

Predicting Objective Function Change in Pose-Graph Optimization

Fang Bai¹, Teresa Vidal-Calleja¹, Shoudong Huang¹, and Rong Xiong²

Abstract—Robust online incremental SLAM applications require metrics to evaluate the impact of current measurements. Despite its prevalence in graph pruning, information-theoretic metrics solely are insufficient to detect outliers. The optimal value of the objective function is a better choice to detect outliers but cannot be computed unless the problem is solved. In this paper, we show how the objective function change can be predicted in an incremental pose-graph optimization scheme, without actually solving the problem. The predicted objective function change can be used to guide online decisions or detect outliers. Experiments validate the accuracy of the predicted objective function, and an application to outlier detection is also provided, showing its advantages over M-estimators.

I. INTRODUCTION

The simultaneous localization and mapping problem (SLAM) [1] is to find a consistent representation of robot poses given the sensory data. SLAM is fundamental to allow full autonomous robot applications, for its ability to provide accurate localization in absence of the global reference. The pose-graph optimization (PGO) [2] is a compact formulation of SLAM, where all the features/landmarks are fused into the relative pose constraints. Recent techniques based on sparse linear algebra [3], or Riemannian trust region methods [4] can provide efficient solutions to the PGO problem. However, the robustness and reliability still remain to be an issue [1].

To achieve robustness and reliability, one of the key problems is how to formulate a metric that can be used to quantify the impact caused by a certain measurement, especially in an online incremental algorithm. The most widely used metric is the information-theoretic metric [5], which have been used in active decision making [6][7], graph pruning [8][9], constructing sparse pose-graph in online applications [10][11] and so forth. However, the information-theoretic metric does not explicitly take the measurement error into consideration, thus is not sufficient to detect incorrect loop-closures (i.e., outliers). Please see Section VII for a detailed discussion on this.

The optimal value of the objective function of the optimization problem, which is attained at the global minimum, can be used as a metric to benchmark the quality of the solution [12][13]. The main issue of this metric is that it cannot be obtained unless you solve the problem, see [14], [15] for a use of outlier detection. This is extremely undesirable, since in real applications, we want to compute

the metric prior to the solving of the problem, thus it can help making better decisions, rather than use the metric to “regret” mistakes after taking an action. Thus, the optimal value of the objective function is mostly used as a posterior metric to check the validity of the solution. Besides, if the computed solution is considered invalid according to the corresponding objective function, then all the computational resources used to calculate it are wasted.

The focus of this paper is how to predict/compute the optimal value of the objective function without solving the problem. Thus the objective function can be used as a metric to help making decisions. In particular, we focus on an online incremental PGO problem, where the active decision making based on the predicted objective function is more valuable. In this scenario, the PGO problem is incrementally augmented through time with new measurements. The task is how to predict/compute the change of the optimal value of the objective function after including the new measurements.

Contributions: In this paper, we propose a metric to predict the objective function change after including new measurements into the PGO problem. An application of using the predicted objective function change to detect outliers is also provided. Although we take PGO as an instance, it is possible to extend the technique to other least squares optimization problems.

This paper is organized as follows. Section II briefly introduces the PGO problem. The exact metric to predict the objective function change in a linear least squares optimization problem (LLS) is proved in Section III. The extension to nonlinear cases is discussed in Section IV. The prediction of objective function change in PGO is presented in Section V. Experimental evaluation on the accuracy of the predicted objective function, and an application to detect outliers are presented in Section VI. Discussions are provided in Section VII, and Section VIII concludes the paper.

Notations and Preliminaries: The upper-case letter and the bold font of lower-case letter are used to represent a matrix and a vector respectively. We denote the squared Mahalanobis distance as $\|\mathbf{e}\|_S^2 = \mathbf{e}^\top \Sigma^{-1} \mathbf{e}$, and use $stack\{H_i\}_{i=m}^n$ to describe the matrix $[H_m^\top \cdots H_n^\top]^\top$. We denote $\text{Exp}(\cdot)$, $\text{Log}(\cdot)$, $\text{Ad}(\cdot)$, $J_l(\cdot)$, $J_r(\cdot)$ as the exponential mapping, logarithm mapping, adjoint, left-hand Jacobian and right-hand Jacobian of $\text{SE}(3)$, respectively. An accessible introduction to these concepts are provided in [16][17].

II. POSE-GRAPH SLAM FORMULATION

Denote the robot/sensor pose sequence to be estimated as $T_i \in \text{SE}(3)$, $i \in [1, n]$. The relative pose measurement model

¹Fang Bai, Teresa Vidal-Calleja and Shoudong Huang are with the Centre for Autonomous Systems (CAS), University of Technology Sydney, Australia. Email: Fang.Bai@student.uts.edu.au, {Teresa.VidalCalleja, Shoudong.Huang}@uts.edu.au.

²Rong Xiong is with the State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou, China. Email: rxiong@zju.edu.cn.

is defined to be

$$T_{i,j} = T_i^{-1} \cdot T_j \cdot \text{Exp}(\epsilon_{i,j}), \quad \epsilon_{i,j} \sim \mathcal{N}(\mathbf{0}, \tilde{\Sigma}_{i,j})$$

where the uncertainty is modeled by a zero-mean Gaussian concentrated at $T_i^{-1}T_j$.

Given a collection of noisy measurements $\tilde{T}_{i,j}$ (along with their covariances $\tilde{\Sigma}_{i,j}$ to model the uncertainty), the vector space of the measurement error $\tau_{i,j} \in \mathbb{R}^6$ is obtained through the logarithm mapping

$$\tau_{i,j} = \text{Log}(\tilde{T}_{i,j}^{-1} \cdot T_i^{-1} \cdot T_j).$$

The cost function of the pose-graph SLAM problem is to minimize the summation of a set of Mahalanobis distances in terms of the measurement error

$$\min \sum_{i,j} \tau_{i,j}^\top \cdot \tilde{\Sigma}_{i,j}^{-1} \cdot \tau_{i,j}.$$

The problem itself is nondeterministic, and usually an anchor is used to fix one of the poses, for example, fixing the first pose to the identity of SE(3).

III. PREDICT OBJECTIVE FUNCTION CHANGE IN THE LINEAR LEAST SQUARES PROBLEM

In this section, we will show that the objective function change after including new measurements in LLS can be exactly computed using the solution (along with its covariance) of the primal problem and the new measurement data.

A. The Primal and Augmented Problem

The linear least squares problem is obtained through a linear Gaussian measurement model, which is formed as

$$\mathbf{z}_i = H_i \mathbf{x} - \mathbf{h}_i + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(\mathbf{0}, \tilde{\Sigma}_i).$$

Here H_i and \mathbf{h}_i are given parameters of the model. The variable \mathbf{z}_i can be measured through sensory technologies, which usually yields measurement data $\tilde{\mathbf{z}}_i$ along with a covariance $\tilde{\Sigma}_i$ indicating the reliability of the measurement.

Problem 1. (Primal Linear Least Squares Optimization Problem): Given a collection of noisy measurements $\{\tilde{\mathbf{z}}_i, \tilde{\Sigma}_i\}_{i=1}^{n_1}$, the linear least squares problem to estimate \mathbf{x} is

$$\mathbf{x}^* = \arg \min \sum_{i=1}^{n_1} \|H_i \mathbf{x} - \mathbf{h}_i - \tilde{\mathbf{z}}_i\|_{\tilde{\Sigma}_i}^2,$$

which can be written in a compact form as

$$\mathbf{x}^* = \arg \min \|A_1 \mathbf{x} - \mathbf{b}_1\|_{\Sigma_1}^2$$

where $A_1 = \text{stack}\{H_i\}_{i=1}^{n_1}$, $\mathbf{b}_1 = \text{stack}\{\mathbf{h}_i + \tilde{\mathbf{z}}_i\}_{i=1}^{n_1}$, and $\Sigma_1 = \text{blkdiag}\{\tilde{\Sigma}_i\}_{i=1}^{n_1}$.

Problem 2. (Augmented Linear Least Squares Optimization Problem): If additional measurements $\{\tilde{\mathbf{z}}_i, \tilde{\Sigma}_i\}_{i=n_1+1}^{n_2}$ are obtained, similarly, by stacking the measurements, we obtain an augmented linear least squares optimization problem as

$$\mathbf{x}^{**} = \arg \min \|A_1 \mathbf{x} - \mathbf{b}_1\|_{\Sigma_1}^2 + \|A_2 \mathbf{x} - \mathbf{b}_2\|_{\Sigma_2}^2$$

where $A_2 = \text{stack}\{H_i\}_{i=n_1+1}^{n_2}$, $\mathbf{b}_2 = \text{stack}\{\mathbf{h}_i + \tilde{\mathbf{z}}_i\}_{i=n_1+1}^{n_2}$, and $\Sigma_2 = \text{blkdiag}\{\tilde{\Sigma}_i\}_{i=n_1+1}^{n_2}$.

B. Predict Objective Function Change

For a linear least squares problem, the solution can be computed in a closed form, with the covariance of the solution obtained through linear approximation of Gaussian.

The optimal solution \mathbf{x}^* of Problem 1 is

$$\mathbf{x}^* = (A_1^\top \Sigma_1^{-1} A_1)^{-1} A_1^\top \Sigma_1^{-1} \mathbf{b}_1,$$

with the covariance of the optimal solution given by

$$\text{cov}(\mathbf{x}^*) = (A_1^\top \Sigma_1^{-1} A_1)^{-1}.$$

The objective function f^* of Problem 1 at the optimal solution \mathbf{x}^* can be computed by $f^* = \|A_1 \mathbf{x}^* - \mathbf{b}_1\|_{\Sigma_1}^2$ whose expansion yields

$$f^* = -\mathbf{b}_1^\top \Sigma_1^{-1} A_1 (A_1^\top \Sigma_1^{-1} A_1)^{-1} A_1^\top \Sigma_1^{-1} \mathbf{b}_1 + \mathbf{b}_1^\top \Sigma_1^{-1} \mathbf{b}_1.$$

With the information obtained from Problem 1, the objective function of Problem 2 at its optimal solution \mathbf{x}^{**} can be exactly predicted, which gives the following proposition.

Proposition 1. (Predict Objective Function Change in the Linear Least Squares Problem): The objective function f^{**} of Problem 2 at the optimal solution \mathbf{x}^{**} which is $f^{**} = \|A \mathbf{x}^{**} - \mathbf{b}\|_{\Sigma}^2$, can be computed directly from the estimate \mathbf{x}^* of Problem 1 along with its covariance $\text{cov}(\mathbf{x}^*)$, without the need to solve Problem 2. Exactly, f^{**} can be computed as,

$$f^{**} = f^* + \mathbf{e}_2^\top \cdot \Omega^{-1} \cdot \mathbf{e}_2 \quad (1)$$

where

$$\mathbf{e}_2 = A_2 \mathbf{x}^* - \mathbf{b}_2 = \text{stack}\{H_i \mathbf{x}^* - \mathbf{h}_i - \tilde{\mathbf{z}}_i\}_{i=n_1+1}^{n_2} \quad (2)$$

$$\Omega = \underbrace{\begin{bmatrix} A_2 & -I \end{bmatrix}}_{J_2} \begin{bmatrix} \text{cov}(\mathbf{x}^*) \\ \Sigma_2 \end{bmatrix} \begin{bmatrix} A_2^\top \\ -I \end{bmatrix}. \quad (3)$$

In other words, $\mathbf{e}_2^\top \cdot \Omega^{-1} \cdot \mathbf{e}_2$ in (1) is a **metric for the objective function change**, defined as

$$\Delta f_{**}^* = f^{**} - f^*.$$

Proof. See Appendix I for the algebraic proof. \square

Remark 1. If we denote the measurement error of the i -th measurement as

$$\tau_i(\mathbf{x}, \mathbf{z}_i) = H_i \mathbf{x} - \mathbf{h}_i - \mathbf{z}_i,$$

then \mathbf{e}_2 is the stacked measurement error evaluated at \mathbf{x}^* , i.e., $\mathbf{e}_2 = \text{stack}\{\tau_i(\mathbf{x}^*, \tilde{\mathbf{z}}_i)\}_{i=n_1+1}^{n_2}$. In addition, J_2 in (3) can be written as

$$J_2 = \text{stack}\left\{ \left[\frac{\partial \tau_i(\mathbf{x}, \mathbf{z}_i)}{\partial \mathbf{x}^\top} \right]_{\mathbf{x}^*} \quad \left[\frac{\partial \tau_i(\mathbf{x}, \mathbf{z}_i)}{\partial \mathbf{z}_i^\top} \right]_{\tilde{\mathbf{z}}_i} \right\}_{i=n_1+1}^{n_2},$$

which implies Ω can be regarded as the covariance of the measurement error $\tau_i(\mathbf{x}, \mathbf{z}_i)$, which is obtained through the linear transformation of the Gaussian distribution.

Remark 2. It is not surprising that Δf_{**}^* takes a form of so-called Mahalanobis distance. In the multiple hypothesis tracking literature [18], Mahalanobis distance is used as a metric to detect feature correspondences, under the name of

Mahalanobis gating. However, the methods are mostly used in filtering approaches [19] rather than pose-graph optimization, and to the best of our knowledge, the relation with the objective function change in least squares optimization has not been studied yet in the literature.

IV. PREDICT OBJECTIVE FUNCTION CHANGE IN THE NONLINEAR LEAST SQUARES PROBLEM

In this section, we will extend the results of the last section to the nonlinear cases based on linearization.

A. Original Nonlinear Least Squares Problems

Assume the nonlinear least squares problem is obtained through the nonlinear measurement model with Gaussian noise added as

$$\mathbf{z}_i = \Gamma_i(\mathbf{x}) + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(\mathbf{0}, \tilde{\Sigma}_i).$$

Given noisy measurements $\tilde{\mathbf{z}}_i$, we can analogously define the primal and augmented problem.

Problem 3. (Primal Nonlinear Least Squares Optimization Problem): Given a collection of noisy measurements $\{\tilde{\mathbf{z}}_i, \tilde{\Sigma}_i\}_{i=1}^{n_1}$, the nonlinear least squares to estimate \mathbf{x} is

$$\mathbf{x}^* = \arg \min \sum_{i=1}^{n_1} \|\Gamma_i(\mathbf{x}) - \tilde{\mathbf{z}}_i\|_{\tilde{\Sigma}_i}^2.$$

Problem 4. (Augmented Nonlinear Least Squares Optimization Problem): The augmented nonlinear least squares with additional noisy measurements $\{\tilde{\mathbf{z}}_i, \tilde{\Sigma}_i\}_{i=n_1+1}^{n_2}$ is

$$\mathbf{x}^{**} = \arg \min \sum_{i=1}^{n_1} \|\Gamma_i(\mathbf{x}) - \tilde{\mathbf{z}}_i\|_{\tilde{\Sigma}_i}^2 + \sum_{i=n_1+1}^{n_2} \|\Gamma_i(\mathbf{x}) - \tilde{\mathbf{z}}_i\|_{\tilde{\Sigma}_i}^2.$$

B. Linearized Problems

To adopt the results in linear cases, we turn to look at the linearized versions of Problem 3 and Problem 4. At \mathbf{x}^* and \mathbf{x}^{**} , $\Gamma_i(\mathbf{x})$ can be linearized as

$$\Gamma_i(\mathbf{x}) \approx \begin{cases} \Gamma_i(\mathbf{x}^*) + \frac{\partial \Gamma_i(\mathbf{x})}{\partial \mathbf{x}^\top} \big|_{\mathbf{x}^*} (\mathbf{x} - \mathbf{x}^*) = \Pi_i \mathbf{x} - \pi_i \\ \Gamma_i(\mathbf{x}^{**}) + \frac{\partial \Gamma_i(\mathbf{x})}{\partial \mathbf{x}^\top} \big|_{\mathbf{x}^{**}} (\mathbf{x} - \mathbf{x}^{**}) = \Delta_i \mathbf{x} - \delta_i \end{cases}$$

where $\Pi_i = \frac{\partial \Gamma_i(\mathbf{x})}{\partial \mathbf{x}^\top} \big|_{\mathbf{x}^*}$, $\pi_i = \Pi_i \mathbf{x}^* - \Gamma_i(\mathbf{x}^*)$, $\Delta_i = \frac{\partial \Gamma_i(\mathbf{x})}{\partial \mathbf{x}^\top} \big|_{\mathbf{x}^{**}}$, $\delta_i = \Delta_i \mathbf{x}^{**} - \Gamma_i(\mathbf{x}^{**})$.

The linearized versions of Problem 3 and Problem 4 that we will discuss later is presented as below.

Problem 5. (Linearized Version of Problem 3 at \mathbf{x}^*): The linearized version of Problem 3 at \mathbf{x}^* is

$$\mathbf{x}_{l_*}^* = \arg \min \|\Lambda_1 \mathbf{x} - \lambda_1\|_{\Sigma_1}^2.$$

where $\Lambda_1 = \text{stack}\{\Pi_i\}_{i=1}^{n_1}$, $\lambda_1 = \text{stack}\{\pi_i + \tilde{\mathbf{z}}_i\}_{i=1}^{n_1}$, and $\Sigma_1 = \text{blkdiag}\{\tilde{\Sigma}_i\}_{i=1}^{n_1}$.

Problem 6. (Linearized Version of Problem 4 at \mathbf{x}^*): The linearized version of Problem 4 at \mathbf{x}^* is

$$\mathbf{x}_{l_*}^{**} = \arg \min \|\Lambda_1 \mathbf{x} - \lambda_1\|_{\Sigma_1}^2 + \|\Lambda_2 \mathbf{x} - \lambda_2\|_{\Sigma_2}^2.$$

where $\Lambda_2 = \text{stack}\{\Pi_i\}_{i=n_1+1}^{n_2}$, $\lambda_2 = \text{stack}\{\pi_i + \tilde{\mathbf{z}}_i\}_{i=n_1+1}^{n_2}$, and $\Sigma_2 = \text{blkdiag}\{\tilde{\Sigma}_i\}_{i=n_1+1}^{n_2}$.

Problem 7. (Linearized Version of Problem 4 at \mathbf{x}^{**}): The linearized version of Problem 4 at \mathbf{x}^{**} is

$$\mathbf{x}_{l_*}^{**} = \arg \min \|\mathbf{K} \mathbf{x} - \kappa\|_{\Sigma}^2.$$

where $\mathbf{K} = \text{stack}\{\Delta_i\}_{i=1}^{n_2}$, $\kappa = \text{stack}\{\delta_i + \tilde{\mathbf{z}}_i\}_{i=1}^{n_2}$, and $\Sigma = \text{blkdiag}\{\tilde{\Sigma}_i\}_{i=1}^{n_2}$.

C. Relationship between Nonlinear and Linearized Problems

It can be shown that Problem 3 and Problem 5 have the same solution and the same objective function, i.e., $\mathbf{x}^* = \mathbf{x}_{l_*}^*$, $f^* = f_{l_*}^*$. The proof is straightforward by the fact that

$$\Gamma_i(\mathbf{x}^*) - \tilde{\mathbf{z}}_i = \Pi_i \mathbf{x}^* - \pi_i - \tilde{\mathbf{z}}_i$$

then obviously $f^* = f_{l_*}^*$ holds because

$$\begin{aligned} \sum_{i=1}^{n_1} \|\Gamma_i(\mathbf{x}^*) - \tilde{\mathbf{z}}_i\|_{\tilde{\Sigma}_i}^2 &= \sum_{i=1}^{n_1} \|\Pi_i \mathbf{x}^* - \pi_i - \tilde{\mathbf{z}}_i\|_{\tilde{\Sigma}_i}^2 \\ &= \|\Lambda_1 \mathbf{x}^* - \lambda_1\|_{\Sigma_1}^2. \end{aligned}$$

The solution satisfies $\mathbf{x}^* = \mathbf{x}_{l_*}^*$ because Problem 3 and Problem 5 have the same gradient at \mathbf{x}^* , shown as below,

$$\begin{aligned} \text{grad}_3 &= 2 \sum_{i=1}^{n_1} \left[\frac{\partial \Gamma_i(\mathbf{x})}{\partial \mathbf{x}^\top} \big|_{\mathbf{x}^*} \right]^\top \cdot \tilde{\Sigma}_i^{-1} \cdot (\Gamma_i(\mathbf{x}^*) - \tilde{\mathbf{z}}_i) \\ &= 2 \sum_{i=1}^{n_1} \Pi_i^\top \cdot \tilde{\Sigma}_i^{-1} \cdot (\Pi_i \mathbf{x}^* - \pi_i - \tilde{\mathbf{z}}_i) \\ &= 2 \Lambda_1^\top \cdot \Sigma_1^{-1} \cdot (\Lambda_1 \mathbf{x}^* - \lambda_1) = \text{grad}_5. \end{aligned}$$

Analogously Problem 4 and Problem 7 have the same solution and objective function, i.e., $\mathbf{x}_{l_*}^{**} = \mathbf{x}^{**}$, $f_{l_*}^{**} = f^{**}$.

D. Objective Function Change between Linearized Problems

Then the objective function change from Problem 3 to Problem 4 equals to the objective function change from Problem 5 to Problem 7, namely

$$\Delta f_{l_*}^{**} = f^{**} - f^* = f_{l_*}^{**} - f_{l_*}^*.$$

However the linearization point \mathbf{x}^{**} cannot be obtained until we actually solve Problem 4. Instead, we consider Problem 6 as an intermediate problem, which is linearized at a given solution \mathbf{x}^* . The objective function change satisfies,

$$\Delta f_{l_*}^{**} = f_{l_*}^{**} - f_{l_*}^* = (f_{l_*}^{**} - f_{l_*}^*) + (f_{l_*}^{**} - f_{l_*}^{**}).$$

Obviously if condition

$$\mathcal{C}: \quad |f_{l_*}^{**} - f_{l_*}^*| \gg |f_{l_*}^{**} - f_{l_*}^{**}| \quad (4)$$

holds, we can safely approximate $\Delta f_{l_*}^{**} \simeq f_{l_*}^{**} - f_{l_*}^*$ which can be exactly computed from Proposition 1 in Section III.

E. Metric to Predict Objective Function Change

Proposition 2. Define the i -th measurement error as

$$\tau_i(\mathbf{x}, \mathbf{z}_i) = \Gamma_i(\mathbf{x}) - \mathbf{z}_i.$$

Denote the solution of Problem 3 as \mathbf{x}^* with covariance $\text{cov}(\mathbf{x}^*)$. If condition C holds, after including noisy measurements $\{\tilde{\mathbf{z}}_i, \tilde{\Sigma}_i\}_{i=n_1+1}^{n_2}$, the objective function change from Problem 3 to Problem 4 can be approximated by

$$\Delta f_{\star}^{**} \simeq f_{l_{\star}}^{**} - f_{l_{\star}}^* = \mathbf{e}_{\star}^{\top} \cdot \Xi^{-1} \cdot \mathbf{e}_{\star}$$

where

$$\mathbf{e}_{\star} = \text{stack}\{\tau_i(\mathbf{x}^*, \tilde{\mathbf{z}}_i)\}_{i=n_1+1}^{n_2}$$

$$\Xi = J_{\star} \cdot \text{blkdiag}\{\text{cov}(\mathbf{x}^*), \text{blkdiag}\{\tilde{\Sigma}_i\}_{i=n_1+1}^{n_2}\} \cdot J_{\star}^{\top}$$

with Jacobian $J_{\star} = \text{stack}\left\{\left[\frac{\partial \tau_i(\mathbf{x}, \mathbf{z}_i)}{\partial \mathbf{x}^{\top}}\right]_{\mathbf{x}^*} \quad -I\right\}_{i=n_1+1}^{n_2}$.

The proof is straightforward from Proposition 1, with the fact that $\tau_i(\mathbf{x}^*, \tilde{\mathbf{z}}_i) = \Gamma_i(\mathbf{x}^*) - \tilde{\mathbf{z}}_i = \Pi_i \mathbf{x}^* - \pi_i - \tilde{\mathbf{z}}_i$, and $\text{stack}\left\{\left[\frac{\partial \tau_i(\mathbf{x}, \mathbf{z}_i)}{\partial \mathbf{x}^{\top}}\right]_{\mathbf{x}^*}\right\}_{i=n_1+1}^{n_2} = \Lambda_2$.

V. PREDICT OBJECTIVE FUNCTION CHANGE IN POSE-GRAPH OPTIMIZATION

Given a relative pose measurement $\tilde{T}_{i,j}$, the measurement error is defined as

$$\tau_{i,j}(T_i, T_j) = \text{Log}(\tilde{T}_{i,j}^{-1} \cdot T_i^{-1} \cdot T_j).$$

In this section, we will seek to predict the objective function change using the linear approximation in the first iteration.

A. Linearization

The linearization of the SE(3) manifold is achieved by adding a perturbation in the tangent space with the exponential mapping, and then compute the differential with respect to the perturbation.

At the current step, the estimates of T_i and T_j are given as $T_{i_{\star}}, T_{j_{\star}}$. Applying perturbations ξ_i and ξ_j to the right-hand side of T_i and T_j respectively, after linearization, $\tau_{i,j}(T_i, T_j)$ takes the form

$$\tau_{i,j}(T_i, T_j) \approx J_i \cdot \xi_i + J_j \cdot \xi_j + \tau_{i,j}(T_{i_{\star}}, T_{j_{\star}})$$

where

$$J_i = \left. \frac{\partial \tau_{i,j}(T_{i_{\star}} \cdot \text{Exp}(\xi_i), T_{j_{\star}})}{\partial \xi_i^{\top}} \right|_{\xi_i=0}$$

$$J_j = \left. \frac{\partial \tau_{i,j}(T_{i_{\star}}, T_{j_{\star}} \cdot \text{Exp}(\xi_j))}{\partial \xi_j^{\top}} \right|_{\xi_j=0}.$$

The analytical Jacobian can be easily computed as

$$J_i = -J_l^{-1}(\eta) \cdot \text{Ad}(\tilde{T}_{i,j}^{-1}), \quad J_j = J_r^{-1}(\eta).$$

where $\eta = \tau_{i,j}(T_{i_{\star}}, T_{j_{\star}}) = \text{Log}(\tilde{T}_{i,j}^{-1} \cdot T_{i_{\star}}^{-1} \cdot T_{j_{\star}})$. See Appendix II for the detailed calculation.

B. Metric to Predict Objective Function Change

For PGO, the measurement can be classified into two categories: odometry and loop-closure. The odometry measurement induces new poses while having no impacts on previous poses, so the objective function change after including an odometry measurement is always zero. For a loop-closure measurement, the resulting objective function change can be predicted by the metric described in Section IV, since PGO is a special case of nonlinear least squares.

We consider the case of an online PGO, where the loop-closure measurements are obtained one by one. Given the current solution \mathbf{x}^* and its covariance $\text{cov}(\mathbf{x}^*)$, if a new loop-closure measurement $\tilde{T}_{i,j}$ from pose i to j arrives, the metric to predict the objective function change after including the relative pose measurement $\tilde{T}_{i,j}$ into the PGO, can be formulated as,

$$\Delta f_{\star}^{**} \simeq \tau_{i,j}(T_{i_{\star}}, T_{j_{\star}})^{\top} \cdot [A_{i,j} \cdot \text{cov}(\mathbf{x}^*) \cdot A_{i,j}^{\top} + \tilde{\Sigma}_{i,j}]^{-1} \cdot \tau_{i,j}(T_{i_{\star}}, T_{j_{\star}})$$

where the linearized measurement matrix at \mathbf{x}^* is

$$A_{i,j} = [\cdots \quad \mathbf{0} \quad -J_i \quad \mathbf{0} \quad \cdots \quad \mathbf{0} \quad -J_j \quad \mathbf{0} \quad \cdots].$$

Since $A_{i,j}$ contains only two non-zero blocks, simple matrix multiplication shows that $A_{i,j} \cdot \text{cov}(\mathbf{x}^*) \cdot A_{i,j}^{\top}$ is only related to the i -th and j -th block row and block column of $\text{cov}(\mathbf{x}^*)$. As a result, $A_{i,j} \cdot \text{cov}(\mathbf{x}^*) \cdot A_{i,j}^{\top}$ can be simplified as

$$A_{i,j} \cdot \text{cov}(\mathbf{x}^*) \cdot A_{i,j}^{\top} = [-J_i \quad -J_j] \cdot \mathcal{M}_{i,j} \cdot [-J_i \quad -J_j]^{\top}$$

where

$$\mathcal{M}_{i,j} = \begin{bmatrix} \Sigma_{i,i} & \Sigma_{i,j} \\ \Sigma_{j,i} & \Sigma_{j,j} \end{bmatrix}$$

is the marginal covariance in terms of pose i and j .

C. Increments and Solution Space

There is a subtle difference between the linearization technique in this section and that in the previous section. For PGO, we adopt Lie group tools and linearize the problem in the perturbation space (i.e., increment space). However, it can be easily shown that the linearized problem and the original nonlinear problem still take the same objective function, if the linearization occurs at the optimal solution of the nonlinear problem. Following the discussion in Section IV, then the objective function change between the primal and augmented problems can be represented by the objective function change between the linearized versions. Thus the discussion in Section IV still holds for the PGO problem.

D. Condition C and Metric to Quantify Prediction Error

Recall that if we want to use the objective function change in the linear system to predict that in the nonlinear system, the condition C (defined in (4)) must hold. Even though it is not yet mathematically proved, we will show empirically that this is usually the case particularly for an online incremental SLAM problem.

To quantify the correctness of the prediction, we propose to use the relative error γ between the predicted objective function change $\Delta f_{*p}^{**} = f_{*}^{**} - f_{l_*}^{**}$ and the real objective function change $\Delta f_{*r}^{**} = f_{*}^{**} - f_{l_*}^{**}$, defined as

$$\beta = \frac{\Delta f_{*r}^{**} - \Delta f_{*p}^{**}}{\Delta f_{*p}^{**}} = \frac{f_{*}^{**} - f_{l_*}^{**}}{f_{*}^{**} - f_{l_*}^{**}}. \quad (5)$$

Please note that the correctness of the condition \mathcal{C} can also be quantified by the metric β . If β is close to zero, it means that the condition \mathcal{C} holds and the prediction is accurate.

VI. RESULTS

A. Experimental Setting

1) *Implementation:* We implement our approach on the SLAM++ graph optimization library [3][20], which provides state-of-the-art strategy to compute the marginal covariance with real-time performance, particularly in an incremental setting. We opt to use a batch solver (Lambda solver) in the SLAM++ setting and set the error tolerance for convergence to $1e-5$ which attains a good balance between the accuracy and timing. An incremental solver (like FastL solver in SLAM++) with larger error tolerance can be significantly faster, however with some compromise on the accuracy. All the experiments are carried out on an Intel i5-5300U CPU. This is a quad core CPU with 2.30GHz for each, and there is no independent GPU.

2) *Datasets:* The evaluation is carried out on two real datasets, INTEL [21] and MITb [22], and four simulated datasets, Manhattan [23], City10k [24], Sphere2500 [24] and Torus [24]. Sphere2500 and Torus are 3D datasets while the rest are 2D datasets. The raw data are available at [25], and the details for the data processing are given in [21] and [22] respectively. The brief information of these datasets is reported in Table I.

B. Prediction of the Objective Function Change

For each dataset, we reorder relative pose measurements in an incremental online sequence, to mimic the measurement data sequence obtained in the real online SLAM application. The objective function change after including an odometry measurement is trivially zero, so we only evaluate the objective function change after including a loop-closure measurement.

TABLE I
BRIEF INFORMATION OF THE DATASETS USED FOR EVALUATION.

Dataset	Poses	Loop-Closures	Noise-Level	Timing
INTEL	1228	256	High	5.6s
MITb	808	20	High	1.5s
Manhattan	3500	2099	Medium	150.0s
City10k	10000	10688	Medium	6754.9s
Sphere2500	2500	2450	Medium	848.3s
Torus	10000	12281	Medium	15582.8s

1) *Accuracy:* We feed the data one by one into the SLAM++ solver, and record the predicted and real objective function change after including a loop-closure measurement, along with the relative error (defined in (5)) in Fig. 1. It can be seen from Fig. 1 that most of the time, the relative error is pretty close to zero. The largest relative error of the prediction lies around 5% ~ 10%, which is accurate enough to be used for applications like outlier detection.

2) *Timing:* The total timing for each dataset is reported in the last column of Table I. The computation is roughly in real-time for each measurement data. However, as already mentioned, we use a batch solver to benchmark the accuracy of the prediction. For real applications, the timing can be improved by using an incremental solver. According to our tests on SLAM++, FastL solver can be at least 2-3 times faster if using a larger error tolerance for example 0.001 (instead of $1e-5$).

C. An Application to Outlier Detection

At last, we show an application of the predicted objective function change by detecting outliers in the INTEL dataset. The original INTEL dataset is outlier-free, and we take the optimal solution as the ground truth. The outliers are generated by randomly selecting two poses, computing their relative pose measurement with the ground truth, and then assigning the computed relative pose measurement to two other poses. The absolute trajectory error (ATE) is computed as the Euclidean norm of the Cartesian difference between the solution and the ground truth.

1) *Illustration of Outlier Detection:* At first, we add 100 outliers to the INTEL dataset, and organize the data sequentially. Then we incrementally solve the dataset containing outliers, using the predicted objective function change as a benchmark to discern outliers. The predicted objective function change for the inlier and outlier loop-closure measurement is presented in Fig. 2a. It can be seen that the predicted objective function change for an outlier is usually orders of magnitude larger than that for an inlier. Thus the inlier and outlier measurement can be easily classified by a simple χ^2 difference test [26][27].

The computational time for each measurement is reported in Fig. 2b, where the timing for outliers is negligible compared with that for inliers. The advantage benefits from the fact that, for an outlier measurement, we only need to compute the metric to predict the objective function change, without any additional expense. This is substantially different from the exiting χ^2 test based methods [14][15], where a problem containing outliers is solved which degenerates the efficiency of the algorithm.

2) *Robustness and Timing on Monte-Carlo Simulation:* To evaluate the robustness of our method against outliers, we use a 20 run Monte-Carlo simulation, with respect to different quantity of outliers. We compare our method with Cauchy M-estimator (with tuning parameter set to 1) in an online incremental scheme. We disable the covariance computation for Cauchy (which is not required for M-estimators and saves a large amount of time), and set the error tolerance for

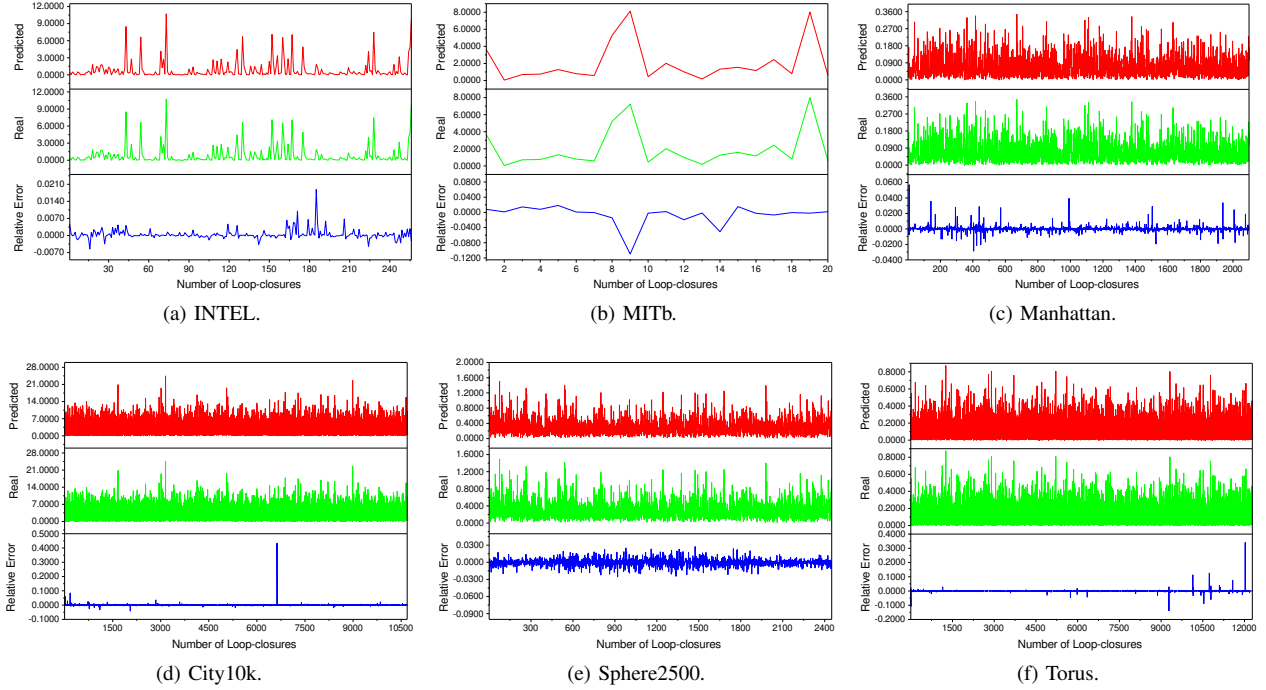


Fig. 1. The predicted and real objective function change after including each loop-closure measurement. The peak in City10k (the point with relative error 43%) is caused by the numerical precision, where the predicted and real objective function change are 0.001462 and 0.002097 respectively. However, by using smaller error tolerance for convergence like $1e-6$, the values can be reduced to 0.001462 and 0.001410, with relative error around 3.6%.

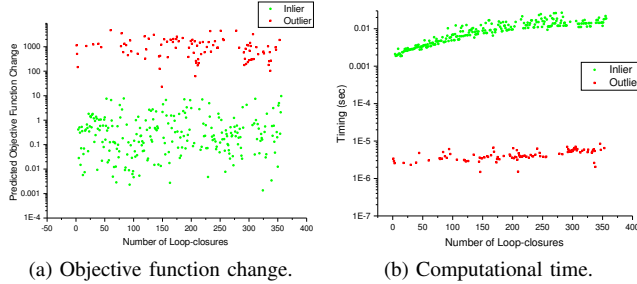


Fig. 2. An example of the predicted objective function change for the inlier and outlier loop-closure measurement and the overall computational time for each measurement on the INTEL dataset.

convergence to $1e-5$ for a fair comparison. The resulting ATE is plotted in Fig. 3. It can be seen that our method (shown in Fig. 3a) diverges less and produces more accurate results than Cauchy (shown in Fig. 3b). The overall computational time of our method and Cauchy for each Monte-Carlo trial is sketched in Fig. 4. In general, our method is much faster than Cauchy M-estimator in an online incremental scheme, and the method scales well with the growing number of outliers, contrasted to the radical growth in timing of Cauchy.

VII. DISCUSSION

A. Why Information-theoretic Approach is Insufficient

Consider the case of adding a new edge to pose-graph \mathcal{G} , and denote the resulting pose-graph as \mathcal{G}' . The information gain (described by the mutual information) [6][10][28] after

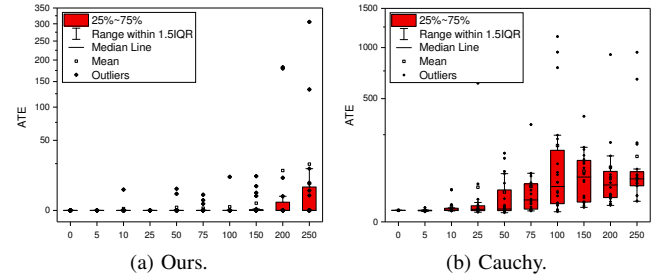


Fig. 3. The robustness of our method and Cauchy M-estimator on the INTEL dataset, shown as the absolute trajectory error (ATE) with respect to different number of outliers on a 20 run Monte-Carlo simulation.

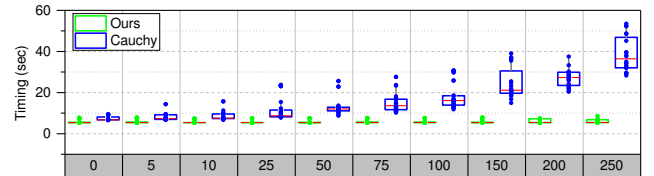


Fig. 4. The overall computational time of our method and Cauchy M-estimators for each Monte-Carlo run on the INTEL dataset.

adding a new link/edge is

$$\mathcal{MI} = \frac{1}{2} \ln \frac{\det(\Sigma_{\mathcal{G}})}{\det(\Sigma_{\mathcal{G}'})} = \frac{1}{2} \ln \frac{\det(\mathcal{I}_{\mathcal{G}'})}{\det(\mathcal{I}_{\mathcal{G}})} \quad (6)$$

where \mathcal{I} stands for the Fisher information matrix which is the inverse of the covariance matrix [29].

It has been shown in [29] that for SLAM instances where the rotation is known, (named “compass SLAM” in

[29], which is actually a linear measurement model), the determinant of the covariance (or Fisher information) is closely related to the number of spanning trees (denoted by t) in the graph. In this case, (6) can be transformed to

$$\mathcal{MI} = \frac{1}{2} \ln \frac{\det(\mathcal{I}_{\mathcal{G}'})}{\det(\mathcal{I}_{\mathcal{G}})} \propto \ln \frac{t(\mathcal{G}')}{t(\mathcal{G})}. \quad (7)$$

In (7), the information gain \mathcal{MI} is only related to the topology of graph, thus cannot discern an inlier and outlier link, because the resulting graph \mathcal{G}' would have the same number of spanning trees $t(\mathcal{G}')$ in both scenarios.

For nonlinear cases, the Fisher information matrix is related to the linearization point, denoted as $\mathcal{I}(\mathbf{x})$. The spanning tree index can be used to derive a lower bound for $\det(\mathcal{I}(\mathbf{x}))$ [30]. The lower bound is very close to $\det(\mathcal{I}(\mathbf{x}))$ in experiments [30], which reveals a strong relationship to graph topology (i.e., number of spanning trees) as well. In real applications, the linearization point is usually chosen at the current estimate of the problem [10][9]. Applying matrix determinant lemma [31], \mathcal{MI} can be computed from the measurement covariance of the link and the marginal covariance of the estimate to achieve efficiency [10]; however the measurement value itself is not considered.

B. Relationship and Difference between Objective Function Change and Information Gain

Consider the linear case in Proposition 1, the expansion of Ω shows

$$\Omega = A_2 \cdot \text{cov}(\mathbf{x}^*) \cdot A_2^T + \Sigma_2. \quad (8)$$

Applying the Woodbury matrix identity to (8), Ω^{-1} can be written in the information form

$$\Omega^{-1} = \Sigma_2^{-1} - \Sigma_2^{-1} A_2 \cdot [\text{cov}^{-1}(\mathbf{x}^*) + A_2^T \Sigma_2^{-1} A_2]^{-1} \cdot A_2^T \Sigma_2^{-1}.$$

where the Fisher information matrix without the new measurement is $\mathcal{I} = \text{cov}^{-1}(\mathbf{x}^*)$, while the Fisher information matrix after including the new measurement is $\mathcal{I}' = \text{cov}^{-1}(\mathbf{x}^*) + A_2^T \Sigma_2^{-1} A_2$. This shows the main difference between information-theoretic metric like information gain $\mathcal{MI} = \frac{1}{2} \ln \frac{\det(\mathcal{I}')}{\det(\mathcal{I})}$ and the objective function change $f^{**} = \mathbf{e}_2^T \cdot \Omega^{-1} \cdot \mathbf{e}_2$. It is clear that the objective function change considers the measurement error \mathbf{e}_2 , while the information gain does not.

C. Drawbacks of Objective Function Metric

There are drawbacks of the objective function metric as well. In our previous contribution [27], we use the objective function change to quantify the change of solution in a constrained optimization framework. However, in cases where the optimization problem is strongly nonlinear, which has long shallow valleys (see [12] for a discussion of this phenomenon), the solution can change dramatically while the objective function does not change much. In this adverse case, the objective function (or its changes) might not be a good metric to benchmark the impact of new measurements (or detect outliers).

VIII. CONCLUSION

To summarize, we propose a metric to predict the change of the optimal value of the objective function after including new measurements in least squares optimization. We take pose-graph optimization as an instance and experimental results validate the accuracy of the predicted objective function change. The predicted objective function change can be used to detect outliers. Although the method requires the computation of the marginal covariance, our experiments on SLAM++ show that our method is much faster than M-estimators in an online incremental scheme, meanwhile yielding enhanced robustness.

APPENDIX I

THE PROOF OF OBJECTIVE FUNCTION CHANGE IN LLS

Let $N = A_1^T \Sigma_1^{-1} A_1$, $M = N + A_2^T \Sigma_2^{-1} A_2$.

The optimal solution of the augmented system can be computed in a closed form, and some trivial math shows that the objective function f^{**} is

$$\begin{aligned} f^{**} = & -\mathbf{b}_1^T \Sigma_1^{-1} A_1 M^{-1} A_1^T \Sigma_1^{-1} \mathbf{b}_1 + \mathbf{b}_1^T \Sigma_1^{-1} \mathbf{b}_1 \\ & -\mathbf{b}_2^T \Sigma_2^{-1} A_2 M^{-1} A_2^T \Sigma_2^{-1} \mathbf{b}_2 + \mathbf{b}_2^T \Sigma_2^{-1} \mathbf{b}_2 \\ & -2\mathbf{b}_1^T \Sigma_1^{-1} A_1 M^{-1} A_2^T \Sigma_2^{-1} \mathbf{b}_2. \end{aligned}$$

The difference $f^{**} - f^*$ is

$$\begin{aligned} f^{**} - f^* = & \mathbf{b}_1^T \Sigma_1^{-1} A_1 \cdot [(A_1^T \Sigma_1^{-1} A_1)^{-1} - M^{-1}] \cdot A_1^T \Sigma_1^{-1} \mathbf{b}_1 \\ & + \mathbf{b}_2^T \cdot [\Sigma_2^{-1} - \Sigma_2^{-1} A_2 M^{-1} A_2^T \Sigma_2^{-1}] \cdot \mathbf{b}_2 \\ & - 2\mathbf{b}_1^T \Sigma_1^{-1} A_1 M^{-1} A_2^T \Sigma_2^{-1} \mathbf{b}_2. \end{aligned}$$

Substituting \mathbf{x}^* into $\mathbf{e}_2 = A_2 \mathbf{x}^* - \mathbf{b}_2$, we have

$$\mathbf{e}_2 = A_2 N^{-1} A_1^T \Sigma_1^{-1} \mathbf{b}_1 - \mathbf{b}_2$$

then

$$\begin{aligned} \mathbf{e}_2^T \cdot \Omega^{-1} \cdot \mathbf{e}_2 = & \mathbf{b}_1^T \Sigma_1^{-1} A_1 N^{-1} A_2^T \Omega^{-1} A_2 N^{-1} A_1^T \Sigma_1^{-1} \mathbf{b}_1 \\ & + \mathbf{b}_2^T \Omega^{-1} \mathbf{b}_2 - 2\mathbf{b}_1^T \Sigma_1^{-1} A_1 N^{-1} A_2^T \Omega^{-1} \mathbf{b}_2. \end{aligned}$$

To show the equality of $f^{**} - f^* = \mathbf{e}_2^T \cdot \Omega^{-1} \cdot \mathbf{e}_2$, we just need to show below three equalities.

$$\Omega^{-1} = \Sigma_2^{-1} - \Sigma_2^{-1} A_2 M^{-1} A_2^T \Sigma_2^{-1} \quad (9)$$

$$N^{-1} A_2^T \Omega^{-1} A_2 N^{-1} = N^{-1} - M^{-1} \quad (10)$$

$$M^{-1} A_2^T \Sigma_2^{-1} = N^{-1} A_2^T \Omega^{-1} \quad (11)$$

Equality (9) can be examined by the fact that

$$\Omega = \Sigma_2 + A_2 N^{-1} A_2^T$$

and applying the Woodbury matrix identity [32] to Ω^{-1} . Equality (10) can be obtained by applying Woodbury matrix identity to M^{-1} . Equality (11) holds because the fact that

$$N^{-1} M = I + N^{-1} A_2^T \Sigma_2^{-1} A_2 \quad (12)$$

then right-multiply (12) by $M^{-1} A_2^T \Sigma_2^{-1}$ to reach

$$\begin{aligned} M^{-1} A_2^T \Sigma_2^{-1} = & N^{-1} A_2^T \Sigma_2^{-1} - N^{-1} A_2^T \Sigma_2^{-1} A_2 M^{-1} A_2^T \Sigma_2^{-1} \\ = & N^{-1} A_2^T (\Sigma_2^{-1} - \Sigma_2^{-1} A_2 M^{-1} A_2^T \Sigma_2^{-1}). \end{aligned}$$

The proof is immediate by considering (9).

APPENDIX II

CALCULATION OF THE ANALYTICAL JACOBIAN

The expansion of $J_i = \frac{\partial \tau_{i,j}(T_{i*} \cdot \text{Exp}(\xi_i), T_{j*})}{\partial \xi_i^T} \Big|_{\xi_i=0}$ yields,

$$J_i = \frac{\partial \{\text{Log}(\tilde{T}_{i,j}^{-1} \cdot \text{Exp}(-\xi_i) \cdot T_{i*}^{-1} \cdot T_{j*})\}}{\partial \xi_i^T} \Big|_{\xi_i=0}.$$

Then by the property of adjoint matrix, the exponential term of ξ_i can be exchanged to the left-most,

$$J_i = \frac{\partial \{\text{Log}(\text{Exp}(-\text{Ad}(\tilde{T}_{i,j}^{-1}) \cdot \xi_i) \cdot \tilde{T}_{i,j}^{-1} \cdot T_{i*}^{-1} \cdot T_{j*})\}}{\partial \xi_i^T} \Big|_{\xi_i=0}.$$

Denote $\eta = \tau_{i,j}(T_{i*}, T_{j*}) = \text{Log}(\tilde{T}_{i,j}^{-1} \cdot T_{i*}^{-1} \cdot T_{j*})$. The two matrix exponentials can be concatenated with approximated BCH formula, if $\xi \rightarrow 0$,

$$J_i \approx \frac{\partial \{\text{Log}(\text{Exp}(J_l^{-1}(\eta) \cdot (-\text{Ad}(\tilde{T}_{i,j}^{-1}) \cdot \xi_i) + \eta))\}}{\partial \xi_i^T} \Big|_{\xi_i=0}.$$

The outside logarithm mapping would cancel the exponential mapping inside. As a result, J_i takes the form,

$$\begin{aligned} J_i &= \frac{\partial \{-J_l^{-1}(\eta) \cdot \text{Ad}(\tilde{T}_{i,j}^{-1}) \cdot \xi_i + \eta\}}{\partial \xi_i^T} \Big|_{\xi_i=0} \\ &= -J_l^{-1}(\eta) \cdot \text{Ad}(\tilde{T}_{i,j}^{-1}). \end{aligned}$$

The calculation of matrix J_j is analogous. By expanding $\tau_{i,j}(T_{i*}, T_{j*} \cdot \text{Exp}(\xi_j))$, we reach two matrix exponentials which could be concatenated by the approximated BCH, as

$$\begin{aligned} J_j &= \frac{\partial \{\text{Log}(\tilde{T}_{i,j}^{-1} \cdot T_{i*}^{-1} \cdot T_{j*} \cdot \text{Exp}(\xi_j))\}}{\partial \xi_j^T} \Big|_{\xi_j=0} \\ &\approx \frac{\partial \{\text{Log}(\text{Exp}(\eta + J_r^{-1}(\eta) \cdot \xi_j))\}}{\partial \xi_j^T} \Big|_{\xi_j=0}. \end{aligned}$$

Then the logarithm and exponential would cancel, resulting,

$$J_j = \frac{\partial \{\eta + J_r^{-1}(\eta) \cdot \xi_j\}}{\partial \xi_j^T} \Big|_{\xi_j=0} = J_r^{-1}(\eta).$$

REFERENCES

- [1] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, and J. J. Leonard, "Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age," *IEEE Transactions on Robotics*, vol. 32, no. 6, pp. 1309–1332, 2016.
- [2] F. Lu and E. Milios, "Globally consistent range scan alignment for environment mapping," *Autonomous robots*, vol. 4, no. 4, pp. 333–349, 1997.
- [3] V. Ila, L. Polok, M. Solony, and P. Svoboda, "SLAM++ 1-A highly efficient and temporally scalable incremental SLAM framework," *The International Journal of Robotics Research*, vol. 36, no. 2, pp. 210–230, 2017.
- [4] D. M. Rosen, L. Carlone, A. S. Bandeira, and J. J. Leonard, "A certifiably correct algorithm for synchronization over the special Euclidean group," in *the Algorithmic Foundations of Robotics, The 12th International Workshop on*, 2016.
- [5] D. J. MacKay, *Information theory, inference and learning algorithms*. Cambridge university press, 2003.
- [6] A. J. Davison, "Active search for real-time vision," in *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, vol. 1. IEEE, 2005, pp. 66–73.
- [7] T. A. Vidal-Calleja, A. Sanfeliu, and J. Andrade-Cetto, "Action selection for single-camera slam," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 40, no. 6, pp. 1567–1581, 2010.
- [8] H. Kretschmar and C. Stachniss, "Information-theoretic compression of pose graphs for laser-based slam," *The International Journal of Robotics Research*, vol. 31, no. 11, pp. 1219–1230, 2012.
- [9] N. Carlevaris-Bianco, M. Kaess, and R. M. Eustice, "Generic node removal for factor-graph slam," *IEEE Transactions on Robotics*, vol. 30, no. 6, pp. 1371–1385, 2014.
- [10] V. Ila, J. M. Porta, and J. Andrade-Cetto, "Information-based compact pose slam," *IEEE Transactions on Robotics*, vol. 26, no. 1, pp. 78–93, 2010.
- [11] H. Johannsson, M. Kaess, M. Fallon, and J. J. Leonard, "Temporally scalable visual slam using a reduced pose graph," in *Robotics and Automation (ICRA), 2013 IEEE International Conference on*. IEEE, 2013, pp. 54–61.
- [12] E. Olson and M. Kaess, "Evaluating the performance of map optimization algorithms," in *RSS Workshop on Good Experimental Methodology in Robotics*, vol. 15, 2009.
- [13] G. Hu, S. Huang, and G. Dissanayake, "Evaluation of pose only slam," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*. IEEE, 2010, pp. 3732–3737.
- [14] Y. Latif, C. Cadena, and J. Neira, "Robust loop closing over time for pose graph SLAM," *The International Journal of Robotics Research*, vol. 32, no. 14, pp. 1611–1626, 2013.
- [15] M. C. Graham, J. P. How, and D. E. Gustafson, "Robust incremental slam with consistency-checking," in *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*. IEEE, 2015, pp. 117–124.
- [16] G. S. Chirikjian, *Stochastic Models, Information Theory, and Lie Groups, Volume 2: Analytic Methods and Modern Applications*. Springer Science & Business Media, 2011, vol. 2.
- [17] T. D. Barfoot, *State Estimation for Robotics*. Cambridge University Press, 2017.
- [18] C. Kim, F. Li, A. Ciptadi, and J. M. Rehg, "Multiple hypothesis tracking revisited," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 4696–4704.
- [19] I. J. Cox and S. L. Hingorani, "An efficient implementation of Reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 18, no. 2, pp. 138–150, 1996.
- [20] [Online]. Available: <https://sourceforge.net/projects/slam-plus-plus/>
- [21] L. Carlone, R. Aragues, J. A. Castellanos, and B. Bona, "A linear approximation for graph-based simultaneous localization and mapping," *Robotics: Science and Systems VII*, pp. 41–48, 2012.
- [22] —, "A first-order solution to simultaneous localization and mapping with graphical models," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1764–1771.
- [23] E. Olson, J. Leonard, and S. Teller, "Fast iterative alignment of pose graphs with poor initial estimates," in *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*. IEEE, 2006, pp. 2262–2269.
- [24] M. Kaess, A. Ranganathan, and F. Dellaert, "iSAM: Incremental smoothing and mapping," *IEEE Transactions on Robotics*, vol. 24, no. 6, pp. 1365–1378, 2008.
- [25] [Online]. Available: <http://www.ipb.uni-bonn.de/datasets/>
- [26] K. Schermelleh-Engel, H. Moosbrugger, and H. Müller, "Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures," *Methods of psychological research online*, vol. 8, no. 2, pp. 23–74, 2003.
- [27] F. Bai, T. Vidal-Calleja, and S. Huang, "Robust incremental slam under constrained optimization formulation," *IEEE Robotics and Automation Letters*, vol. 3, no. 2, pp. 1207–1214, 2018.
- [28] M. G. Dissanayake, P. Newman, S. Clark, H. F. Durrant-Whyte, and M. Csorba, "A solution to the simultaneous localization and map building (slam) problem," *IEEE Transactions on robotics and automation*, vol. 17, no. 3, pp. 229–241, 2001.
- [29] K. Khosoussi, S. Huang, and G. Dissanayake, "Novel insights into the impact of graph structure on slam," in *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*. IEEE, 2014, pp. 2707–2714.
- [30] —, "Tree-connectivity: Evaluating the graphical structure of slam," in *Robotics and Automation (ICRA), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1316–1322.
- [31] D. A. Harville, *Matrix algebra from a statistician's perspective*. Springer, 1997, vol. 1.
- [32] W. W. Hager, "Updating the inverse of a matrix," *SIAM review*, vol. 31, no. 2, pp. 221–239, 1989.