the MLE in the anchored setting and validates the quality of the anchored intrinsic Cramér–Rao bound derived in Theorem (5.2.2).



Graph 6.8: Anchored case: Frobenius norm of the difference between the empirical covariance of the intrinsic MLE error and the anchored intrinsic Cramér–Rao bound $\left(\Pi\, \mathbf{F}(\Theta)_m\, \Pi\right)^\dagger$, as a function of the number of measurements per directed arc m (log-log scale).

Once the gauge is removed by restricting the problem to $\bar{\mathcal{P}}_A$, the matrix $\left(\Pi\, \mathbf{F}(\Theta)_m\, \Pi\right)^\dagger$ plays the role of the Fisher-based lower bound on the covariance, and the quantity plotted in Graph (6.8) is precisely the norm of the difference between the left-hand side and the right-hand side of (5.5).

Quotient formulation on $\bar{\mathcal{P}}_{\mathcal{O}}$

We now repeat the same empirical validation in the quotient formulation, where no anchors are imposed on the absolute poses. In this case, the natural parameter space is the quotient manifold

$$\bar{\mathcal{P}}_{\mathcal{O}} = \mathrm{SE}(2)^n / \mathrm{SE}(2)^c, \qquad n = 5,\ c = 2,$$

reflecting the presence of two connected components in the undirected measurement graph. The quotient construction removes the $3c$-dimensional gauge freedom associated with independent rigid motions acting on each component, yielding a 9-dimensional identifiable space.

Neglecting curvature terms, the intrinsic Cramér–Rao bound in this setting is given by equation (5.5), namely

$$\mathrm{Cov}_{\Theta}(\hat{\Theta}) \succeq \mathbf{F}(\Theta)^{\dagger}_{m}.$$

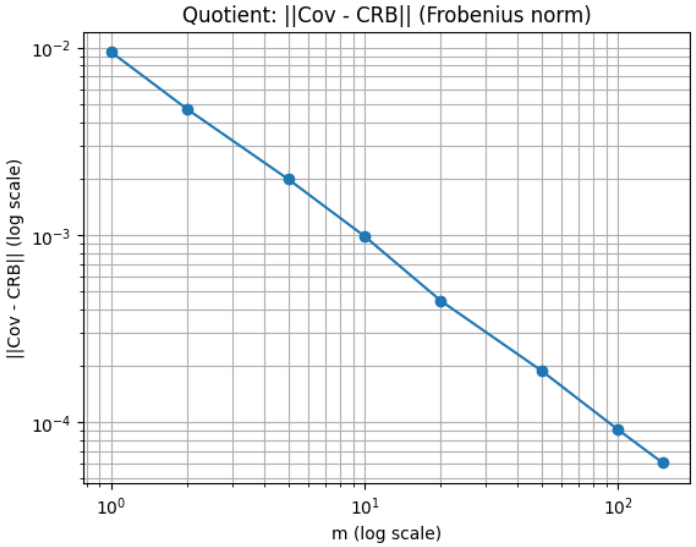
Although the quotient manifold is abstract, the empirical validation can be carried out by working in local coordinates and exploiting the vertical/horizontal decomposition of the tangent space. The log-likelihood is interpreted as a function on $\bar{\mathcal{P}}_{\mathcal{O}}$, and only variations along the horizontal subspace $\mathcal{H}_{\Theta}$ correspond to observable directions. The vertical directions, associated with gauge transformations, are intrinsically uninformative.

For each Monte Carlo trial, the full MLE is first computed in the ambient space $\mathrm{SE}(2)^5$. A right gauge alignment is then applied independently to each connected component in order to obtain a consistent representative of the equivalence class. The intrinsic estimation error is subsequently projected onto the horizontal subspace, yielding a 9-dimensional error vector. Further implementation details are provided in the Appendix.

From the resulting Monte Carlo samples, the empirical covariance on the horizontal subspace is estimated and compared with the quotient intrinsic Cramér–Rao bound (5.5). As in the anchored case, we summarize the discrepancy by the Frobenius norm of the difference between the empirical covariance and the theoretical bound:

$$\left\| \widehat{\mathrm{Cov}}^{\mathcal{H}}_{\Theta}(\hat{\Theta}) - \mathbf{F}(\Theta)^{\dagger}_{m} \right\|_F.$$

The results are shown in Graph (6.9). Also in this case, the norm of the difference decays rapidly as the number of measurements per directed arc increases, exhibiting an approximately linear trend on a logarithmic scale. This exponential convergence confirms that the MLE is asymptotically efficient along all identifiable directions even in the quotient formulation and that the Fisher-based bound (5.5) becomes increasingly tight as m grows.



Graph 6.9: Quotient case: Frobenius norm of the difference between the empirical covariance of the intrinsic MLE error (restricted to the horizontal identifiable directions) and the intrinsic Cramér–Rao bound $\mathbf{F}(\Theta)_m^\dagger$ from (5.5), as a function of the number of measurements per directed arc m (log-log scale).

Chapter 7

Conclusions

The analysis developed throughout this thesis highlights the importance of addressing pose-graph estimation problems within a framework that is consistent with the intrinsic geometry of the underlying parameter space. In estimation problems on Lie groups, ignoring this geometric structure can lead not only to conceptual inconsistencies, but also to misleading interpretations of uncertainty and performance limits.

Within this context, the results presented in this thesis show that maximum likelihood estimation on Lie groups, when formulated in a geometrically consistent manner, provides a remarkably effective tool for pose-graph estimation from noisy relative measurements. Both the analytical developments and the numerical experiments indicate that the MLE is able to exploit the structure of the measurement graph to produce globally consistent estimates, accurately reconstructing relative poses even in sparse configurations. When the intrinsic gauge ambiguity is removed by anchoring at least one pose per connected component, the same framework naturally extends to the reliable estimation of absolute poses, confirming that the loss of identifiability in the unanchored problem is structural rather than algorithmic.

A key outcome of this work is the empirical validation of the intrinsic Cramér–Rao bounds derived for both the anchored and quotient formulations. In both settings, the observed convergence of the empirical covariance of the MLE to the corresponding bounds supports their interpretation as meaningful and tight performance limits. This agreement reinforces the view that intrinsic bounds, when properly formulated on Riemannian manifolds and quotient spaces, offer a faithful description of estimation uncertainty in problems with symmetry and gauge freedom.

Beyond these results, several directions naturally emerge for future investigation. A first extension concerns a more refined treatment of the intrinsic Cramér–Rao bounds by explicitly accounting for curvature terms. While the present analysis focuses on first-order approximations, incorporating curvature-dependent corrections could provide sharper bounds and a deeper understanding of the role played by the geometry of $\mathrm{SE}(d)^n$, especially in regimes of moderate noise.

Closely related to this point is the interpretation of scalar measures of uncertainty. In particular, the role of the trace of the covariance matrix as a proxy for geodesic dispersion deserves further attention, both in the anchored setting and in the quotient formulation, following the approach originally proposed in [19]. A systematic investigation of how such scalar quantities relate to intrinsic distances on the manifold could lead to more geometrically meaningful performance metrics.

The quotient formulation opens an especially interesting perspective in this regard. Rather than interpreting the bound solely in terms of covariance on identifiable directions, one may seek an intrinsic characterization of the limit as the minimal geodesic distance between equivalence classes of the true parameter and of an unbiased estimator. Such an interpretation, analogous to the one proposed by N.Boumal in his synchronization example on SO(3), would further clarify the geometric meaning of the bound and strengthen the conceptual link between estimation theory and Riemannian geometry.

From an application-oriented viewpoint, the proposed framework is well suited to realistic scenarios arising in camera network calibration and multi-view geometry. In particular, configurations in which multiple cameras observe a common object from different viewpoints often give rise to incomplete measurement graphs, since the shape of the object may prevent some cameras from directly observing each other. This naturally leads to sparse and directed measurement graphs with missing edges, precisely of the type considered in this work. In such settings, the ability of the MLE to reconstruct consistent relative poses and to provide rigorous uncertainty quantification through intrinsic bounds is of practical relevance and suggests promising applications in real-world synchronization and calibration problems.

In conclusion, this work highlights how respecting the intrinsic geometry of the parameter space and the symmetry structure of the problem leads to both theoretically sound and practically effective estimation methods. The combination of Lie group modeling, graph-based representations, and intrinsic statistical bounds provides a coherent framework that can be further extended and refined, offering a solid basis for future research at the intersection of geometry, estimation theory, and applications in computer vision.

Appendix A

Simulation Code

This appendix reports the Python script used to generate the simulations and plots presented in Chapter 6.

A.1 Python script

Listing A.1: Python script for case-study

```
"""
Sparse pose-graph estimation on SE(2) with 5 nodes, heteroscedastic noise,
directed incomplete measurement graph, anchored MLE, Fisher Information Matrix,
and Monte Carlo validation of CRB (anchored and quotient).

This version prints the main numerical objects before the related plots:
- T_true, T_ij_true
- injected noise xi_ij
- noisy measurements H_ij
- anchored MLE poses T_MLE and T_ij_MLE
- relative errors per edge and per component
- Fisher matrix at truth, its eigenvalues (showing 6 zeros)
- Monte Carlo summaries per m for anchored and quotient:
  eigenvalues of Sym(Cov - CRB), minimum eigenvalue, trace(Cov), trace(CRB)
- Anchored: || Cov_emp - CRB || (matrix norm, all directions) vs m
- Quotient:  || Cov_emp   - CRB || (matrix norm, all directions) vs m
Expected to decay toward 0 as m increases.

Author: Michele Bertini
"""

import numpy as np
import matplotlib.pyplot as plt
from scipy.linalg import expm
from scipy.optimize import minimize


# ============================
```

```python
# Configuration
# ============================

RNG_SEED = 7
M_DEMO = 1                          # one noisy measurement per directed edge for the demo plots
M_VALUES = [1, 2, 5, 10, 20, 50, 100, 150]
N_MC = 60

PRINT_PREC = 6
ZERO_EIG_TOL = 1e-9


# ============================
# SE(2) utilities
# ============================

def rot2(theta: float) -> np.ndarray:
    c, s = np.cos(theta), np.sin(theta)
    return np.array([[c, -s], [s, c]])

def T_from_params(theta: float, tx: float, ty: float) -> np.ndarray:
    T = np.eye(3)
    T[:2, :2] = rot2(theta)
    T[:2, 2] = [tx, ty]
    return T

def inv_se2(T: np.ndarray) -> np.ndarray:
    R = T[:2, :2]
    t = T[:2, 2]
    Ti = np.eye(3)
    Ti[:2, :2] = R.T
    Ti[:2, 2] = -R.T @ t
    return Ti

def hat_se2(xi: np.ndarray) -> np.ndarray:
    w, vx, vy = xi
    return np.array([[0.0, -w,  vx],
                     [w,   0.0, vy],
                     [0.0, 0.0, 0.0]])

def Exp_SE2(xi: np.ndarray) -> np.ndarray:
    w, vx, vy = xi
    v = np.array([vx, vy])
    T = np.eye(3)
    R = rot2(w)
    T[:2, :2] = R

    if abs(w) < 1e-12:
        V = np.eye(2) + 0.5 * np.array([[0.0, -w],
                                        [w,   0.0]])
    else:
        A = np.sin(w) / w
        B = (1.0 - np.cos(w)) / w
        V = np.array([[A, -B],
                      [B,  A]])
```

```
    T[:2, 2] = V @ v
    return T

def Log_SE2(T: np.ndarray) -> np.ndarray:
    R = T[:2, :2]
    t = T[:2, 2]
    w = np.arctan2(R[1, 0], R[0, 0])

    if abs(w) < 1e-12:
        Vinv = np.eye(2) - 0.5 * np.array([[0.0, -w],
                                           [w,   0.0]])
    else:
        A = np.sin(w) / w
        B = (1.0 - np.cos(w)) / w
        denom = A*A + B*B
        Vinv = (1.0/denom) * np.array([[ A,  B],
                                       [-B,  A]])
    v = Vinv @ t
    return np.array([w, v[0], v[1]])

def Ad_SE2(T: np.ndarray) -> np.ndarray:
    R = T[:2, :2]
    p = T[:2, 2]
    Ad = np.zeros((3, 3))
    Ad[0, 0] = 1.0
    Ad[1, 0] = p[1]
    Ad[2, 0] = -p[0]
    Ad[1:3, 1:3] = R
    return Ad


# ============================
# Right Jacobian (numeric)
# ============================

def Jr_numeric(xi: np.ndarray, eps: float = 1e-8) -> np.ndarray:
    T0 = Exp_SE2(xi)
    T0inv = inv_se2(T0)
    Jr = np.zeros((3, 3))
    for k in range(3):
        d = np.zeros(3)
        d[k] = eps
        Tk = Exp_SE2(xi + d)
        Delta = T0inv @ Tk
        Jr[:, k] = Log_SE2(Delta) / eps
    return Jr

def Jr_inv_numeric(xi: np.ndarray) -> np.ndarray:
    return np.linalg.inv(Jr_numeric(xi))


# ============================
# Problem instance
# ============================

def build_instance():
```

```
    nodes = (1, 2, 3, 4, 5)
    E = [(1, 2), (2, 1), (3, 4), (4, 5), (5, 4), (5, 3)]  # directed sparse set

    T_true = {
        1: T_from_params(theta=0.55,  tx=-0.6, ty=0.2),
        2: T_from_params(theta=1.05,  tx= 0.9, ty=0.1),
        3: T_from_params(theta=-0.75, tx=-0.2, ty=1.2),
        4: T_from_params(theta=0.30,  tx= 1.1, ty=1.6),
        5: T_from_params(theta=-0.15, tx= 2.0, ty=0.9),
    }

    # Consistent with residual r_ij = Log( H_ij T_j T_i^{-1} ), noise-free measurement is T_i T_j^{-1}
    Trel_true = {(i, j): (T_true[i] @ inv_se2(T_true[j])) for (i, j) in E}

    Sigma = {
        (1, 2): np.diag([0.0060, 0.0010, 0.0080]),
        (2, 1): np.diag([0.0020, 0.0090, 0.0010]),
        (3, 4): np.diag([0.0050, 0.0070, 0.0090]),
        (4, 5): np.diag([0.0070, 0.0060, 0.0100]),
        (5, 4): np.diag([0.0090, 0.0080, 0.0060]),
        (5, 3): np.diag([0.0065, 0.0055, 0.0075]),
    }
    Omega = {e: np.linalg.inv(Sigma[e]) for e in E}

    # Connected components in undirected sense
    components = [tuple([1, 2]), tuple([3, 4, 5])]

    return nodes, E, components, T_true, Trel_true, Sigma, Omega


# ============================
# Measurement model and residuals
# ============================

def generate_measurements(rng, E, Trel_true, Sigma, m: int):
    H_list, xi_list = {}, {}
    for (i, j) in E:
        Hijs, xis = [], []
        for _ in range(m):
            xi = rng.multivariate_normal(mean=np.zeros(3), cov=Sigma[(i, j)])
            Hij = expm(hat_se2(xi)) @ Trel_true[(i, j)]
            Hijs.append(Hij)
            xis.append(xi)
        H_list[(i, j)] = Hijs
        xi_list[(i, j)] = xis
    return H_list, xi_list

def residual_ij(Hij: np.ndarray, Ti: np.ndarray, Tj: np.ndarray) -> np.ndarray:
    return Log_SE2(Hij @ Tj @ inv_se2(Ti))

def se2_delta(T_true: np.ndarray, T_est: np.ndarray) -> np.ndarray:
```

```
    return Log_SE2(T_true @ inv_se2(T_est))

def sym(A: np.ndarray) -> np.ndarray:
    return 0.5 * (A + A.T)

def stable_pinv(A: np.ndarray, rcond: float = 1e-12) -> np.ndarray:
    return np.linalg.pinv(A, rcond=rcond)


# ============================
# Plot helpers (poses)
# ============================

def draw_pose(ax, T, label=None, arrow_len=0.35, marker='o'):
    p = T[:2, 2]
    R = T[:2, :2]
    d = (R @ np.array([1.0, 0.0])) * arrow_len
    ax.plot(p[0], p[1], marker=marker)
    ax.quiver(p[0], p[1], d[0], d[1], angles='xy', scale_units='xy',
        scale=1)
    if label is not None:
        ax.text(p[0] + 0.03, p[1] + 0.03, label)

def setup_pose_axes(ax, title):
    ax.set_title(title)
    ax.set_aspect('equal', adjustable='box')
    ax.grid(True)
    ax.set_xlabel("x")
    ax.set_ylabel("y")

def autoset_equal_limits(ax, points_xy: np.ndarray, pad_frac: float =
    0.25, min_pad: float = 0.6):
    x = points_xy[:, 0]
    y = points_xy[:, 1]
    xmin, xmax = float(np.min(x)), float(np.max(x))
    ymin, ymax = float(np.min(y)), float(np.max(y))
    span = max(xmax - xmin, ymax - ymin, 1e-9)
    pad = max(min_pad, pad_frac * span)
    cx = 0.5 * (xmin + xmax)
    cy = 0.5 * (ymin + ymax)
    half = 0.5 * span + pad
    ax.set_xlim(cx - half, cx + half)
    ax.set_ylim(cy - half, cy + half)


# ============================
# MLE solvers
# ============================

def solve_mle_anchor_T1_T3(H_list, E, Omega, T_true, x0=None, maxiter
    =900):
    if x0 is None:
        x0 = np.zeros(9)

    def unpack_x(x):
        T2x = T_from_params(x[0], x[1], x[2])
```

```
        T4x = T_from_params(x[3], x[4], x[5])
        T5x = T_from_params(x[6], x[7], x[8])
        return T2x, T4x, T5x

    cost_hist = []

    def objective(x):
        T2x, T4x, T5x = unpack_x(x)
        T = {1: T_true[1], 2: T2x, 3: T_true[3], 4: T4x, 5: T5x}
        total = 0.0
        for (i, j) in E:
            for Hij in H_list[(i, j)]:
                r = residual_ij(Hij, T[i], T[j])
                total += 0.5 * (r.T @ Omega[(i, j)] @ r)
        cost_hist.append(total)
        return total

    res = minimize(
        objective, x0, method="L-BFGS-B",
        options=dict(maxiter=maxiter, ftol=1e-11, gtol=1e-9, maxfun
            =9000)
    )

    T2_mle, T4_mle, T5_mle = unpack_x(res.x)
    Tmle = {1: T_true[1], 2: T2_mle, 3: T_true[3], 4: T4_mle, 5:
        T5_mle}
    return Tmle, res, np.array(cost_hist)

def solve_mle_full_15(H_list, nodes, E, Omega, x0=None, maxiter=1100):
    if x0 is None:
        x0 = np.zeros(15)
    node_to_idx = {n: k for k, n in enumerate(nodes)}

    def unpack(x):
        T = {}
        for n in nodes:
            k = node_to_idx[n]
            th, tx, ty = x[3*k:3*k+3]
            T[n] = T_from_params(th, tx, ty)
        return T

    def objective(x):
        T = unpack(x)
        total = 0.0
        for (i, j) in E:
            for Hij in H_list[(i, j)]:
                r = residual_ij(Hij, T[i], T[j])
                total += 0.5 * (r.T @ Omega[(i, j)] @ r)
        return total

    res = minimize(
        objective, x0, method="L-BFGS-B",
        options=dict(maxiter=maxiter, ftol=1e-11, gtol=1e-8, maxfun
            =12000)
    )
    return unpack(res.x), res
```

```python


# ============================
# Gauge alignment for disconnected components (quotient case)
# ============================

def align_right_gauge_components(T_est: dict, T_true: dict, components
    , maxiter=300):
    c = len(components)

    node_to_comp = {}
    for k, comp in enumerate(components):
        for node in comp:
            node_to_comp[node] = k

    def objective(rvec_all):
        total = 0.0
        G = []
        for k in range(c):
            gk = rvec_all[3*k:3*k+3]
            G.append(Exp_SE2(gk))
        for node, Tn_est in T_est.items():
            k = node_to_comp[node]
            Tn_al = Tn_est @ G[k]
            e = Log_SE2(T_true[node] @ inv_se2(Tn_al))
            total += 0.5 * (e @ e)
        return total

    res = minimize(
        objective, np.zeros(3*c), method="L-BFGS-B",
        options=dict(maxiter=maxiter, ftol=1e-12, gtol=1e-9, maxfun
           =8000)
    )

    Gopt = []
    for k in range(c):
        gk = res.x[3*k:3*k+3]
        Gopt.append(Exp_SE2(gk))

    Tal = {node: (T_est[node] @ Gopt[node_to_comp[node]]) for node in
       T_est.keys()}
    return Tal, res


# ============================
# Fisher at truth
# ============================

def fisher_full_at_truth(nodes, E, Omega, T_true, Trel_true):
    node_to_idx = {n: k for k, n in enumerate(nodes)}
    dim = 3 * len(nodes)

    def blk(n):
        k = node_to_idx[n]
        return slice(3*k, 3*k+3)

```

```
    F = np.zeros((dim, dim))

    for (i, j) in E:
        Ti = T_true[i]
        Tj = T_true[j]
        Hij = Trel_true[(i, j)]  # noise-free measurement

        r = residual_ij(Hij, Ti, Tj)     # should be ~ 0
        Jr_inv = Jr_inv_numeric(r)       # explicit J_r^{-1}(r)
        A = Jr_inv @ Ad_SE2(Ti)

        J = np.zeros((3, dim))
        J[:, blk(i)] = -A
        J[:, blk(j)] = +A

        F += J.T @ Omega[(i, j)] @ J

    return F


# ============================
# Printing helpers
# ============================

def fmt_mat(M: np.ndarray) -> str:
    return np.array2string(M, precision=PRINT_PREC, suppress_small=
        False)

def print_header(title: str):
    print("\n" + "=" * 80)
    print(title)
    print("=" * 80)

def edge_str(i, j) -> str:
    return f"{i}->{j}"


# ============================
# Main
# ============================

def main():
    np.set_printoptions(precision=PRINT_PREC, suppress=True)
    rng = np.random.default_rng(RNG_SEED)

    nodes, E, components, T_true, Trel_true, Sigma, Omega =
        build_instance()

    # ---------
    # Print: T_true and T_ij_true
    # ---------
    print_header("True poses (T_true)")
    for k in nodes:
        print(f"T_{k}^true =\n{fmt_mat(T_true[k])}\n")

    print_header("True relative transforms (T_ij_true)")
```

```
    for (i, j) in E:
        print(f"T_{edge_str(i,j)}^true =\n{fmt_mat(Trel_true[(i, j)])}\n")

    # ---------
    # 1) Plot: true poses
    # ---------
    fig = plt.figure()
    ax = fig.add_subplot(111)
    for k in nodes:
        draw_pose(ax, T_true[k], label=str(k), marker='o')
    setup_pose_axes(ax, "True poses")
    pts = np.array([T_true[k][:2, 2] for k in nodes])
    autoset_equal_limits(ax, pts)
    plt.show()

    # ---------
    # 2-3) One noisy measurement per edge + print noise and H_ij
    # ---------
    H_demo, xi_demo = generate_measurements(rng, E, Trel_true, Sigma, m=M_DEMO)

    print_header("Injected noise per arc (xi_ij)")
    for (i, j) in E:
        xi = xi_demo[(i, j)][0]
        print(f"xi_{edge_str(i,j)} = {xi}")

    print_header("Noisy measurements per arc (H_ij)")
    for (i, j) in E:
        Hij = H_demo[(i, j)][0]
        print(f"H_{edge_str(i,j)} =\n{fmt_mat(Hij)}\n")

    # Plot: injected noise
    xi_first = np.vstack([xi_demo[e][0] for e in E])
    labels = [edge_str(i, j) for (i, j) in E]
    x = np.arange(len(E))

    fig = plt.figure()
    ax = fig.add_subplot(111)
    ax.plot(x, xi_first[:, 0], marker='o', label=" Delta theta")
    ax.plot(x, xi_first[:, 1], marker='o', label="Delta x")
    ax.plot(x, xi_first[:, 2], marker='o', label="Delta y")
    ax.set_title("Injected noise per arc")
    ax.set_xlabel("Arcs")
    ax.set_ylabel("Noise value")
    ax.set_xticks(x)
    ax.set_xticklabels(labels)
    ax.grid(True)
    ax.legend()
    plt.show()

    # ---------
    # 4) Anchored MLE: print T_MLE and show cost + poses
    # ---------
    Tmle_demo, res_demo, cost_hist = solve_mle_anchor_T1_T3(
        H_demo, E, Omega, T_true, x0=np.zeros(9), maxiter=1100
```

```
458     )
459
460     print_header("Anchored MLE result (T_MLE)")
461     print(f"Optimization success: {res_demo.success} | message: {
            res_demo.message}\n")
462     for k in nodes:
463         print(f"T_{k}^MLE =\n{fmt_mat(Tmle_demo[k])}\n")
464
465     fig = plt.figure()
466     ax = fig.add_subplot(111)
467     ax.plot(cost_hist)
468     ax.set_title("Optimization cost")
469     ax.set_xlabel("Iteration")
470     ax.set_ylabel("Cost")
471     ax.grid(True)
472     plt.show()
473
474     fig = plt.figure()
475     ax = fig.add_subplot(111)
476     for k in nodes:
477         draw_pose(ax, Tmle_demo[k], label=str(k), marker='x')
478     setup_pose_axes(ax, "Estimated poses (anchored MLE)")
479     pts_mle = np.array([Tmle_demo[k][:2, 2] for k in nodes])
480     autoset_equal_limits(ax, pts_mle)
481     plt.show()
482
483     # ---------
484     # 5-6) Demo: T_ij_MLE, errors per arc + components (print + plots)
485     # ---------
486     Trel_mle_demo = {(i, j): (Tmle_demo[i] @ inv_se2(Tmle_demo[j]))
            for (i, j) in E}
487     delta_demo = {(i, j): Log_SE2(Trel_true[(i, j)] @ inv_se2(
            Trel_mle_demo[(i, j)])) for (i, j) in E}
488
489     print_header("Relative transforms from MLE (T_ij_MLE)")
490     for (i, j) in E:
491         print(f"T_{edge_str(i,j)}^MLE =\n{fmt_mat(Trel_mle_demo[(i, j)
                ])}\n")
492
493     print_header("Relative errors per arc")
494     for (i, j) in E:
495         d = delta_demo[(i, j)]
496         print(f"delta_{edge_str(i,j)} = [ Delta theta,  Delta x, Delta
                y] = {d} | norm = {np.linalg.norm(d):.6e}")
497
498     norm_demo = np.array([np.linalg.norm(delta_demo[e]) for e in E])
499
500     fig = plt.figure()
501     ax = fig.add_subplot(111)
502     ax.bar(x, norm_demo)
503     ax.set_title("Relative error norm per arc")
504     ax.set_xlabel("Arcs")
505     ax.set_ylabel("||log error||")
506     ax.set_xticks(x)
507     ax.set_xticklabels(labels)
508     ax.grid(True, axis="y")
```

```
    plt.show()

    Dmat = np.vstack([delta_demo[e] for e in E])
    fig = plt.figure()
    ax = fig.add_subplot(111)
    ax.plot(x, Dmat[:, 0], marker='o', label="Delta theta")
    ax.plot(x, Dmat[:, 1], marker='o', label="Delta x")
    ax.plot(x, Dmat[:, 2], marker='o', label="Delta y")
    ax.set_title("Relative error components")
    ax.set_xlabel("Arcs")
    ax.set_ylabel("Component value")
    ax.set_xticks(x)
    ax.set_xticklabels(labels)
    ax.grid(True)
    ax.legend()
    plt.show()

    # ---------
    # 7) Fisher at truth: print matrix + eigenvalues, count zeros;
       then plot heatmap and spectrum
    # ---------
    F_true = fisher_full_at_truth(nodes, E, Omega, T_true, Trel_true)
    eigF = np.linalg.eigvalsh(sym(F_true))

    print_header("Fisher information matrix at truth (F(Theta_true))")
    print(fmt_mat(F_true))

    print_header("Fisher eigenvalues (sorted)")
    for k, val in enumerate(eigF, start=1):
        print(f"lambda_{k:02d} = {val:.12e}")
    n_zeros = int(np.sum(np.abs(eigF) < ZERO_EIG_TOL))
    print(f"\nZero eigenvalues (|lambda| < {ZERO_EIG_TOL:g}): {n_zeros
       } (expected: 6)")

    fig = plt.figure()
    ax = fig.add_subplot(111)
    im = ax.imshow(F_true, aspect='auto')
    ax.set_title("Fisher information matrix")
    ax.set_xlabel("Index")
    ax.set_ylabel("Index")
    plt.colorbar(im, ax=ax)
    plt.show()

    fig = plt.figure()
    ax = fig.add_subplot(111)
    ax.stem(np.arange(1, 3*len(nodes) + 1), eigF)
    ax.set_title("Fisher eigenvalues")
    ax.set_xlabel("Eigenvalue index")
    ax.set_ylabel("Eigenvalue")
    ax.grid(True)
    plt.show()

    # ---------
    # 8-9) Monte Carlo: anchored CRB validation using Fisher at truth
    # ---------
    free_nodes = [2, 4, 5]
```

```
    node_to_idx = {n: k for k, n in enumerate(nodes)}

    def blk_full(n_):
        k_ = node_to_idx[n_]
        return slice(3*k_, 3*k_ + 3)

    idx_free = []
    for nn in free_nodes:
        s = blk_full(nn)
        idx_free.extend(range(s.start, s.stop))
    idx_free = np.array(idx_free, dtype=int)

    F_free_1 = F_true[np.ix_(idx_free, idx_free)]

    trace_cov_A = []
    trace_crb_A = []
    eigs_diff_A = []
    min_eig_A = []
    norm_diff_A = []
    x0_warm = np.zeros(9)

    print_header("Anchored Monte Carlo summary")
    for m in M_VALUES:
        errs = []
        ok = 0
        for _ in range(N_MC):
            H_list, _ = generate_measurements(rng, E, Trel_true, Sigma
                , m)
            Tmle, res, _ = solve_mle_anchor_T1_T3(H_list, E, Omega,
                T_true, x0=x0_warm, maxiter=700)
            if not res.success:
                Tmle, res, _ = solve_mle_anchor_T1_T3(H_list, E, Omega
                    , T_true, x0=np.zeros(9), maxiter=950)
                if not res.success:
                    continue
            x0_warm = res.x.copy()

            e = np.hstack([se2_delta(T_true[nn], Tmle[nn]) for nn in
                free_nodes])
            errs.append(e)
            ok += 1

        errs = np.array(errs)
        if errs.shape[0] < 10:
            trace_cov_A.append(np.nan)
            trace_crb_A.append(np.nan)
            eigs_diff_A.append(np.full(9, np.nan))
            min_eig_A.append(np.nan)
            norm_diff_A.append(np.nan)
            print(f"m={m:3d} | successful trials={ok:3d} |
                insufficient samples")
            continue

        mu = np.mean(errs, axis=0)
        Xc = errs - mu
        Cov_emp = (Xc.T @ Xc) / (errs.shape[0] - 1)
```

```
        CRB = stable_pinv(m * F_free_1, rcond=1e-12)

        Dm = sym(Cov_emp - CRB)
        eigs = np.linalg.eigvalsh(Dm)

        tr_cov = float(np.trace(Cov_emp))
        tr_crb = float(np.trace(CRB))
        ndiff = float(np.linalg.norm(Dm, ord='fro'))

        trace_cov_A.append(tr_cov)
        trace_crb_A.append(tr_crb)
        eigs_diff_A.append(eigs)
        min_eig_A.append(float(np.min(eigs)))
        norm_diff_A.append(ndiff)

        eigs_str = ", ".join([f"{v:+.3e}" for v in eigs])
        print(f"m={m:3d} | successful trials={ok:3d} | trace(Cov)={
            tr_cov:.3e} | trace(CRB)={tr_crb:.3e} | min eig={min_eig_A
            [-1]:+.3e}")
        print(f"  eigenvalues(Sym(Cov-CRB)) = [{eigs_str}]")

    m_arr = np.array(M_VALUES, dtype=float)
    eigs_diff_A = np.vstack(eigs_diff_A)

    # --- REMOVED: Anchored trace comparison plot ---

    # Anchored eigenvalue difference plot
    fig = plt.figure()
    ax = fig.add_subplot(111)
    for k in range(9):
        ax.plot(m_arr, eigs_diff_A[:, k], marker='o', label=rf"$\
            lambda_{k+1}$")
    ax.axhline(0.0, linewidth=1.0)
    ax.set_xscale("log")
    ax.set_title("Anchored: eigenvalues of Sym(Cov - CRB)")
    ax.set_xlabel("m (log scale)")
    ax.set_ylabel("Eigenvalue")
    ax.grid(True, which="both")
    ax.legend(ncols=3, fontsize=8)
    plt.show()

    # ---------
    # 10-11) Monte Carlo: quotient validation
    # ---------
    evals_F, evecs_F = np.linalg.eigh(sym(F_true))
    U9 = evecs_F[:, -9:]

    trace_cov_Q = []
    trace_crb_Q = []
    eigs_diff_Q = []
    min_eig_Q = []
    norm_diff_Q = []
    x0_warm_full = np.zeros(15)

    print_header("Quotient Monte Carlo summary")
```

```
    for m in M_VALUES:
        errs_full = []
        ok = 0
        for _ in range(N_MC):
            H_list, _ = generate_measurements(rng, E, Trel_true, Sigma
                , m)

            Tmle, res = solve_mle_full_15(H_list, nodes, E, Omega, x0=
                x0_warm_full, maxiter=900)
            if not res.success:
                Tmle, res = solve_mle_full_15(H_list, nodes, E, Omega,
                    x0=np.zeros(15), maxiter=1200)
                if not res.success:
                    continue
            x0_warm_full = res.x.copy()

            Tal, res_align = align_right_gauge_components(Tmle, T_true
                , components, maxiter=260)
            if not res_align.success:
                continue

            e_stack = np.hstack([se2_delta(T_true[nn], Tal[nn]) for nn
                in nodes])
            errs_full.append(e_stack)
            ok += 1

        errs_full = np.array(errs_full)
        if errs_full.shape[0] < 10:
            trace_cov_Q.append(np.nan)
            trace_crb_Q.append(np.nan)
            eigs_diff_Q.append(np.full(9, np.nan))
            min_eig_Q.append(np.nan)
            norm_diff_Q.append(np.nan)
            print(f"m={m:3d} | successful trials={ok:3d} |
                insufficient samples")
            continue

        mu = np.mean(errs_full, axis=0)
        Xc = errs_full - mu
        Cov_full = sym((Xc.T @ Xc) / (errs_full.shape[0] - 1))

        CRB_full = sym(stable_pinv(m * F_true, rcond=1e-12))

        Cov_9 = sym(U9.T @ Cov_full @ U9)
        CRB_9 = sym(U9.T @ CRB_full @ U9)

        Dm = sym(Cov_9 - CRB_9)
        eigs = np.linalg.eigvalsh(Dm)

        tr_cov = float(np.trace(Cov_9))
        tr_crb = float(np.trace(CRB_9))
        ndiff = float(np.linalg.norm(Dm, ord='fro'))

        trace_cov_Q.append(tr_cov)
        trace_crb_Q.append(tr_crb)
        eigs_diff_Q.append(eigs)
```

```
717             min_eig_Q.append(float(np.min(eigs)))
718             norm_diff_Q.append(ndiff)
719 
720             eigs_str = ", ".join([f"{v:+.3e}" for v in eigs])
721             print(f"m={m:3d} | successful trials={ok:3d} | trace(Cov)={
                    tr_cov:.3e} | trace(CRB)={tr_crb:.3e} | min eig={min_eig_Q
                    [-1]:+.3e}")
722             print(f"  eigenvalues(Sym(Cov-CRB)) = [{eigs_str}]")
723 
724     eigs_diff_Q = np.vstack(eigs_diff_Q)
725     # Quotient eigenvalue difference plot
726     fig = plt.figure()
727     ax = fig.add_subplot(111)
728     for k in range(9):
729         ax.plot(m_arr, eigs_diff_Q[:, k], marker='o', label=rf"$\
                lambda_{k+1}$")
730     ax.axhline(0.0, linewidth=1.0)
731     ax.set_xscale("log")
732     ax.set_title("Quotient: eigenvalues of Sym(Cov - CRB)")
733     ax.set_xlabel("m (log scale)")
734     ax.set_ylabel("Eigenvalue")
735     ax.grid(True, which="both")
736     ax.legend(ncols=3, fontsize=8)
737     plt.show()
738 
739     # ---------
740     # Final plots: Frobenius norm of Cov - CRB vs m (anchored and
            quotient)
741     # ---------
742     fig = plt.figure()
743     ax = fig.add_subplot(111)
744     ax.plot(m_arr, norm_diff_A, marker='o')
745     ax.set_xscale("log")
746     ax.set_yscale("log")
747     ax.set_title("Anchored: ||Cov - CRB|| (Frobenius norm)")
748     ax.set_xlabel("m (log scale)")
749     ax.set_ylabel("||Cov - CRB|| (log scale)")
750     ax.grid(True, which="both")
751     plt.show()
752 
753     fig = plt.figure()
754     ax = fig.add_subplot(111)
755     ax.plot(m_arr, norm_diff_Q, marker='o')
756     ax.set_xscale("log")
757     ax.set_yscale("log")
758     ax.set_title("Quotient: ||Cov - CRB|| (Frobenius norm)")
759     ax.set_xlabel("m (log scale)")
760     ax.set_ylabel("||Cov - CRB|| (log scale)")
761     ax.grid(True, which="both")
762     plt.show()
763 
764 
765 if __name__ == "__main__":
766     main()
```

Bibliography

[1] D. S. Dummit and R. M. Foote, *Abstract Algebra*, 3rd ed., Wiley, 2004.

[2] J. J. Rotman, *An Introduction to the Theory of Groups*, 4th ed., Springer, 1995.

[3] S. Lang, *Algebra*, Revised 3rd ed., Springer, 2002.

[4] M. A. Armstrong, *Groups and Symmetry*, Springer, 1988.

[5] J. M. Lee, *Introduction to Smooth Manifolds*, Graduate Texts in Mathematics, vol. 218, Springer, 2013.

[6] J. M. Lee, *Riemannian Manifolds: An Introduction to Curvature*, Graduate Texts in Mathematics, vol. 176, Springer, 1997.

[7] L. W. Tu, *An Introduction to Manifolds*, Springer, 2011.

[8] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*, Academic Press, 1978.

[9] F. W. Warner, *Foundations of Differentiable Manifolds and Lie Groups*, Springer, 1983.

[10] B. C. Hall, *Lie Groups, Lie Algebras, and Representations*, 2nd ed., Springer, 2015.

[11] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, 2008.

[12] G. S. Chirikjian, *Stochastic Models, Information Theory, and Lie Groups*, vols. 1–2, Birkhäuser, 2009–2011.

[13] S.-I. Amari and H. Nagaoka, *Methods of Information Geometry*, Translations of Mathematical Monographs, vol. 191, American Mathematical Society, 2000.

[14] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*, Prentice Hall, 1993.

[15] P. Stoica and T. Marzetta, "Parameter estimation problems with singular information matrices," *IEEE Transactions on Signal Processing*, vol. 49, no. 1, pp. 87–90, 2001.

[16] S. Smith, "Covariance, subspace, and intrinsic Cramér–Rao bounds," *IEEE Transactions on Signal Processing*, vol. 53, no. 5, pp. 1610–1630, 2005.

[17] J. Xavier and V. Barroso, "Intrinsic Variance Lower Bound (IVLB): An extension of the Cramér–Rao bound to Riemannian manifolds," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2004.

[18] J. Xavier and V. Barroso, "The Riemannian geometry of certain parameter estimation problems with singular Fisher information matrices," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 2, pp. 1021–1024, 2004.

[19] N. Boumal, "On Intrinsic Cramér–Rao Bounds for Riemannian Submanifolds and Quotient Manifolds," arXiv preprint arXiv:1305.0041, 2013.

[20] N. Boumal, A. Singer, P.-A. Absil, and V. D. Blondel, "Cramér–Rao bounds for synchronization of rotations," *IEEE Transactions on Signal Processing*, vol. 62, no. 21, pp. 5613–5627, 2014.

[21] R. Tron and R. Vidal, "Barankin, McAulay–Seidman and Cramér–Rao bounds on matrix Lie groups," *IEEE Transactions on Signal Processing*, vol. 60, no. 3, pp. 1063–1079, 2012.

[22] S. Labsir, A. Renaux, J. Vilà-Valls, and E. Chaumette, "Barankin, McAulay–Seidman and Cramér–Rao bounds on matrix Lie groups," *Automatica*, vol. 156, p. 111199, 2023.

[23] C. Chahbazian, K. Dahia, N. Merlinge, B. Winter-Bonnet, A. Blanc, and C. Musso, "Recursive posterior Cramér–Rao lower bound on Lie groups," *Signal Processing*, vol. 196, p. 108512, 2022.

[24] X. Pennec, "Probabilities and statistics on Riemannian manifolds: Basic tools for geometric measurements," in *Proceedings of the IEEE Workshop on Nonlinear Signal and Image Processing*, 1999.

[25] X. Pennec, P. Fillard, and N. Ayache, "A Riemannian framework for tensor computing," *International Journal of Computer Vision*, vol. 66, no. 1, pp. 41–66, 2006.

[26] R. Bhattacharya and V. Patrangenaru, "Large sample theory of intrinsic and extrinsic sample means on manifolds," *The Annals of Statistics*, vol. 31, no. 1, pp. 1–29, 2003.

[27] N.-D. Ho and P. Van Dooren, "On the pseudo-inverse of the Laplacian of a bipartite graph," *Applied Mathematics Letters*, vol. 18, no. 8, pp. 917–922, 2005.

[28] A. Singer, "Angular synchronization by eigenvectors and semidefinite programming," *Applied and Computational Harmonic Analysis*, vol. 30, no. 1, pp. 20–36, 2011.

[29] S. D. Howard, D. Cochran, W. Moran, and F. R. Cohen, "Estimation and registration on graphs," arXiv preprint arXiv:1010.2983, 2010.

[30] T. D. Barfoot, *State Estimation for Robotics*, Cambridge University Press, 2017.

[31] G. Grisetti, R. Kümmerle, C. Stachniss, and W. Burgard, "A tutorial on graph-based SLAM," *IEEE Robotics & Automation Magazine*, 2010.

[32] R. Smith, M. Self, and P. Cheeseman, "Estimating uncertain spatial relationships in robotics," in *Proceedings of the IEEE International Conference on Robotics and Automation*, 1987.

[33] A. Barrau and S. Bonnabel, "The intrinsic filtering on Lie groups," *IEEE Transactions on Automatic Control*, vol. 62, no. 12, pp. 6349–6364, 2017.

[34] A. Barrau and S. Bonnabel, "The invariant extended Kalman filter as a stable observer," *IEEE Transactions on Automatic Control*, vol. 65, no. 4, pp. 1360–1375, 2020.

[35] A. Barrau and S. Bonnabel, "An intrinsic approach to invariant filtering on Lie groups," *IEEE Control Systems Magazine*, vol. 37, no. 6, pp. 81–96, 2017.