Supplementary: Graph-based SLAM-Aware Exploration with Prior Topo-metric Information

Date: September 2, 2023

Abstract

This note provides the detailed derivation of the Equation (5) in the paper "*Graph-based SLAM-Aware Exploration with Prior Topo-metric Information*", which relates the pose graph reliability in SLAM with the pose graph topology. It is important to note that the supplemental materials are not considered contributions of this paper; rather, they are intended to offer a more comprehensive background and facilitate a better understanding. The Equation (5) is replicated as follows:

$$\mathbb{I}(\boldsymbol{x}) = \mathbf{H} = \frac{1}{2}\sum_{k=1}^{m} \mathbf{J}^\top \boldsymbol{\Sigma}^{-1}\mathbf{J} = \frac{1}{2}\sum_{k=1}^{m} \mathbf{B}_k \mathbf{B}_k^\top \otimes \widetilde{\boldsymbol{\Sigma}}_k^{-1},$$

1 Relationship between Pose Graph Optimization and Graph Topology

1.1 Preliminary of Lie Algebra

The capitalized Exp and Log maps are convenient shortcuts to map vector elements $\boldsymbol{\tau} \in \mathbb{R}^m$ ($\widetilde{=} T_{\mathcal{E}}\mathcal{M}$, the tangent space of manifold $\mathcal{M}$ at identity $\mathcal{E}$) directly with elements $\mathcal{X} \in \mathcal{M}$.

$$\begin{aligned} \mathrm{Exp}: \quad & \mathbb{R}^m \to \mathcal{M} \quad ; \quad \boldsymbol{\tau} \mapsto \mathcal{X} = \mathrm{Exp}(\boldsymbol{\tau}) \\ \mathrm{Log}: \quad & \mathcal{M} \to \mathbb{R}^m \quad ; \quad \mathcal{X} \mapsto \boldsymbol{\tau} = \mathrm{Log}(\mathcal{X}) \end{aligned} \tag{1}$$

We also have:

$$\mathcal{X} = \mathrm{Exp}(\boldsymbol{\tau}) \triangleq \exp\left(\boldsymbol{\tau}^\wedge\right) \tag{2}$$

$$\boldsymbol{\tau} = \mathrm{Log}(\mathcal{X}) \triangleq \log(\mathcal{X})^\vee. \tag{3}$$

Note here the definition of $\exp$ and $\log$ on the matrix are clear, so both of them only take effect on the matrix, rather than on the vector.

1.2 Formulation of Pose Graph Optimization

Pose graph is the simplest form of back-end optimization in SLAM, but it is efficient and reliable. Here we derive the uncertainty evaluation of the robot's pose graph in SLAM, especially establishing the relationship between the uncertainty of the pose graph and the connectivity of the pose graph. The uncertainty of the pose graph actually means how reliable are the poses in the pose graph, after applying nonlinear optimization like Gauss-Newton or LM methods. Intuitively, this requires each pose in the pose graph should be well-constrained.

The pose graph optimization problem can be formulated as following equation:

$$\begin{aligned} \boldsymbol{x}^* &= \arg\min_{\boldsymbol{x}} F(\boldsymbol{x}), \\ \text{s.t.} \quad F(\boldsymbol{x}) &= \tfrac{1}{2}\sum_{k=1}^{m} F_k(\boldsymbol{x}) = \tfrac{1}{2}\sum_{j=1}^{m} \boldsymbol{e}_k^\top(\boldsymbol{x})\boldsymbol{\Sigma}_k^{-1}\boldsymbol{e}_k(\boldsymbol{x}) \end{aligned} \tag{4}$$

here $\boldsymbol{x}$ represents a column vector of robot poses, $F_k(\boldsymbol{x})$ is the error corresponding to edge $edge_k$ in pose graph $\mathcal{G}$, and it is computed by the difference between measurement $\boldsymbol{T}_{ij}$ and the value of $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ in $\boldsymbol{x}$, where $edge_k = \langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle$. $\boldsymbol{\Sigma}_k$ is the covariance of measurement $\boldsymbol{z}_k$. In 3D case, each $\boldsymbol{x}_i$ can be represented by a transformation matrix $\boldsymbol{T}_i$. The $\boldsymbol{e}_k(\boldsymbol{x})$ can be computed as:

$$\boldsymbol{e}_k(\boldsymbol{x}) = \ln\left(\boldsymbol{T}_{ij}^{-1}\boldsymbol{T}_i^{-1}\boldsymbol{T}_j\right)^\vee \tag{5}$$

Here $\boldsymbol{T}_i^{-1}\boldsymbol{T}_j$ is the expected relative measurement according to current value of T_i and T_j. However, we know that both $\boldsymbol{T}_i$ and $\boldsymbol{T}_j$ have some accumulated error in SLAM odometry, so the actual measurement $\boldsymbol{T}_{ij}$ is slightly different from $\boldsymbol{T}_i^{-1}\boldsymbol{T}_j$. Such difference forms the error term $\boldsymbol{e}_k(\boldsymbol{x})$ in Eq. (4). Note in Eq. (5), $\ln$ is taken over the transformation matrix $\boldsymbol{T}_{ij}^{-1}\boldsymbol{T}_i^{-1}\boldsymbol{T}_j$, and then the resulting matrix is transformed into vector space by $\vee$.

First, let's describe the mathematical modeling of the pose graph optimization problem. So for Eq. (5), we want to optimize the two poses $\boldsymbol{T}_i$ and $\boldsymbol{T}_j$, so that $\boldsymbol{e}_k(\boldsymbol{x})$ is minimized. In order to take the derivative, we use Lie algebra by applying left disturbance to these two transformation matrixes. So we have:

$$\boldsymbol{T}_i \rightarrow \exp(\xi_i^\wedge)\boldsymbol{T}_i, \quad \boldsymbol{T}_j \rightarrow \exp(\xi_j^\wedge)\boldsymbol{T}_j \tag{6}$$

Note here the perturbance ξ_i and ξ_j is defined in the world frame, because it pre-multiplies $\boldsymbol{T}_i$ and $\boldsymbol{T}_j$ (which means we can write as $\xi_i = \xi_{wi}$). Now Eq. (5) becomes to:

$$\boldsymbol{e}_k(\boldsymbol{x}) = \ln\left(\boldsymbol{T}_{ij}^{-1}\boldsymbol{T}_i^{-1}\exp(-\xi_i^\wedge)\exp(\xi_j^\wedge)\boldsymbol{T}_j\right)^\vee \tag{7}$$

We want all disturbance terms move to the left or right side for better handling. This can be achieved by applying adjoint property of the Lie algebra, i.e.,

$$\exp\left((\mathrm{Ad}_T\,\xi)^\wedge\right) = \boldsymbol{T}\exp(\xi^\wedge)\boldsymbol{T}^{-1} \tag{8}$$

And further, we have:

$$\exp(\xi^\wedge)\boldsymbol{T} = \boldsymbol{T}\exp\left((\mathrm{Ad}_{\boldsymbol{T}^{-1}}\,\xi)^\wedge\right) \tag{9}$$

By using this, we can move all disturbance term in Eq. (7) to the right side of $\boldsymbol{T}_j$. So we get:

$$
\begin{aligned}
\boldsymbol{e}_k(\boldsymbol{x}) = \hat{\boldsymbol{e}}_{ij} &= \ln\left(\boldsymbol{T}_{ij}^{-1}\boldsymbol{T}_i^{-1}\exp\left((-\boldsymbol{\xi}_i)^\wedge\right)\exp\left(\boldsymbol{\xi}_j^\wedge\right)\boldsymbol{T}_j\right)^\vee \\
&= \ln\left(\boldsymbol{T}_{ij}^{-1}\boldsymbol{T}_i^{-1}\boldsymbol{T}_j\exp\left(\left(-\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_i\right)^\wedge\right)\exp\left(\left(\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_j\right)^\wedge\right)\right)^\vee \\
&\approx \ln\left(\boldsymbol{T}_{ij}^{-1}\boldsymbol{T}_i^{-1}\boldsymbol{T}_j\left[\boldsymbol{I} - \left(\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_i\right)^\wedge + \left(\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_j\right)^\wedge\right]\right)^\vee \\
&\approx \boldsymbol{e}_{ij} + \frac{\partial \boldsymbol{e}_k(\boldsymbol{x})}{\partial \boldsymbol{\xi}_i}\boldsymbol{\xi}_i + \frac{\partial \boldsymbol{e}_k(\boldsymbol{x})}{\partial \boldsymbol{\xi}_j}\boldsymbol{\xi}_j
\end{aligned}
\tag{10}
$$

Step 2 to 3: In the 2nd step, the two terms $\exp\left(\left(-\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_i\right)^\wedge\right)$ and $\exp\left(\left(\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_j\right)^\wedge\right)$ are close to identity matrix, and so can be approximated by the first order Taylor expansion:

$$
\exp\left(\left(-\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_i\right)^\wedge\right) = \boldsymbol{I} + \left(-\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_i\right)^\wedge \tag{11}
$$

$$
\exp\left(\left(\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_j\right)^\wedge\right) = \boldsymbol{I} + \left(\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_j\right)^\wedge \tag{12}
$$

Taken the multiplication of these two terms, we have:

$$
\begin{aligned}
\exp\left(\left(-\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_i\right)^\wedge\right)\exp\left(\left(\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_j\right)^\wedge\right) &= \boldsymbol{I} + \left(-\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_i\right)^\wedge + \left(\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_j\right)^\wedge \\
&\quad + \left(-\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_i\right)^\wedge\left(\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_j\right)^\wedge
\end{aligned}
\tag{13}
$$

The second-order term is omitted in 3rd step of Eq. (10).

Step 3 to 4: In the 4th step, first we define the Lie algebra corresponds to $\boldsymbol{T}_{ij}^{-1}\boldsymbol{T}_i^{-1}\boldsymbol{T}_j$ to be e_{ij}. So we have:

$$
\boldsymbol{T}_{ij}^{-1}\boldsymbol{T}_i^{-1}\boldsymbol{T}_j = \exp\left(e_{ij}^\wedge\right) \tag{14}
$$

We also define:

$$
\exp\left(x^\wedge\right) = \boldsymbol{I} + \left(-\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_i\right)^\wedge + \left(\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_j\right)^\wedge \tag{15}
$$

So we get:

$$
\begin{aligned}
\boldsymbol{e}_k(\boldsymbol{x}) = \hat{\boldsymbol{e}}_{ij} &\approx \ln\left(\boldsymbol{T}_{ij}^{-1}\boldsymbol{T}_i^{-1}\boldsymbol{T}_j\left[\boldsymbol{I} - \left(\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_i\right)^\wedge + \left(\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_j\right)^\wedge\right]\right)^\vee \\
&= \ln\left(\exp\left(e_{ij}^\wedge\right)\exp\left(x^\wedge\right)\right)^\vee
\end{aligned}
\tag{16}
$$

Here $x^\wedge$ can be treated as smaller term in the equation, so we obtain following according to the BCH right-multiplication approximation formula:

$$
\hat{e}_{ij} = \ln\left(\exp\left(e_{ij}^\wedge\right)\exp\left(x^\wedge\right)\right)^\vee \approx \mathcal{J}_r^{-1}\left(e_{ij}\right)x + e_{ij} \tag{17}
$$

Let's further evaluate the $x^\wedge$ in Eq. (15). According to the Taylor expansion of $\ln$ function, we have:

$$
\ln(A) = (A - \boldsymbol{I}) - \frac{(A - \boldsymbol{I})^2}{2} + \frac{(A - \boldsymbol{I})^3}{3} + \cdots \tag{18}
$$

We have:

$$
\begin{aligned}
x^\wedge &= \ln\left(\boldsymbol{I} + \left(-\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_i\right)^\wedge + \left(\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_j\right)^\wedge\right) \\
&\approx \left(-\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_i\right)^\wedge + \left(\operatorname{Ad}_{\boldsymbol{T}_j^{-1}}\boldsymbol{\xi}_j\right)^\wedge
\end{aligned}
\tag{19}
$$

So we have:

$$x \approx -\operatorname{Ad}_{\boldsymbol{T}_j^{-1}} \boldsymbol{\xi}_i + \operatorname{Ad}_{\boldsymbol{T}_j^{-1}} \boldsymbol{\xi}_j \tag{20}$$

Take into Eq. (17), we have:

$$\begin{aligned} \hat{e}_{ij} &\approx \mathcal{J}_r^{-1}(e_{ij})\, x + e_{ij} \\ &= e_{ij} \underbrace{-\mathcal{J}_r^{-1}(e_{ij}) \operatorname{Ad}_{\boldsymbol{T}_j^{-1}}}_{\frac{\partial \boldsymbol{e}_{ij}}{\partial \boldsymbol{\xi}_i}} \boldsymbol{\xi}_i + \underbrace{\mathcal{J}_r^{-1}(e_{ij}) \operatorname{Ad}_{\boldsymbol{T}_j^{-1}}}_{\frac{\partial \boldsymbol{e}_{ij}}{\partial \boldsymbol{\xi}_j}} \boldsymbol{\xi}_j \\ &= \boldsymbol{e}_{ij} + \frac{\partial \boldsymbol{e}_{ij}}{\partial \boldsymbol{\xi}_i} \boldsymbol{\xi}_i + \frac{\partial \boldsymbol{e}_{ij}}{\partial \boldsymbol{\xi}_j} \boldsymbol{\xi}_j \end{aligned} \tag{21}$$

Note here $\frac{\partial \boldsymbol{e}_{ij}}{\partial \boldsymbol{\xi}_i}$ and $\frac{\partial \boldsymbol{e}_{ij}}{\partial \boldsymbol{\xi}_j}$ differ only in sign. Note in $SE(3)$, the adjoint matrix is a 6×6 matrix with the form:

$$\boldsymbol{Ad}(T) = \begin{bmatrix} R & t^\wedge R \\ 0 & R \end{bmatrix} \tag{22}$$

Up to now, we formulate each error term as a linear function of $\boldsymbol{\xi}_i$ and $\boldsymbol{\xi}_j$, according to Eq. (10). Then we can use the nonlinear optimization methods to iteratively update $\boldsymbol{\xi}_i$ (and so $\exp(\xi_i^\wedge) T_i$) to minimize the total error in the pose graph.

For $\mathcal{J}_r^{-1}(e_{ij})$ in above equation, it is the Jacobi $\mathcal{J}_r$ of e_{ij} in BCH approximation. Typically, it is approximated as:

$$\mathcal{J}_r^{-1}(e_{ij}) \approx \boldsymbol{I} + \frac{1}{2} \begin{bmatrix} \phi_e^\wedge & \rho_e^\wedge \\ 0 & \phi_e^\wedge \end{bmatrix} \tag{23}$$

If we take $\mathcal{J}_r^{-1}(e_{ij}) \approx \boldsymbol{I}$, then the Eq. (21) will be the same as in [Placed and Castellanos 2023].

The objective is to minimize the Eq. (4) so that measurement error is minimized. It becomes to:

$$\begin{aligned} F(\boldsymbol{x}) &= \frac{1}{2} \sum_{k=1}^{m} \boldsymbol{e}_k^\top(\boldsymbol{x}) \boldsymbol{\Sigma}_k^{-1} \boldsymbol{e}_k(\boldsymbol{x}) \\ &= \frac{1}{2} \sum_{k=1}^{m} \left[e_k - \mathcal{J}_r^{-1}(e_k) \operatorname{Ad}_{\boldsymbol{T}_j^{-1}} \boldsymbol{\xi}_i + \mathcal{J}_r^{-1}(e_k) \operatorname{Ad}_{\boldsymbol{T}_j^{-1}} \boldsymbol{\xi}_j \right]^\top \boldsymbol{\Sigma}_k^{-1} \\ &\quad \left[e_k - \mathcal{J}_r^{-1}(e_k) \operatorname{Ad}_{\boldsymbol{T}_j^{-1}} \boldsymbol{\xi}_i + \mathcal{J}_r^{-1}(e_k) \operatorname{Ad}_{\boldsymbol{T}_j^{-1}} \boldsymbol{\xi}_j \right] \end{aligned} \tag{24}$$

If we represent $-\mathcal{J}_r^{-1}(e_k) \operatorname{Ad}_{\boldsymbol{T}_j^{-1}} \boldsymbol{\xi}_i + \mathcal{J}_r^{-1}(e_k) \operatorname{Ad}_{\boldsymbol{T}_j^{-1}} \boldsymbol{\xi}_j$ as

$$\begin{bmatrix} 0 & \dots & -\mathcal{J}_r^{-1}(e_k) \operatorname{Ad}_{\boldsymbol{T}_j^{-1}} & \dots & \mathcal{J}_r^{-1}(e_k) \operatorname{Ad}_{\boldsymbol{T}_j^{-1}} & \dots & 0 \end{bmatrix}_{6 \times 6m} \begin{bmatrix} 0 \\ \dots \\ \boldsymbol{\xi}_i \\ \dots \\ \boldsymbol{\xi}_j \\ \dots \\ 0 \end{bmatrix}_{6m \times 1} = \mathbf{Ad}^k \boldsymbol{\xi} \tag{25}$$

We can write Eq. (24) as:

$$
\begin{aligned}
F(\boldsymbol{x}) &= \frac{1}{2}\sum_{k=1}^{m}\left(e_k^\top \boldsymbol{\Sigma}_k^{-1} e_k + 2e_k^\top \boldsymbol{\Sigma}_k^{-1}\,\mathrm{Ad}^k\,\boldsymbol{\xi} + \left(\mathrm{Ad}^k\,\boldsymbol{\xi}\right)^\top \boldsymbol{\Sigma}_k^{-1}\left(\mathrm{Ad}^k\,\boldsymbol{\xi}\right)\right)\\
&= \frac{1}{2}\sum_{k=1}^{m}\left(e_k^\top \boldsymbol{\Sigma}_k^{-1} e_k + 2e_k^\top \boldsymbol{\Sigma}_k^{-1}\,\mathrm{Ad}^k\,\boldsymbol{\xi} + \boldsymbol{\xi}^\top (\mathrm{Ad}^k)^\top \boldsymbol{\Sigma}_k^{-1}\,\mathrm{Ad}^k\,\boldsymbol{\xi}\right)
\end{aligned}
\tag{26}
$$

In the above formulation, the Hessian matrix is:

$$
\begin{aligned}
\mathbf{H} &= \frac{1}{2}\sum_{k=1}^{m}(\mathbf{Ad}^k)^\top \boldsymbol{\Sigma}_k^{-1}\mathbf{Ad}^k\\
&= \frac{1}{2}\sum_{k=1}^{m}\boldsymbol{E}_k \otimes (\mathcal{J}_r^{-1}(e_k)\,\mathrm{Ad}_{\boldsymbol{T}_j^{-1}})^\top \boldsymbol{\Sigma}_k^{-1}(\mathcal{J}_r^{-1}(e_k)\,\mathrm{Ad}_{\boldsymbol{T}_j^{-1}})\\
&= \frac{1}{2}\sum_{k=1}^{m}\boldsymbol{E}_k \otimes \widetilde{\boldsymbol{\Sigma}}_k^{-1}
\end{aligned}
\tag{27}
$$

where $\boldsymbol{E}_j$ is the Laplacian factor of the j-th edge. Note here the disturbance $\boldsymbol{\xi}$ is defined in world coordinate, so we have adjoint terms $\mathrm{Ad}_{\boldsymbol{T}_j^{-1}}$ in H. Here $\mathcal{J}_r^{-1}(e_k)$ can be approximated as $\mathbb{I}_{6\times 6}$ according to Eq. (23).

According to the definition of Fisher Information Matrix (FIM), $\boldsymbol{H}$ calculated at the optimal solution of Eq. (4) is the FIM. In fact, we can only get the MAP solution of Eq. (4), and the obtained $\boldsymbol{H}$ thus only approximates the actual FIM [Kitanov and Indelman 2019].

Bibliography

[1] Andrej Kitanov and Vadim Indelman. *Topological Information-Theoretic Belief Space Planning with Optimality Guarantees*. _eprint: 1903.00927. 2019.

[2] Julio A. Placed and José A. Castellanos. "A General Relationship Between Optimality Criteria and Connectivity Indices for Active Graph-SLAM". In: *IEEE Robotics and Automation Letters* 8.2 (2023), pp. 816–823. DOI: 10.1109/LRA.2022.3233230.