

Supplementary Material for Night-Voyager: Consistent and Efficient Nocturnal Vision-Aided State Estimation in Object Maps

This document provides the supplementary material for the manuscript: “Night-Voyager: Consistent and Efficient Nocturnal Vision-Aided State Estimation in Object Maps”.

APPENDIX A DERIVATION OF STATE PROPAGATION FUNCTION

In this section, we derive the IMU propagation function for MSC-InEKF-FDRC. For simplicity, we only investigate the states \mathbf{X}_{I_i} and \mathbf{B}_i since other states are not correlated with the covariance propagation. We first rewrite the dynamic propagation model in [Eq. \(7\) of Section IV-A. State Propagation](#) in the manuscript as:

$$\mathbf{X}_{I_{i+1}} = \Gamma_i \Psi(\mathbf{X}_{I_i}) \Upsilon_i \quad (1)$$

where Γ_i , $\Psi(\mathbf{X}_{I_i})$, and Υ_i are formulated as:

$$\begin{aligned} \Gamma_i &= \begin{bmatrix} \mathbf{I}_{3 \times 3} & \frac{1}{2}\mathbf{g}\delta t^2 & \delta t \\ \mathbf{0}_{1 \times 3} & 1 & 0 \\ \mathbf{0}_{1 \times 3} & 0 & 1 \end{bmatrix}, \quad \Upsilon_i = \begin{bmatrix} \Delta\mathbf{R}_i & \Delta\mathbf{p}_i & \Delta\mathbf{v}_i \\ \mathbf{0}_{1 \times 3} & 1 & 0 \\ \mathbf{0}_{1 \times 3} & 0 & 1 \end{bmatrix} \\ \Psi(\mathbf{X}_{I_i}) &= \begin{bmatrix} {}^L\mathbf{R}_{I_i} & {}^L\mathbf{p}_{I_i} + {}^L\mathbf{v}_{I_i}\delta t & {}^L\mathbf{v}_{I_i} \\ \mathbf{0}_{1 \times 3} & 1 & 0 \\ \mathbf{0}_{1 \times 3} & 0 & 1 \end{bmatrix}. \end{aligned} \quad (2)$$

For $\Psi(\mathbf{X}_{I_i})$, it has the following properties:

$$\Psi(\mathbf{X}_1 \mathbf{X}_2) = \Psi(\mathbf{X}_1) \Psi(\mathbf{X}_2) \quad (3)$$

$$\Psi(\text{Exp}(\xi)) \approx \text{Exp}(\mathbf{F}\xi), \quad \mathbf{F} = \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \delta t \mathbf{I}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{bmatrix}. \quad (4)$$

Using [Eq. \(3\) of Section III-A. Theoretical Background](#) in the manuscript, (1) can be formulated as:

$$\begin{aligned} \text{Exp}(\xi_{I_{i+1}}) \hat{\mathbf{X}}_{I_{i+1}} &= \Gamma_i \Psi(\text{Exp}(\xi_i) \hat{\mathbf{X}}_{I_i}) \Upsilon_i \\ &= \Gamma_i \Psi(\text{Exp}(\xi_i)) \Psi(\hat{\mathbf{X}}_{I_i}) \Upsilon_i \approx \Gamma_i \text{Exp}(\mathbf{F}\xi_i) \Psi(\hat{\mathbf{X}}_{I_i}) \Upsilon_i \\ &= \text{Exp}(\text{Ad}(\Gamma_i)\mathbf{F}\xi_i) \Gamma_i \Psi(\hat{\mathbf{X}}_{I_i}) \Upsilon_i \\ &= \text{Exp}(\text{Ad}(\Gamma_i)\mathbf{F}\xi_i) \hat{\mathbf{X}}_{I_{i+1}}. \end{aligned} \quad (5)$$

Therefore, we can derive the formulation of Φ_{II} as:

$$\Phi_{II} = \text{Ad}(\Gamma_i)\mathbf{F} = \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \frac{1}{2}(\mathbf{g}) \times \delta t^2 & \mathbf{I}_{3 \times 3} & \mathbf{I}_{3 \times 3} \delta t \\ (\mathbf{g}) \times \delta t & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{bmatrix}. \quad (6)$$

To derive the relationship between $\xi_{I_{i+1}}$ and ζ_{I_i} , we first rewrite Υ_i as:

$$\Upsilon_i = \hat{\Upsilon}_i \tilde{\Upsilon}_i = \begin{bmatrix} \Delta\hat{\mathbf{R}}_i & \Delta\hat{\mathbf{p}}_i & \Delta\hat{\mathbf{v}}_i \\ \mathbf{0}_{1 \times 3} & 1 & 0 \\ \mathbf{0}_{1 \times 3} & 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta\tilde{\mathbf{R}}_i & \Delta\tilde{\mathbf{R}}_i^\top & \Delta\tilde{\mathbf{p}}_i & \Delta\hat{\mathbf{R}}_i^\top & \Delta\tilde{\mathbf{v}}_i \\ \mathbf{0}_{1 \times 3} & 1 & 0 \\ \mathbf{0}_{1 \times 3} & 0 & 1 \end{bmatrix}$$

$$= \hat{\Upsilon}_i \text{Exp} \left(\begin{bmatrix} \delta\Delta\theta_i \\ \mathbf{J}_l^{-1}(\delta\Delta\theta_i) \Delta\hat{\mathbf{R}}_i^\top \Delta\tilde{\mathbf{p}}_i \\ \mathbf{J}_l^{-1}(\delta\Delta\theta_i) \Delta\hat{\mathbf{R}}_i^\top \Delta\tilde{\mathbf{v}}_i \end{bmatrix} \right) \approx \hat{\Upsilon}_i \text{Exp} \left(\begin{bmatrix} \delta\Delta\theta_i \\ \Delta\hat{\mathbf{R}}_i^\top \Delta\tilde{\mathbf{p}}_i \\ \Delta\hat{\mathbf{R}}_i^\top \Delta\tilde{\mathbf{v}}_i \end{bmatrix} \right) \quad (7)$$

where $\Delta\tilde{\mathbf{R}}_i$, $\Delta\tilde{\mathbf{p}}_i$ and $\Delta\tilde{\mathbf{v}}_i$ are respectively formulated as:

$$\begin{aligned} \Delta\tilde{\mathbf{R}}_i &= \text{Exp}(-\mathbf{J}_r(\tilde{\omega}_i \delta t)(\delta\mathbf{b}_{g_i} + \mathbf{n}_{dg})\delta t) = \text{Exp}(\delta\Delta\theta_i) \\ \Delta\tilde{\mathbf{p}}_i &= \Xi_4(\delta\mathbf{b}_{g_i} + \mathbf{n}_{dg}) - \Xi_2(\delta\mathbf{b}_{a_i} + \mathbf{n}_{da}) \\ \Delta\tilde{\mathbf{v}}_i &= \Xi_3(\delta\mathbf{b}_{g_i} + \mathbf{n}_{dg}) - \Xi_1(\delta\mathbf{b}_{a_i} + \mathbf{n}_{da}) \end{aligned} \quad (8)$$

where $\mathbf{n}_{dg} \sim \mathcal{N}(\mathbf{0}_{3 \times 1}, \frac{\sigma_g^2}{\delta t} \mathbf{I}_{3 \times 3})$ and $\mathbf{n}_{da} \sim \mathcal{N}(\mathbf{0}_{3 \times 1}, \frac{\sigma_a^2}{\delta t} \mathbf{I}_{3 \times 3})$ are the discretized noises. The measurements $\tilde{\omega}_i$ and $\tilde{\mathbf{a}}_i$ are assumed to be constant during propagation, thus we have $\Delta\hat{\theta}_i = \tilde{\omega}_i \delta t = \text{Log}(\Delta\hat{\mathbf{R}}_i)$. Detailed derivation can be referred to [1], [2]. By substituting (7) into (1), we can derive:

$$\begin{aligned} \text{Exp}(\xi_{I_{i+1}}) \hat{\mathbf{X}}_{I_{i+1}} &= \Gamma_i \Psi(\text{Exp}(\mathbf{X}_i) \hat{\Upsilon}_i \tilde{\Upsilon}_i) \\ \Rightarrow \text{Exp}(\xi_{I_{i+1}}) &= \hat{\mathbf{X}}_{I_{i+1}} \tilde{\Upsilon}_i \hat{\mathbf{X}}_{I_{i+1}}^{-1} = \text{Exp}(\text{Ad}_{\hat{\mathbf{X}}_{I_{i+1}}} \begin{bmatrix} \delta\Delta\theta_i \\ \Delta\hat{\mathbf{R}}_i^\top \Delta\tilde{\mathbf{p}}_i \\ \Delta\hat{\mathbf{R}}_i^\top \Delta\tilde{\mathbf{v}}_i \end{bmatrix}) \\ &= \text{Exp}(\text{Ad}_{\hat{\mathbf{X}}_{I_{i+1}}} \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \Delta\hat{\mathbf{R}}_i^\top & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \Delta\hat{\mathbf{R}}_i^\top \end{bmatrix} \begin{bmatrix} \delta\Delta\theta_i \\ \Delta\tilde{\mathbf{p}}_i \\ \Delta\tilde{\mathbf{v}}_i \end{bmatrix}) \quad (9) \\ &= \text{Exp}(\text{Ad}_{\hat{\mathbf{X}}_{I_{i+1}}} \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \Delta\hat{\mathbf{R}}_i^\top & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \Delta\hat{\mathbf{R}}_i^\top \end{bmatrix} \begin{bmatrix} -\mathbf{J}_r(\Delta\hat{\theta}_i) \delta t \mathbf{0}_{3 \times 3} \\ \Xi_4 & -\Xi_2 \\ \Xi_3 & -\Xi_1 \end{bmatrix} \begin{bmatrix} \delta\mathbf{b}_{g_i} + \mathbf{n}_{dg} \\ \delta\mathbf{b}_{a_i} + \mathbf{n}_{da} \end{bmatrix}). \end{aligned}$$

Then we can derive Φ_{IB} and \mathbf{G}_{II} as:

$$\begin{aligned} \Phi_{IB} = \mathbf{G}_{II} &= \text{Ad}_{\hat{\mathbf{X}}_{I_{i+1}}} \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \Delta\hat{\mathbf{R}}_i^\top & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \Delta\hat{\mathbf{R}}_i^\top \end{bmatrix} \begin{bmatrix} -\mathbf{J}_r(\Delta\hat{\theta}_i) \mathbf{0}_{3 \times 3} \\ \Xi_4 & -\Xi_2 \\ \Xi_3 & -\Xi_1 \end{bmatrix} \\ &= \begin{bmatrix} -{}^L\hat{\mathbf{R}}_{I_{i+1}} \mathbf{J}_r(\Delta\hat{\theta}_i) \delta t & \mathbf{0}_{3 \times 3} \\ -({}^L\hat{\mathbf{p}}_{I_{i+1}}) \times {}^L\hat{\mathbf{R}}_{I_{i+1}} \mathbf{J}_r(\Delta\hat{\theta}_i) \delta t + \hat{\mathbf{R}}_{I_i} \Xi_4 & -\hat{\mathbf{R}}_{I_i} \Xi_2 \\ -({}^L\hat{\mathbf{v}}_{I_{i+1}}) \times {}^L\hat{\mathbf{R}}_{I_{i+1}} \mathbf{J}_r(\Delta\hat{\theta}_i) \delta t + \hat{\mathbf{R}}_{I_i} \Xi_3 & -\hat{\mathbf{R}}_{I_i} \Xi_1 \end{bmatrix}. \quad (10) \end{aligned}$$

On the other hand, the biases can be discretized as:

$$\mathbf{b}_{g_{i+1}} = \mathbf{b}_{g_i} + \mathbf{n}_{dbg} \delta t \quad (11)$$

$$\mathbf{b}_{a_{i+1}} = \mathbf{b}_{a_i} + \mathbf{n}_{dba} \delta t \quad (12)$$

where $\mathbf{n}_{dbg} \sim \mathcal{N}(\mathbf{0}_{3 \times 1}, \frac{\sigma_{bg}^2}{\delta t} \mathbf{I}_{3 \times 3})$ and $\mathbf{n}_{dba} \sim \mathcal{N}(\mathbf{0}_{3 \times 1}, \frac{\sigma_{ba}^2}{\delta t} \mathbf{I}_{3 \times 3})$ are the discretized noises. We can then derive:

$$\delta\mathbf{b}_{g_{i+1}} = \delta\mathbf{b}_{g_i} + \mathbf{n}_{dbg} \delta t \quad (13)$$

$$\delta\mathbf{b}_{a_{i+1}} = \delta\mathbf{b}_{a_i} + \mathbf{n}_{dba} \delta t. \quad (14)$$

Therefore, Φ_i^{i+1} and \mathbf{G}_i are formulated as:

$$\Phi_i^{i+1} = \begin{bmatrix} \Phi_{II} & \Phi_{IB} \\ \mathbf{0}_{6 \times 9} & \mathbf{I}_{6 \times 6} \end{bmatrix}, \quad \mathbf{G}_i = \begin{bmatrix} \mathbf{G}_{II} & \mathbf{0}_{9 \times 6} \\ \mathbf{0}_{6 \times 6} & \mathbf{I}_{6 \times 6} \delta t \end{bmatrix}. \quad (15)$$

APPENDIX B DERIVATION OF OBSERVATION JACOBIAN MATRICES

A. Odometer Observation

By replacing the state in error function **Eq. (14)** of **Section IV-B. Odometer Observation** with **Eq. (5)** of **Section III-B. Notation and State Definition** in the manuscript, the linearized observation error function is derived as:

$$\begin{aligned} \mathbf{z}_{o_k} &= \tilde{\mathbf{y}}_{o_k} - \hat{\mathbf{y}}_{o_k} \\ &= {}^O\mathbf{R}_I(\text{Exp}(\xi_{R_{LI_k}})^L \hat{\mathbf{R}}_{I_k})^\top \left[\text{Exp}(\xi_{R_{LI_k}})^L \hat{\mathbf{v}}_{I_k} \right. \\ &\quad \left. + \mathbf{J}_l(\xi_{R_{LI_k}}) \xi_{v_{LI_k}} \right] - {}^O\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top L \hat{\mathbf{v}}_{I_k} + \mathbf{n}_{o_k} \\ &\approx {}^O\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top \text{Exp}(\xi_{R_{LI_k}})^\top \left[\text{Exp}(\xi_{R_{LI_k}})^L \hat{\mathbf{v}}_{I_k} + \xi_{v_{LI_k}} \right] \\ &\quad - {}^O\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top L \hat{\mathbf{v}}_{I_k} + \mathbf{n}_{o_k} \\ &\approx {}^O\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top \text{Exp}(\xi_{R_{LI_k}})^\top \xi_{v_{LI_k}} + \mathbf{n}_{o_k} \\ &\approx {}^O\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top \xi_{v_{LI_k}} + \mathbf{n}_{o_k}. \end{aligned} \quad (16)$$

Then the Jacobian matrix \mathbf{H}_{o_k} of \mathbf{z}_{o_k} with respect to ξ_k is formulated as:

$$\mathbf{H}_{o_k} = \begin{bmatrix} \frac{\partial \mathbf{z}_{o_k}}{\partial \xi_{R_{LI_k}}} & \frac{\partial \mathbf{z}_{o_k}}{\partial \xi_{p_{LI_k}}} & \frac{\partial \mathbf{z}_{o_k}}{\partial \xi_{v_{LI_k}}} & \frac{\partial \mathbf{z}_{o_k}}{\partial \xi_{B_k}} & \frac{\partial \mathbf{z}_{o_k}}{\partial \xi_{C_k}} & \frac{\partial \mathbf{z}_{o_k}}{\partial \xi_{G_k}} \end{bmatrix} \quad (17)$$

$$= \begin{bmatrix} \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & {}^O\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top & \mathbf{0}_{3 \times 6} & \mathbf{0}_{3 \times 6c} & \mathbf{0}_{3 \times 9} \end{bmatrix}. \quad (18)$$

B. Feature Observation

By replacing the state in error function **Eq. (16)** of **Section IV-C. Feature Observation** with **Eq. (5)** of **Section III-B. Notation and State Definition** in the manuscript, the linearized observation error function is derived as:

$$\begin{aligned} \mathbf{z}_{f_{jk}} &= \tilde{\mathbf{y}}_{f_{jk}} - \hat{\mathbf{y}}_{f_{jk}} \\ &= h \left[{}^C\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top ({}^L \hat{\mathbf{p}}_{f_k} - {}^L \mathbf{p}_{I_k}) + {}^C\mathbf{p}_I \right] - \hat{\mathbf{y}}_{f_{jk}} + \mathbf{n}_{f_k} \\ &\approx \mathcal{H}_{j_k} \left[{}^C\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top ({}^L \hat{\mathbf{p}}_{f_k} - {}^L \mathbf{p}_{I_k}) + {}^C\mathbf{p}_I \right. \\ &\quad \left. - {}^C\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top ({}^L \hat{\mathbf{p}}_{f_k} - {}^L \hat{\mathbf{p}}_{I_k}) - {}^C\mathbf{p}_I \right] + \mathbf{n}_{f_k} \\ &= \mathcal{H}_{j_k} \left[{}^C\mathbf{R}_I (\text{Exp}(\xi_{R_{LI_k}})^L \hat{\mathbf{R}}_{I_k})^\top \left[\text{Exp}(\xi_{R_{LG_k}})^L \hat{\mathbf{p}}_{f_k} + \right. \right. \\ &\quad \left. \left. \mathbf{J}_l(\xi_{R_{LG_k}}) \xi_{p_{Lf_k}} - \text{Exp}(\xi_{R_{LI_k}})^L \hat{\mathbf{p}}_{I_k} - \mathbf{J}_l(\xi_{R_{LI_k}}) \xi_{p_{LI_k}} \right] \right. \\ &\quad \left. - {}^C\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top ({}^L \hat{\mathbf{p}}_{f_k} - {}^L \hat{\mathbf{p}}_{I_k}) \right] + \mathbf{n}_{f_k} \\ &\approx \mathcal{H}_{j_k} \left[{}^C\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top (\mathbf{I}_{3 \times 3} + (\xi_{R_{LI_k}}) \times) [(\mathbf{I}_{3 \times 3} \right. \\ &\quad \left. - (\xi_{R_{LG_k}}) \times) {}^L \hat{\mathbf{p}}_{f_k} + \xi_{p_{Lf_k}} - (\mathbf{I}_{3 \times 3} + (\xi_{R_{LI_k}}) \times) {}^L \hat{\mathbf{p}}_{I_k} \right. \\ &\quad \left. - \xi_{p_{LI_k}}] - {}^C\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top ({}^L \hat{\mathbf{p}}_{f_k} - {}^L \hat{\mathbf{p}}_{I_k}) \right] + \mathbf{n}_{f_k} \\ &\approx \mathcal{H}_{j_k} {}^C\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top \left[({}^L \hat{\mathbf{p}}_{f_k}) \times \xi_{R_{LI_k}} - ({}^L \hat{\mathbf{p}}_{f_k}) \times \xi_{R_{LG_k}} + \xi_{p_{LI_k}} \right. \\ &\quad \left. - \xi_{p_{Lf_k}} \right] + \mathbf{n}_{f_k}. \end{aligned} \quad (19)$$

The Jacobian matrices $\mathbf{H}_{f_{jk}}$ of $\mathbf{z}_{f_{jk}}$ with respect to ξ_k and $\xi_{p_{Lf_k}}$ are formulated as:

$$\begin{aligned} \mathbf{H}_{f_{jk}} &= \begin{bmatrix} \frac{\partial \mathbf{z}_{f_{jk}}}{\partial \xi_{I_k}} & \frac{\partial \mathbf{z}_{f_{jk}}}{\partial \xi_{B_k}} & \frac{\partial \mathbf{z}_{f_{jk}}}{\partial \xi_{C_k}} & \frac{\partial \mathbf{z}_{f_{jk}}}{\partial \xi_{T_{LG_k}}} & \frac{\partial \mathbf{z}_{f_{jk}}}{\partial \xi_{p_{Lf_k}}} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{0}_{2 \times 9} & \mathbf{0}_{2 \times 6} & \mathbf{H}_{f_{jk}}^{\mathbf{X}_C} & \mathbf{H}_{f_{jk}}^{\mathbf{X}_G} & \mathbf{H}_{f_{jk}}^{\mathbf{p}_{Lf_k}} \end{bmatrix} \end{aligned} \quad (20)$$

$$\mathbf{H}_{f_{jk}}^{\mathbf{X}_C} = \mathcal{H}_{j_k} \left[\mathbf{0}_{3 \times 6} \dots {}^C\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k} ({}^L \hat{\mathbf{p}}_f) \times \dots \right. \\ \left. - {}^C\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k} \dots \mathbf{0}_{3 \times 6} \right] \quad (21)$$

C. Streetlight Observation

By replacing the state in error function **Eq. (25)** of **Section V-B. Data Association and Streetlight Observation** with **Eq. (5)** of **Section III-B. Notation and State Definition** in the manuscript, the observation error function is linearized as:

$$\begin{aligned} \mathbf{z}_{s_k} &= \tilde{\mathbf{y}}_{s_k} - \hat{\mathbf{y}}_{s_k} \\ &= h \left[{}^C\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top ({}^L \mathbf{R}_{G_k} G \tilde{\mathbf{c}}_s + {}^L \mathbf{p}_{G_k} - {}^L \mathbf{p}_{I_k}) + {}^C\mathbf{p}_I \right] - \hat{\mathbf{y}}_{s_k} + \mathbf{n}_{s_k} \\ &\approx \mathcal{H}_{k_k} \left[{}^C\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top ({}^L \mathbf{R}_{G_k} G \tilde{\mathbf{c}}_s + {}^L \mathbf{p}_{G_k} - {}^L \mathbf{p}_{I_k}) + {}^C\mathbf{p}_I \right. \\ &\quad \left. - {}^C\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top ({}^L \hat{\mathbf{R}}_{G_k} G \tilde{\mathbf{c}}_s + {}^L \hat{\mathbf{p}}_{G_k} - {}^L \hat{\mathbf{p}}_{I_k}) - {}^C\mathbf{p}_I \right] + \mathbf{n}_{s_k} \\ &= \mathcal{H}_{k_k} \left[{}^C\mathbf{R}_I (\text{Exp}(\xi_{R_{LI_k}})^L \hat{\mathbf{R}}_{I_k})^\top \left[\text{Exp}(\xi_{R_{LG_k}})^L \hat{\mathbf{R}}_{G_k} G \tilde{\mathbf{c}}_s + \right. \right. \\ &\quad \left. \left. \text{Exp}(\xi_{R_{LG_k}})^L \hat{\mathbf{p}}_{G_k} + \mathbf{J}_l(\xi_{R_{LG_k}}) \xi_{p_{LG_k}} - \text{Exp}(\xi_{R_{LI_k}})^L \hat{\mathbf{p}}_{I_k} \right. \right. \\ &\quad \left. \left. - \mathbf{J}_l(\xi_{R_{LI_k}}) \xi_{p_{LI_k}} \right] - {}^C\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top ({}^L \hat{\mathbf{R}}_{G_k} G \tilde{\mathbf{c}}_s + {}^L \hat{\mathbf{p}}_{G_k} \right. \\ &\quad \left. \left. - {}^L \hat{\mathbf{p}}_{I_k}) \right] + \mathbf{n}_{s_k} \right. \\ &\approx \mathcal{H}_{k_k} \left[{}^C\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top (\mathbf{I}_{3 \times 3} - (\xi_{R_{LI_k}}) \times) \left[(\mathbf{I}_{3 \times 3} + (\xi_{R_{LG_k}}) \times) {}^L \hat{\mathbf{R}}_{G_k} G \tilde{\mathbf{c}}_s \right. \right. \\ &\quad \left. \left. + (\mathbf{I}_{3 \times 3} + (\xi_{R_{LG_k}}) \times) {}^L \hat{\mathbf{p}}_{G_k} + \xi_{p_{LG_k}} - (\mathbf{I}_{3 \times 3} + (\xi_{R_{LI_k}}) \times) {}^L \hat{\mathbf{p}}_{I_k} \right. \right. \\ &\quad \left. \left. - \xi_{p_{LI_k}} \right] - {}^C\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top ({}^L \hat{\mathbf{R}}_{G_k} G \tilde{\mathbf{c}}_s + {}^L \hat{\mathbf{p}}_{G_k} - {}^L \hat{\mathbf{p}}_{I_k}) \right] + \mathbf{n}_{s_k} \\ &\approx \mathcal{H}_{k_k}^C {}^C\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top \left[({}^L \hat{\mathbf{R}}_{G_k} G \tilde{\mathbf{c}}_s + {}^L \hat{\mathbf{p}}_{G_k}) \times (\xi_{R_{LI_k}} - \xi_{R_{LG_k}}) \right. \\ &\quad \left. + \xi_{p_{LG_k}} - \xi_{p_{LI_k}} \right] + \mathbf{n}_{s_k}. \end{aligned} \quad (22)$$

Then the Jacobian matrix \mathbf{H}_{s_k} of \mathbf{z}_{s_k} with respect to ξ_k is formulated as:

$$\begin{aligned} \mathbf{H}_{s_k} &= \begin{bmatrix} \frac{\partial \mathbf{z}_{s_k}}{\partial \xi_{I_k}} & \frac{\partial \mathbf{z}_{s_k}}{\partial \xi_{B_k}} & \frac{\partial \mathbf{z}_{s_k}}{\partial \xi_{C_k}} & \frac{\partial \mathbf{z}_{s_k}}{\partial \xi_{T_{LG_k}}} & \frac{\partial \mathbf{z}_{s_k}}{\partial \xi_{p_{Lf_k}}} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{H}_{s_k}^{\mathbf{X}_I} & \mathbf{0}_{2 \times 6} & \mathbf{0}_{2 \times 6c} & \mathbf{H}_{s_k}^{\mathbf{X}_{TG}} & \mathbf{0}_{2 \times 3} \end{bmatrix} \end{aligned} \quad (23)$$

$$\mathbf{H}_{s_k}^{\mathbf{X}_I} = \mathcal{H}_{k_k} \left[{}^C\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top ({}^L \hat{\mathbf{c}}_s) \times \dots \right. \\ \left. - {}^C\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top \dots \mathbf{0}_{3 \times 3} \right] \quad (24)$$

$$\mathbf{H}_{s_k}^{\mathbf{X}_{TG}} = \mathcal{H}_{k_k} \left[- {}^C\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top ({}^L \hat{\mathbf{c}}_s) \times \dots \right. \\ \left. {}^C\mathbf{R}_I^L \hat{\mathbf{R}}_{I_k}^\top \dots \right]. \quad (24)$$

D. Prior Pose Observation

By replacing the state in error function **Eq. (27)** of **Section V-C. Prior Pose Observation** with **Eq. (5)** of **Section III-B. Notation and State Definition** in the manuscript, the linearized observation error function is derived as:

$$\begin{aligned} \mathbf{z}_{p_k} &= \tilde{\mathbf{y}}_{p_k} - \hat{\mathbf{y}}_{p_k} \\ &= \begin{bmatrix} \mathbf{e}_3^\top {}^O\mathbf{R}_I {}^G\tilde{\mathbf{R}}_{I'}^\top ({}^L \mathbf{R}_{G_k}^\top ({}^L \mathbf{p}_{I_k} - {}^L \mathbf{p}_{G_k}) - {}^G \tilde{\mathbf{p}}_{I'}) \\ \mathbf{e}_3^\top {}^O\mathbf{R}_I {}^G\tilde{\mathbf{R}}_{I'}^\top {}^L \mathbf{R}_{G_k}^\top {}^O\mathbf{R}_I^\top \mathbf{e}_3 \end{bmatrix} \end{aligned} \quad (25)$$

$$-\begin{bmatrix} \mathbf{e}_3^\top \mathbf{R}_I^G \tilde{\mathbf{R}}_{I'}^\top (\mathbf{L}\hat{\mathbf{R}}_{G_k}^\top (\mathbf{L}\hat{\mathbf{p}}_{I_k} - \mathbf{L}\hat{\mathbf{p}}_{G_k}) - {}^G\tilde{\mathbf{p}}_{I'}) \\ \mathbf{e}_3^\top \mathbf{R}_I^G \tilde{\mathbf{R}}_{I'}^\top \mathbf{L}\hat{\mathbf{R}}_{G_k}^\top {}^L\hat{\mathbf{R}}_{I_k} \mathbf{R}_I^\top \mathbf{e}_3 \end{bmatrix} + \mathbf{n}_{p_k}.$$

For the first element of \mathbf{z}_{p_k} , we have:

$$\begin{aligned} z_{p_k,1} &= \mathbf{e}_3^\top \mathbf{R}_I^G \tilde{\mathbf{R}}_{I'}^\top (\mathbf{L}\hat{\mathbf{R}}_{G_k}^\top (\mathbf{L}\hat{\mathbf{p}}_{I_k} - \mathbf{L}\hat{\mathbf{p}}_{G_k}) - {}^G\tilde{\mathbf{p}}_{I'}) - \\ &\quad - \mathbf{e}_3^\top \mathbf{R}_I^G \tilde{\mathbf{R}}_{I'}^\top (\mathbf{L}\hat{\mathbf{R}}_{G_k}^\top (\mathbf{L}\hat{\mathbf{p}}_{I_k} - \mathbf{L}\hat{\mathbf{p}}_{G_k}) - {}^G\tilde{\mathbf{p}}_{I'}) \\ &= \mathbf{e}_3^\top \mathbf{R}_I^G \tilde{\mathbf{R}}_{I'}^\top \left[(\text{Exp}(\boldsymbol{\xi}_{R_{LI_k}}) \mathbf{L}\hat{\mathbf{R}}_{G_k})^\top \left[(\text{Exp}(\boldsymbol{\xi}_{R_{LI_k}}) \mathbf{L}\hat{\mathbf{p}}_{I_k} \right. \right. \\ &\quad \left. + \mathbf{J}_l(\boldsymbol{\xi}_{R_{LI_k}}) \boldsymbol{\xi}_{p_{LI_k}} - \text{Exp}(\boldsymbol{\xi}_{R_{LG_k}}) \mathbf{L}\hat{\mathbf{p}}_{G_k} - \mathbf{J}_l(\boldsymbol{\xi}_{R_{LG_k}}) \boldsymbol{\xi}_{p_{LG_k}} \right. \\ &\quad \left. - {}^G\tilde{\mathbf{p}}_{I'} - {}^L\hat{\mathbf{R}}_{G_k}^\top (\mathbf{L}\hat{\mathbf{p}}_{I_k} - \mathbf{L}\hat{\mathbf{p}}_{G_k}) + {}^G\tilde{\mathbf{p}}_{I'} \right] + n_{p_k,1} \\ &\approx \mathbf{e}_3^\top \mathbf{R}_I^G \tilde{\mathbf{R}}_{I'}^\top \left[{}^L\hat{\mathbf{R}}_{G_k}^\top (\mathbf{I}_{3 \times 3} - (\boldsymbol{\xi}_{R_{LG_k}}) \times) \right] [(\mathbf{I}_{3 \times 3} - \\ &\quad + (\boldsymbol{\xi}_{R_{LI_k}}) \times) \mathbf{L}\hat{\mathbf{p}}_{I_k} + \boldsymbol{\xi}_{p_{LI_k}} - (\mathbf{I}_{3 \times 3} + (\boldsymbol{\xi}_{R_{LG_k}}) \times) \mathbf{L}\hat{\mathbf{p}}_{G_k} \\ &\quad - \boldsymbol{\xi}_{p_{LG_k}}] - {}^L\hat{\mathbf{R}}_{G_k}^\top (\mathbf{L}\hat{\mathbf{p}}_{I_k} - \mathbf{L}\hat{\mathbf{p}}_{G_k}) \right] + n_{p_k,1} \\ &\approx \mathbf{e}_3^\top \mathbf{R}_I^G \tilde{\mathbf{R}}_{I'}^\top {}^L\hat{\mathbf{R}}_{G_k}^\top \left[(\mathbf{L}\hat{\mathbf{p}}_{I_k}) \times \boldsymbol{\xi}_{R_{LG_k}} - (\mathbf{L}\hat{\mathbf{p}}_{I_k}) \times \boldsymbol{\xi}_{R_{LI_k}} \right. \\ &\quad \left. + \boldsymbol{\xi}_{p_{LI_k}} - \boldsymbol{\xi}_{p_{LG_k}} \right] + n_{p_k,1}. \end{aligned} \quad (26)$$

For the second element of \mathbf{z}_{p_k} , we have:

$$\begin{aligned} z_{p_k,2} &= \mathbf{e}_3^\top \mathbf{R}_I^G \tilde{\mathbf{R}}_{I'}^\top {}^L\hat{\mathbf{R}}_{G_k}^\top {}^L\mathbf{R}_{I_k} \mathbf{R}_I^\top \mathbf{e}_3 \\ &\quad - \mathbf{e}_3^\top \mathbf{R}_I^G \tilde{\mathbf{R}}_{I'}^\top {}^L\hat{\mathbf{R}}_{G_k}^\top {}^L\mathbf{R}_{I_k} \mathbf{R}_I^\top \mathbf{e}_3 + n_{p_k,2} \\ &= \mathbf{e}_3^\top \mathbf{R}_I^G \tilde{\mathbf{R}}_{I'}^\top (\text{Exp}(\boldsymbol{\xi}_{R_{LG_k}}) \mathbf{L}\hat{\mathbf{R}}_{G_k})^\top (\text{Exp}(\boldsymbol{\xi}_{R_{LI_k}}) \mathbf{L}\hat{\mathbf{R}}_{I_k}) \\ &\quad \mathbf{R}_I^\top \mathbf{e}_3 - \mathbf{e}_3^\top \mathbf{R}_I^G \tilde{\mathbf{R}}_{I'}^\top {}^L\hat{\mathbf{R}}_{G_k}^\top {}^L\mathbf{R}_{I_k} \mathbf{R}_I^\top \mathbf{e}_3 + n_{p_k,2} \\ &\approx \mathbf{e}_3^\top \mathbf{R}_I^G \tilde{\mathbf{R}}_{I'}^\top {}^L\hat{\mathbf{R}}_{G_k}^\top (\mathbf{I}_{3 \times 3} - (\boldsymbol{\xi}_{R_{LG_k}}) \times) (\mathbf{I}_{3 \times 3} + (\boldsymbol{\xi}_{R_{LI_k}}) \times) \\ &\quad {}^L\hat{\mathbf{R}}_{I_k} \mathbf{R}_I^\top \mathbf{e}_3 - \mathbf{e}_3^\top \mathbf{R}_I^G \tilde{\mathbf{R}}_{I'}^\top {}^L\hat{\mathbf{R}}_{G_k}^\top {}^L\hat{\mathbf{R}}_{I_k} \mathbf{R}_I^\top \mathbf{e}_3 + n_{p_k,2} \\ &\approx \mathbf{e}_3^\top \mathbf{R}_I^G \tilde{\mathbf{R}}_{I'}^\top {}^L\hat{\mathbf{R}}_{G_k}^\top ((\boldsymbol{\xi}_{R_{LI_k}}) \times - (\boldsymbol{\xi}_{R_{LG_k}}) \times) {}^L\hat{\mathbf{R}}_{I_k} \mathbf{R}_I^\top \mathbf{e}_3 \\ &\quad + n_{p_k,2} \\ &= \mathbf{e}_3^\top \mathbf{R}_I^G \tilde{\mathbf{R}}_{I'}^\top {}^L\hat{\mathbf{R}}_{G_k}^\top (\mathbf{L}\hat{\mathbf{R}}_{I_k} \mathbf{R}_I^\top \mathbf{e}_3) \times \boldsymbol{\xi}_{R_{LG_k}} - \mathbf{e}_3^\top \mathbf{R}_I^G \tilde{\mathbf{R}}_{I'}^\top \\ &\quad {}^L\hat{\mathbf{R}}_{G_k}^\top (\mathbf{L}\hat{\mathbf{R}}_{I_k} \mathbf{R}_I^\top \mathbf{e}_3) \times \boldsymbol{\xi}_{R_{LI_k}} + n_{p_k,2}. \end{aligned} \quad (27)$$

Then the Jacobian matrix \mathbf{H}_{p_k} of \mathbf{z}_{p_k} with respect to $\boldsymbol{\xi}_k$ is formulated as:

$$\begin{aligned} \mathbf{H}_{p_k} &= \begin{bmatrix} \frac{\partial \mathbf{z}_{p_k}}{\partial \boldsymbol{\xi}_{I_k}} & \frac{\partial \mathbf{z}_{p_k}}{\partial \boldsymbol{\xi}_{B_k}} & \frac{\partial \mathbf{z}_{p_k}}{\partial \boldsymbol{\xi}_{C_k}} & \frac{\partial \mathbf{z}_{p_k}}{\partial \boldsymbol{\xi}_{T_{LG_k}}} & \frac{\partial \mathbf{z}_{p_k}}{\partial \boldsymbol{\xi}_{p_{LF_k}}} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{H}_{p_k}^{\mathbf{X}_I} & \mathbf{0}_{2 \times 6} & \mathbf{0}_{2 \times 6} & \mathbf{H}_{p_k}^{\mathbf{X}_{TG}} & \mathbf{0}_{2 \times 3} \end{bmatrix} \end{aligned} \quad (28)$$

$$\mathbf{H}_{p_k}^{\mathbf{X}_I} = \mathbf{e}_3^\top \mathbf{R}_O^G \tilde{\mathbf{R}}_{I'}^\top \begin{bmatrix} {}^L\hat{\mathbf{R}}_{G_k}^\top (\mathbf{L}\hat{\mathbf{p}}_{I_k}) \times & {}^L\hat{\mathbf{R}}_{G_k} \mathbf{0}_{1 \times 3} \\ - {}^L\hat{\mathbf{R}}_{G_k}^\top (\mathbf{L}\hat{\mathbf{R}}_{I_k} \mathbf{R}_I^\top \mathbf{e}_3) \times & \mathbf{0}_{1 \times 3} \mathbf{0}_{1 \times 3} \end{bmatrix}$$

$$\mathbf{H}_{p_k}^{\mathbf{X}_{TG}} = \mathbf{e}_3^\top \mathbf{R}_O^G \tilde{\mathbf{R}}_{I'}^\top \begin{bmatrix} {}^L\hat{\mathbf{R}}_{G_k}^\top (\mathbf{L}\hat{\mathbf{p}}_{I_k}) \times & - {}^L\hat{\mathbf{R}}_{G_k} \\ {}^L\hat{\mathbf{R}}_{G_k}^\top (\mathbf{L}\hat{\mathbf{R}}_{I_k} \mathbf{R}_I^\top \mathbf{e}_3) \times & \mathbf{0}_{1 \times 3} \end{bmatrix}. \quad (29)$$

APPENDIX C

DERIVATION OF VARIANCE $\sigma_{p,dg}^2$ AND $\sigma_{a,dg}^2$

To derive the formulation of variance in a more compact form, we use the transformation matrix form to represent the pose. Hence the reprojection error in Eq. (20) of Section V-B. Data Association and Streetlight Observation in the manuscript can be rewritten as

$$\mathbf{r}_{p,dg} = \tilde{\mathbf{p}}_{s_{dg}} - \mathbf{h}({}^C\mathbf{T}_I {}^L\hat{\mathbf{T}}_{I_k}^{-1} {}^L\hat{\mathbf{T}}_{G_k} {}^G\tilde{\mathbf{c}}_{s_g}) \quad (30)$$

where ${}^G\tilde{\mathbf{c}}_{s_g}$ is the homogeneous coordinate of ${}^G\mathbf{c}_{s_g}$. The error states of ${}^L\hat{\mathbf{T}}_{I_k}$ and ${}^L\hat{\mathbf{T}}_{G_k}$ are respectively denoted as $\boldsymbol{\xi}_{T_{LI_k}} = [\boldsymbol{\xi}_{R_{LI_k}}^\top \ \boldsymbol{\xi}_{p_{LI_k}}^\top]^\top$ and $\boldsymbol{\xi}_{T_{LG_k}} = [\boldsymbol{\xi}_{R_{LG_k}}^\top \ \boldsymbol{\xi}_{p_{LG_k}}^\top]^\top$. We denote $r_{p,dg} = \|\mathbf{r}_{p,dg}\|_2$. The Jacobian matrices of $\delta r_{p,dg}$ with respect to $\boldsymbol{\xi}_{T_{LI_k}}$ and $\boldsymbol{\xi}_{T_{LG_k}}$ are derived as:

$$\begin{aligned} \frac{\partial \delta r_{p,dg}}{\partial \boldsymbol{\xi}_{T_{LI_k}}} &= - \frac{(\tilde{\mathbf{p}}_{s_{dg}} - \mathbf{h}({}^C\mathbf{T}_I {}^L\hat{\mathbf{T}}_{I_k}^{-1} {}^L\hat{\mathbf{T}}_{G_k} {}^G\tilde{\mathbf{c}}_{s_g}))^\top \mathbf{H}_k \boldsymbol{\Pi}_1 \frac{\partial \delta {}^C\tilde{\mathbf{c}}_{s_g}}{\partial \boldsymbol{\xi}_{T_{LI_k}}}}{r_{p,dg}} \\ \frac{\partial \delta {}^C\tilde{\mathbf{c}}_{s_g}}{\partial \boldsymbol{\xi}_{T_{LI_k}}} &= \lim_{\boldsymbol{\xi}_{T_{LI_k}} \rightarrow 0} \frac{({}^C\mathbf{T}_I \text{Exp}(-\boldsymbol{\xi}_{T_{LI_k}}) {}^L\hat{\mathbf{T}}_{I_k}^{-1} - {}^C\mathbf{T}_I {}^L\hat{\mathbf{T}}_{I_k}^{-1}) {}^L\hat{\mathbf{T}}_{G_k} {}^G\tilde{\mathbf{c}}_{s_g}}{\boldsymbol{\xi}_{T_{LI_k}}} \\ &= \lim_{\boldsymbol{\xi}_{T_{LI_k}} \rightarrow 0} \frac{- {}^C\mathbf{T}_I {}^L\hat{\mathbf{T}}_{I_k}^{-1} {}^L\hat{\mathbf{T}}_{G_k} (\text{Ad}({}^L\hat{\mathbf{T}}_{G_k}^{-1} {}^L\hat{\mathbf{T}}_{I_k})) \boldsymbol{\xi}_{T_{LI_k}}) \wedge {}^G\tilde{\mathbf{c}}_{s_g}}{\boldsymbol{\xi}_{T_{LI_k}}} \\ &= - {}^C\mathbf{T}_I {}^L\hat{\mathbf{T}}_{I_k}^{-1} {}^L\hat{\mathbf{T}}_{G_k} {}^G\tilde{\mathbf{c}}_{s_g}^\odot \text{Ad}({}^L\hat{\mathbf{T}}_{G_k}^{-1} {}^L\hat{\mathbf{T}}_{I_k}) \\ \frac{\partial \delta r_{p,dg}}{\partial \boldsymbol{\xi}_{T_{LG_k}}} &= - \frac{(\tilde{\mathbf{p}}_{s_{dg}} - \mathbf{h}({}^C\mathbf{T}_I {}^L\hat{\mathbf{T}}_{I_k}^{-1} {}^L\hat{\mathbf{T}}_{G_k} {}^G\tilde{\mathbf{c}}_{s_g}))^\top \mathbf{H}_k \boldsymbol{\Pi}_1 \frac{\partial \delta {}^C\tilde{\mathbf{c}}_{s_g}}{\partial \boldsymbol{\xi}_{T_{LG_k}}}}{r_{p,dg}} \\ \frac{\partial \delta {}^C\tilde{\mathbf{c}}_{s_g}}{\partial \boldsymbol{\xi}_{T_{LG_k}}} &= \lim_{\boldsymbol{\xi}_{T_{LG_k}} \rightarrow 0} \frac{{}^C\mathbf{T}_I {}^L\hat{\mathbf{T}}_{I_k}^{-1} (\text{Exp}(\boldsymbol{\xi}_{T_{LG_k}}) {}^L\hat{\mathbf{T}}_{G_k} {}^G\tilde{\mathbf{c}}_{s_g} - {}^L\hat{\mathbf{T}}_{G_k} {}^G\tilde{\mathbf{c}}_{s_g})}{\boldsymbol{\xi}_{T_{LG_k}}} \\ &= \lim_{\boldsymbol{\xi}_{T_{LG_k}} \rightarrow 0} \frac{{}^C\mathbf{T}_I {}^L\hat{\mathbf{T}}_{I_k}^{-1} {}^L\hat{\mathbf{T}}_{G_k} (\text{Ad}({}^L\hat{\mathbf{T}}_{G_k}^{-1})) \boldsymbol{\xi}_{T_{LG_k}}) \wedge {}^G\tilde{\mathbf{c}}_{s_g}}{\boldsymbol{\xi}_{T_{LG_k}}} \\ &= {}^C\mathbf{T}_I {}^L\hat{\mathbf{T}}_{I_k}^{-1} {}^L\hat{\mathbf{T}}_{G_k} {}^G\tilde{\mathbf{c}}_{s_g}^\odot \text{Ad}({}^L\hat{\mathbf{T}}_{G_k}^{-1}) \end{aligned} \quad (31)$$

where ${}^G\tilde{\mathbf{c}}_{s_g}^\odot = \begin{bmatrix} \mathbf{I}_{3 \times 3} & -({}^G\tilde{\mathbf{c}}_{s_g}) \times \\ \mathbf{0}_{1 \times 3} & \mathbf{0}_{1 \times 3} \end{bmatrix}$ and $\boldsymbol{\Pi}_1 = [\mathbf{I}_{3 \times 3} \ \mathbf{0}_{3 \times 1}]$.

$\mathbf{P}_{LG_k}^{LI_k}$ denotes the covariance matrix block related to $\boldsymbol{\xi}_{T_{LI_k}}$ and $\boldsymbol{\xi}_{T_{LG_k}}$. $\text{Ad}(\cdot)$ represents the adjacent operator which can be formulated as the adjoint matrix for $SE(3)$: $\text{Ad}(\mathbf{T}) = \begin{bmatrix} \mathbf{R} & \mathbf{0}_{3 \times 3} \\ (\mathbf{p})_\times \mathbf{R} & \mathbf{R} \end{bmatrix}$. Therefore, the variance is as:

$$\sigma_{p,dg}^2 = \begin{bmatrix} \frac{\partial \delta r_{p,dg}}{\partial \boldsymbol{\xi}_{T_{LI_k}}} & \frac{\partial \delta r_{p,dg}}{\partial \boldsymbol{\xi}_{T_{LG_k}}} \end{bmatrix} \mathbf{P}_{LG_k}^{LI_k} \begin{bmatrix} \frac{\partial \delta r_{p,dg}}{\partial \boldsymbol{\xi}_{T_{LI_k}}} & \frac{\partial \delta r_{p,dg}}{\partial \boldsymbol{\xi}_{T_{LG_k}}} \end{bmatrix}^\top + \boldsymbol{\Sigma}_{s_d} \quad (32)$$

where $\boldsymbol{\Sigma}_{s_d}$ is the covariance corresponding to the box center.

Similar to $\sigma_{p,dg}^2$, $\sigma_{a,dg}^2$ is also related to ${}^L\hat{\mathbf{T}}_{I_k}$ and ${}^L\hat{\mathbf{T}}_{G_k}$. We denote $r_{a,dg} = \|\mathbf{r}_{a,dg}\|_2$ and the Jacobian matrices of $\delta r_{a,dg}$ with respect to $\boldsymbol{\xi}_{T_{LI_k}}$ and $\boldsymbol{\xi}_{T_{LG_k}}$ are:

$$\begin{aligned} \frac{\partial \delta r_{a,dg}}{\partial \boldsymbol{\xi}_{T_{LI_k}}} &= \frac{({}^C\tilde{\mathbf{p}}_{s_{dg}} \times {}^C\tilde{\mathbf{c}}_{s_g})^\top} {\|{}^C\tilde{\mathbf{p}}_{s_{dg}} \times {}^C\tilde{\mathbf{c}}_{s_g}\|_2} ({}^C\tilde{\mathbf{p}}_{s_{dg}}) \times \left(\frac{\mathbf{I}_{3 \times 3}}{\|{}^C\tilde{\mathbf{c}}_{s_g}\|_2} \right. \\ &\quad \left. - \frac{{}^C\tilde{\mathbf{c}}_{s_g} {}^C\tilde{\mathbf{c}}_{s_g}^\top}{\|{}^C\tilde{\mathbf{c}}_{s_g}\|_2^3} \right) \boldsymbol{\Pi}_1 \frac{\partial \delta {}^C\tilde{\mathbf{c}}_{s_g}}{\partial \boldsymbol{\xi}_{T_{LI_k}}} \end{aligned} \quad (33)$$

$$\begin{aligned} \frac{\partial \delta r_{a,dg}}{\partial \boldsymbol{\xi}_{T_{LG_k}}} &= \frac{({}^C\tilde{\mathbf{p}}_{s_{dg}} \times {}^C\tilde{\mathbf{c}}_{s_g})^\top} {\|{}^C\tilde{\mathbf{p}}_{s_{dg}} \times {}^C\tilde{\mathbf{c}}_{s_g}\|_2} ({}^C\tilde{\mathbf{p}}_{s_{dg}}) \times \left(\frac{\mathbf{I}_{3 \times 3}}{\|{}^C\tilde{\mathbf{c}}_{s_g}\|_2} \right. \\ &\quad \left. - \frac{{}^C\tilde{\mathbf{c}}_{s_g} {}^C\tilde{\mathbf{c}}_{s_g}^\top}{\|{}^C\tilde{\mathbf{c}}_{s_g}\|_2^3} \right) \boldsymbol{\Pi}_1 \frac{\partial \delta {}^C\tilde{\mathbf{c}}_{s_g}}{\partial \boldsymbol{\xi}_{T_{LG_k}}}. \end{aligned} \quad (34)$$

The Jacobian matrix of $\delta r_{a,dg}$ with respect to $\delta {}^C\mathbf{p}_{s_{dg}}$ is derived as:

$$\frac{\partial \delta r_{a,dg}}{\partial \delta {}^C\mathbf{p}_{s_{dg}}} = \frac{({}^C\tilde{\mathbf{p}}_{s_{dg}} \times {}^C\tilde{\mathbf{c}}_{s_g})^\top} {\|{}^C\tilde{\mathbf{p}}_{s_{dg}} \times {}^C\tilde{\mathbf{c}}_{s_g}\|_2} ({}^C\tilde{\mathbf{c}}_{s_g}) \times \left(\frac{\mathbf{I}_{3 \times 3}}{\|{}^C\tilde{\mathbf{p}}_{s_{dg}}\|_2} \right)$$

$$-\frac{C\widetilde{\mathbf{p}}_{s_{dk}}^T C\widetilde{\mathbf{p}}_{s_{dk}}^T}{\|C\widetilde{\mathbf{p}}_{s_{dk}}\|_2^3}. \quad (35)$$

The final variance is propagated by:

$$\begin{aligned} \sigma_{a,dg}^2 &= \left[\frac{\partial \delta r_{a,dg}}{\partial \xi_{T_{LI_k}}} \quad \frac{\partial \delta r_{a,dg}}{\partial \xi_{T_{LG_k}}} \right] \mathbf{P}_{LG_k}^{LI_k} \left[\frac{\partial \delta r_{a,dg}}{\partial \xi_{T_{LI_k}}^T} \quad \frac{\partial \delta r_{a,dg}}{\partial \xi_{T_{LG_k}}^T} \right]^T \\ &+ \frac{\partial \delta r_{a,dg}}{\partial \delta^C \mathbf{p}_{s_{dk}}} \boldsymbol{\Pi}_2 \boldsymbol{\Sigma}_{s_d} \boldsymbol{\Pi}_2^T \frac{\partial \delta r_{a,dg}}{\partial \delta^C \mathbf{p}_{s_{dk}}} \end{aligned} \quad (36)$$

where $\boldsymbol{\Pi}_2 = [\mathbf{I}_{2 \times 2} \quad \mathbf{0}_{2 \times 1}]^T$.

APPENDIX D MSC-InEKF-FC

A. State Representation

For MSC-InEKF-FC, the features are associated with IMU navigation state to form $SE_{2+K}(3)$ group. The system state is defined as:

$$\begin{aligned} \mathbf{X}_k &= (\mathbf{X}_{I_k}, \mathbf{B}_k, \mathbf{X}_{C_k}, \mathbf{X}_{G_k}) \\ \mathbf{X}_{I_k} &= \left[\begin{array}{c|c} {}^L\mathbf{R}_{I_k} & {}^L\mathbf{p}_{I_k} \\ \hline \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{array} \right. \left. {}^L\mathbf{v}_{I_k} \quad {}^L\mathbf{p}_{f_k} \right] \\ \mathbf{B}_k &= [\mathbf{b}_{g_k}^\top \quad \mathbf{b}_{a_k}^\top]^T \\ \mathbf{X}_{C_k} &= (\mathbf{X}_{CP_{k_k}}, \dots, \mathbf{X}_{CP_{k-c+1_k}}), \quad \mathbf{X}_{CP_{j_k}} = \left[\begin{array}{c|c} {}^L\mathbf{R}_{I_{j_k}} & {}^L\mathbf{p}_{I_{j_k}} \\ \hline \mathbf{0}_{1 \times 3} & 1 \end{array} \right] \\ \mathbf{X}_{G_k} &= \left[\begin{array}{c|c} {}^L\mathbf{R}_{G_k} & {}^L\mathbf{p}_{G_k} \\ \hline \mathbf{0}_{1 \times 3} & 1 \end{array} \right]. \end{aligned} \quad (37)$$

We keep the error states in the same formula with that in MSC-InEKF-FDRC, i.e., $\xi_k = [\xi_{I_k}^\top \quad \xi_{B_k}^\top \quad \xi_{C_k}^\top \quad \xi_{T_{LG_k}}^\top \quad \xi_{p_{Lf_k}}^\top]^\top$. The feature error state $\xi_{p_{Lf_k}}$ is formulated as:

$${}^L\mathbf{p}_{f_k} = \text{Exp}(\xi_{R_{LI_k}}) {}^L\widehat{\mathbf{p}}_{f_k} + \mathbf{J}_l(\xi_{R_{LI_k}}) \xi_{p_{Lf_k}}. \quad (38)$$

Since $\mathbf{X}_{CP_{k_k}}$ is the clone of current pose, (38) is equivalent to:

$${}^L\mathbf{p}_{f_k} = \text{Exp}(\xi_{R_{CP_{k_k}}}) {}^L\widehat{\mathbf{p}}_{f_k} + \mathbf{J}_l(\xi_{p_{CP_{k_k}}}) \xi_{p_{Lf_k}}. \quad (39)$$

B. State Propagation

We also only investigate the states \mathbf{X}_{I_i} and \mathbf{B}_i for simplicity. Since the feature point is associated with propagation, the dynamic propagation model in Eq. (7) of Section IV-A. State Propagation in the manuscript is rewritten as:

$$\mathbf{X}_{I_{i+1}} = \boldsymbol{\Gamma}_i \boldsymbol{\Psi}(\mathbf{X}_{I_i}) \boldsymbol{\Upsilon}_i \quad (40)$$

where $\boldsymbol{\Gamma}_i$, $\boldsymbol{\Psi}(\mathbf{X}_{I_i})$, and $\boldsymbol{\Upsilon}_i$ are formulated as:

$$\begin{aligned} \boldsymbol{\Gamma}_i &= \left[\begin{array}{c|c} \mathbf{I}_{3 \times 3} & \frac{1}{2}\mathbf{g}\delta t^2 \quad \delta t \quad \mathbf{0}_{3 \times 1} \\ \hline \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{array} \right], \quad \boldsymbol{\Upsilon}_i = \left[\begin{array}{c|c} \Delta \mathbf{R}_i & \Delta \mathbf{p}_i \quad \Delta \mathbf{v}_i \quad \mathbf{0}_{3 \times 1} \\ \hline \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{array} \right] \\ \boldsymbol{\Psi}(\mathbf{X}_{I_i}) &= \left[\begin{array}{c|c} {}^L\mathbf{R}_{I_i} & {}^L\mathbf{p}_{I_i} + {}^L\mathbf{v}_{I_i} \delta t \quad {}^L\mathbf{v}_{I_i} \quad {}^G\mathbf{p}_f \\ \hline \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{array} \right]. \end{aligned} \quad (41)$$

Using Eq. (3) of Section III-A. Theoretical Background in the manuscript, (40) can be formulated as:

$$\begin{aligned} \text{Exp}(\xi_{I_{i+1}}) \widehat{\mathbf{X}}_{I_{i+1}} &= \boldsymbol{\Gamma}_i \boldsymbol{\Psi}(\text{Exp}(\xi_{I_i}) \widehat{\mathbf{X}}_{I_i}) \boldsymbol{\Upsilon}_i \\ &= \boldsymbol{\Gamma}_i \boldsymbol{\Psi}(\text{Exp}(\xi_{I_i})) \boldsymbol{\Psi}(\widehat{\mathbf{X}}_{I_i}) \boldsymbol{\Upsilon}_i \approx \boldsymbol{\Gamma}_i \text{Exp}(\mathbf{F} \xi_{I_i}) \boldsymbol{\Psi}(\widehat{\mathbf{X}}_{I_i}) \boldsymbol{\Upsilon}_i \end{aligned}$$

$$\begin{aligned} &= \text{Exp}(\mathbf{Ad}(\boldsymbol{\Gamma}_i) \mathbf{F} \xi_{I_i}) \boldsymbol{\Gamma}_i \boldsymbol{\Psi}(\widehat{\mathbf{X}}_{I_i}) \boldsymbol{\Upsilon}_i \\ &= \text{Exp}(\mathbf{Ad}(\boldsymbol{\Gamma}_i) \widehat{\mathbf{X}}_{I_{i+1}}) \end{aligned} \quad (42)$$

where the matrix \mathbf{F} is formulated as:

$$\mathbf{F} = \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \delta t \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{bmatrix}. \quad (43)$$

By expanding $\mathbf{Ad}(\boldsymbol{\Gamma}_i) \mathbf{F}$, we can obtain:

$$\mathbf{Ad}(\boldsymbol{\Gamma}_i) \mathbf{F} = \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \frac{1}{2}(\mathbf{g}) \times \delta t^2 & \mathbf{I}_{3 \times 3} & \mathbf{I}_{3 \times 3} \delta t & \mathbf{0}_{3 \times 3} \\ (\mathbf{g}) \times \delta t & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Phi}_{II} & \mathbf{0}_{9 \times 3} \\ \mathbf{0}_{3 \times 9} & \mathbf{I}_{3 \times 3} \end{bmatrix}. \quad (44)$$

To derive the relationship between $\xi_{I_{i+1}}$ and ζ_{I_i} , we first rewrite $\boldsymbol{\Upsilon}_i$ as:

$$\begin{aligned} \boldsymbol{\Upsilon}_i &= \widehat{\boldsymbol{\Upsilon}}_i \widetilde{\boldsymbol{\Upsilon}}_i \\ &= \left[\begin{array}{c|c} \Delta \widehat{\mathbf{R}}_i & \Delta \widehat{\mathbf{p}}_i \quad \Delta \widehat{\mathbf{v}}_i \quad \mathbf{0}_{3 \times 1} \\ \hline \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{array} \right] \left[\begin{array}{c|c} \Delta \widetilde{\mathbf{R}}_i & \Delta \widetilde{\mathbf{R}}_i^\top \quad \Delta \widetilde{\mathbf{p}}_i \quad \Delta \widetilde{\mathbf{R}}_i^\top \quad \Delta \widetilde{\mathbf{v}}_i \quad \mathbf{0}_{3 \times 1} \\ \hline \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{array} \right] \\ &= \widehat{\boldsymbol{\Upsilon}}_i \text{Exp} \left(\begin{bmatrix} \delta \Delta \boldsymbol{\theta}_i \\ \mathbf{J}_l^{-1}(\delta \Delta \boldsymbol{\theta}_i) \Delta \widehat{\mathbf{R}}_i^\top \Delta \widetilde{\mathbf{p}}_i \\ \mathbf{J}_l^{-1}(\delta \Delta \boldsymbol{\theta}_i) \Delta \widehat{\mathbf{R}}_i^\top \Delta \widetilde{\mathbf{v}}_i \\ \mathbf{0}_{3 \times 1} \end{bmatrix} \right) \\ &\approx \widehat{\boldsymbol{\Upsilon}}_i \text{Exp} \left(\begin{bmatrix} \delta \Delta \boldsymbol{\theta}_i \\ \Delta \widehat{\mathbf{R}}_i^\top \Delta \widetilde{\mathbf{p}}_i \\ \Delta \widehat{\mathbf{R}}_i^\top \Delta \widetilde{\mathbf{v}}_i \\ \mathbf{0}_{3 \times 1} \end{bmatrix} \right). \end{aligned} \quad (45)$$

Then similar to the procedure in (9), we can obtain:

$$\begin{aligned} \text{Exp}(\xi_{I_{i+1}}) \widehat{\mathbf{X}}_{i+1} &= \boldsymbol{\Gamma}_i \boldsymbol{\Psi}(\text{Exp}(\mathbf{X}_i) \widehat{\boldsymbol{\Upsilon}}_i \widetilde{\boldsymbol{\Upsilon}}_i) \\ \Rightarrow \text{Exp}(\xi_{I_{i+1}}) &= \widehat{\mathbf{X}}_{i+1} \widetilde{\boldsymbol{\Upsilon}}_i \widehat{\mathbf{X}}_{i+1}^{-1} \\ &= \text{Exp}(\mathbf{Ad}_{\widehat{\mathbf{X}}_{i+1}} \begin{bmatrix} \delta \Delta \boldsymbol{\theta}_i \\ \Delta \widehat{\mathbf{R}}_i^\top \Delta \widetilde{\mathbf{p}}_i \\ \Delta \widehat{\mathbf{R}}_i^\top \Delta \widetilde{\mathbf{v}}_i \\ \mathbf{0}_{3 \times 1} \end{bmatrix}) \\ &= \text{Exp} \left(\mathbf{Ad}_{\widehat{\mathbf{X}}_{i+1}} \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \Delta \widehat{\mathbf{R}}_i^\top & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \Delta \widehat{\mathbf{R}}_i^\top \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \end{bmatrix} \begin{bmatrix} \delta \Delta \boldsymbol{\theta}_i \\ \Delta \widetilde{\mathbf{p}}_i \\ \Delta \widetilde{\mathbf{v}}_i \end{bmatrix} \right). \end{aligned} \quad (46)$$

By expanding the expression in the exponential function, the following equation is obtained:

$$\begin{aligned} \mathbf{Ad}_{\widehat{\mathbf{X}}_{i+1}} & \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \Delta \widehat{\mathbf{R}}_i^\top & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \Delta \widehat{\mathbf{R}}_i^\top \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \end{bmatrix} \begin{bmatrix} \delta \Delta \boldsymbol{\theta}_i \\ \Delta \widetilde{\mathbf{p}}_i \\ \Delta \widetilde{\mathbf{v}}_i \end{bmatrix} \\ &= \mathbf{Ad}_{\widehat{\mathbf{X}}_{i+1}} \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \Delta \widehat{\mathbf{R}}_i^\top & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \Delta \widehat{\mathbf{R}}_i^\top \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \end{bmatrix} \begin{bmatrix} -\mathbf{J}_r(\Delta \widehat{\boldsymbol{\theta}}_i) \mathbf{0}_{3 \times 3} \\ \mathbf{E}_4 & -\mathbf{E}_2 \\ \mathbf{E}_3 & -\mathbf{E}_1 \end{bmatrix} \end{aligned}$$

$$\begin{aligned}
&= \begin{bmatrix} {}^L\widehat{\mathbf{R}}_{I_{i+1}} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ ({}^L\widehat{\mathbf{p}}_{I_{i+1}}) \times {}^L\widehat{\mathbf{R}}_{I_{i+1}} & {}^L\widehat{\mathbf{R}}_{I_{i+1}} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ ({}^L\widehat{\mathbf{v}}_{I_{i+1}}) \times {}^L\widehat{\mathbf{R}}_{I_{i+1}} & \mathbf{0}_{3 \times 3} & {}^L\widehat{\mathbf{R}}_{I_{i+1}} & \mathbf{0}_{3 \times 3} \\ ({}^L\widehat{\mathbf{p}}_{f_{i+1}}) \times {}^L\widehat{\mathbf{R}}_{I_{i+1}} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & {}^L\widehat{\mathbf{R}}_{I_{i+1}} \end{bmatrix} \\
&\quad \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \Delta\widehat{\mathbf{R}}_i^\top & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \Delta\widehat{\mathbf{R}}_i^\top \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \end{bmatrix} \begin{bmatrix} -\mathbf{J}_r(\Delta\widehat{\theta}_i) & \mathbf{0}_{3 \times 3} \\ \Xi_4 & -\Xi_2 \\ \Xi_3 & -\Xi_1 \end{bmatrix} \\
&= \begin{bmatrix} -{}^L\widehat{\mathbf{R}}_{I_{i+1}} \mathbf{J}_r(\Delta\widehat{\theta}_i) \delta t & \mathbf{0}_{3 \times 3} \\ -({}^L\widehat{\mathbf{p}}_{I_{i+1}}) \times {}^L\widehat{\mathbf{R}}_{I_{i+1}} \mathbf{J}_r(\Delta\widehat{\theta}_i) \delta t + \widehat{\mathbf{R}}_{I_i} \Xi_4 - \widehat{\mathbf{R}}_{I_i} \Xi_2 \\ -({}^L\widehat{\mathbf{v}}_{I_{i+1}}) \times {}^L\widehat{\mathbf{R}}_{I_{i+1}} \mathbf{J}_r(\Delta\widehat{\theta}_i) \delta t + \widehat{\mathbf{R}}_{I_i} \Xi_3 - \widehat{\mathbf{R}}_{I_i} \Xi_1 \\ -({}^L\widehat{\mathbf{p}}_{f_{i+1}}) \times {}^L\widehat{\mathbf{R}}_{I_{i+1}} \mathbf{J}_r(\Delta\widehat{\theta}_i) \delta t & \mathbf{0}_{3 \times 3} \end{bmatrix} \\
&= \begin{bmatrix} \Phi_{IB} \\ \Phi_{fB} \end{bmatrix}. \tag{47}
\end{aligned}$$

We can find that the propagation of feature point error state is associated with the biases. Organizing the above equations, the error state propagation function of MSC-InEKF-FC is formulated as:

$$\begin{aligned}
\xi_{i+1} &= \begin{bmatrix} \xi_{I_{i+1}} \\ \xi_{B_{i+1}} \\ \xi_{p_{Lf_{i+1}}} \end{bmatrix} = \Phi_i^{i+1} \begin{bmatrix} \xi_{I_i} \\ \xi_{B_i} \\ \xi_{p_{Lf_i}} \end{bmatrix} + \mathbf{G}_i \begin{bmatrix} \mathbf{n}_{dg} \\ \mathbf{n}_{da} \\ \mathbf{n}_{dbg} \\ \mathbf{n}_{dba} \end{bmatrix} \tag{48} \\
&= \begin{bmatrix} \Phi_{II} & \Phi_{IB} & \mathbf{0}_{9 \times 3} \\ \mathbf{0}_{6 \times 9} & \mathbf{I}_{6 \times 6} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 9} & \Phi_{fB} & \mathbf{I}_{3 \times 3} \end{bmatrix} \begin{bmatrix} \xi_{I_i} \\ \xi_{B_i} \\ \xi_{p_{Lf_i}} \end{bmatrix} + \begin{bmatrix} \mathbf{G}_{II} & \mathbf{0}_{9 \times 6} \\ \mathbf{0}_{6 \times 9} & \mathbf{I}_{6 \times 6} \delta t \\ \mathbf{G}_{fI} & \mathbf{0}_{3 \times 6} \end{bmatrix} \begin{bmatrix} \mathbf{n}_{dg} \\ \mathbf{n}_{da} \\ \mathbf{n}_{dbg} \\ \mathbf{n}_{dba} \end{bmatrix}.
\end{aligned}$$

To analyze the computation cost, we partition Φ_i^{i+1} and \mathbf{G}_i into several matrix blocks:

$$\Phi_i^{i+1} \triangleq \begin{bmatrix} \Phi_{\bar{f}\bar{f}} & \mathbf{0}_{15 \times 3} \\ \Phi_{f\bar{f}} & \mathbf{I}_{3 \times 3} \end{bmatrix} \quad \mathbf{G}_i \triangleq \begin{bmatrix} \mathbf{G}_{\bar{f}\bar{f}} \\ \mathbf{G}_{f\bar{f}} \end{bmatrix} \tag{49}$$

where $\Phi_{\bar{f}\bar{f}} \in \mathbb{R}^{15 \times 15}$ and $\mathbf{G}_{\bar{f}\bar{f}} \in \mathbb{R}^{15 \times 15}$ represent the blocks that are independent of the feature, and $\Phi_{f\bar{f}} \in \mathbb{R}^{3 \times 15}$, $\mathbf{G}_{f\bar{f}} \in \mathbb{R}^{3 \times 15}$ represent the blocks that are associated with the feature. Then the state covariance matrix is:

$$\begin{aligned}
\widehat{\mathbf{P}}_{i+1} &= \Phi_i^{i+1} \begin{bmatrix} \widehat{\mathbf{P}}_{\bar{f}\bar{f}} & \widehat{\mathbf{P}}_{f\bar{f}} \\ \widehat{\mathbf{P}}_{f\bar{f}} & \widehat{\mathbf{P}}_{ff} \end{bmatrix} \Phi_i^{i+1\top} + \mathbf{G}_i \mathbf{Q}_d \mathbf{G}_i^\top \\
&= \begin{bmatrix} \Phi_{\bar{f}\bar{f}} \widehat{\mathbf{P}}_{\bar{f}\bar{f}} \Phi_{\bar{f}\bar{f}}^\top & \Phi_{\bar{f}\bar{f}} \widehat{\mathbf{P}}_{f\bar{f}} \Phi_{f\bar{f}}^\top + \Phi_{f\bar{f}} \widehat{\mathbf{P}}_{\bar{f}\bar{f}} \\ \Phi_{f\bar{f}} \widehat{\mathbf{P}}_{\bar{f}\bar{f}} \Phi_{\bar{f}\bar{f}}^\top + \widehat{\mathbf{P}}_{f\bar{f}} \Phi_{\bar{f}\bar{f}}^\top & \Phi_{f\bar{f}} \widehat{\mathbf{P}}_{f\bar{f}} \Phi_{f\bar{f}}^\top + \Phi_{f\bar{f}} \widehat{\mathbf{P}}_{ff} \\ + \begin{bmatrix} \mathbf{G}_{\bar{f}\bar{f}} \mathbf{Q}_d \mathbf{G}_{\bar{f}\bar{f}}^\top & \mathbf{G}_{\bar{f}\bar{f}} \mathbf{Q}_d \mathbf{G}_{f\bar{f}}^\top \\ \mathbf{G}_{f\bar{f}} \mathbf{Q}_d \mathbf{G}_{\bar{f}\bar{f}}^\top & \mathbf{G}_{f\bar{f}} \mathbf{Q}_d \mathbf{G}_{f\bar{f}}^\top \end{bmatrix} & \end{bmatrix} \tag{50}
\end{aligned}$$

where $\widehat{\mathbf{P}}_{\bar{f}\bar{f}} \in \mathbb{R}^{(21+6c) \times (21+6c)}$, $\widehat{\mathbf{P}}_{f\bar{f}} \in \mathbb{R}^{(21+6c) \times 3}$, $\widehat{\mathbf{P}}_{ff} \in \mathbb{R}^{3 \times (21+6c)}$, $\widehat{\mathbf{P}}_{ff} \in \mathbb{R}^{3 \times 3}$ are the block components of covariance matrix $\widehat{\mathbf{P}}_i$. It can be seen that since $\Phi_{f\bar{f}}$ and $\mathbf{G}_{f\bar{f}}$ are not zero matrices, the calculation of $\Phi_{f\bar{f}} \widehat{\mathbf{P}}_{\bar{f}\bar{f}} \Phi_{\bar{f}\bar{f}}^\top$, $\widehat{\mathbf{P}}_{f\bar{f}} \Phi_{\bar{f}\bar{f}}^\top$, $\Phi_{f\bar{f}} \widehat{\mathbf{P}}_{\bar{f}\bar{f}} \Phi_{\bar{f}\bar{f}}^\top$, $\mathbf{G}_{\bar{f}\bar{f}} \mathbf{Q}_d \mathbf{G}_{\bar{f}\bar{f}}^\top$, $\mathbf{G}_{\bar{f}\bar{f}} \mathbf{Q}_d \mathbf{G}_{f\bar{f}}^\top$, and $\mathbf{G}_{f\bar{f}} \mathbf{Q}_d \mathbf{G}_{\bar{f}\bar{f}}^\top$ is required. If there are K features, computing these terms needs a computation of $O(K^2)$ flops.

C. State Update

Since only the feature points are associated with the IMU navigation state instead of relative transformation, the feature-based state update is different from (19) while the derivation of other state update equations keeps the same. According to the frame where the feature points are observed, there are two cases to be considered.

1) Feature points are observed by current clone pose. By replacing the state in error function Eq. (16) of Section IV-C. Feature Observation in the manuscript with (38), the linearized observation error function is derived as:

$$\begin{aligned}
\mathbf{z}_{f_{k_k}} &= \widetilde{\mathbf{y}}_{f_{k_k}} - \widehat{\mathbf{y}}_{f_{k_k}} \\
&= h \left[{}^C\mathbf{R}_I {}^L\mathbf{R}_{I_{k_k}}^\top ({}^L\mathbf{p}_{f_k} - {}^L\mathbf{p}_{I_{k_k}}) + {}^C\mathbf{p}_I \right] - \widehat{\mathbf{y}}_{f_{k_k}} + \mathbf{n}_{f_k} \\
&\approx \mathcal{H}_{k_k} \left[{}^C\mathbf{R}_I {}^L\mathbf{R}_{I_{k_k}}^\top ({}^L\mathbf{p}_{f_k} - {}^L\mathbf{p}_{I_{k_k}}) + {}^C\mathbf{p}_I \right. \\
&\quad \left. - {}^C\mathbf{R}_I {}^L\mathbf{R}_{I_{k_k}}^\top ({}^L\widehat{\mathbf{p}}_{f_k} - {}^L\widehat{\mathbf{p}}_{I_{k_k}}) - {}^C\mathbf{p}_I \right] + \mathbf{n}_{f_k} \\
&= \mathcal{H}_{k_k} \left[{}^C\mathbf{R}_I (\text{Exp}(\boldsymbol{\xi}_{R_{CP_{k_k}}}) {}^L\widehat{\mathbf{R}}_{I_{k_k}})^\top \left[\text{Exp}(\boldsymbol{\xi}_{R_{CP_{k_k}}}) {}^L\widehat{\mathbf{p}}_{f_k} + \right. \right. \\
&\quad \left. \left. \mathbf{J}_l(\boldsymbol{\xi}_{R_{CP_{k_k}}}) \boldsymbol{\xi}_{p_{Lf_k}} - \text{Exp}(\boldsymbol{\xi}_{R_{CP_{k_k}}}) {}^L\widehat{\mathbf{p}}_{I_{k_k}} - \mathbf{J}_l(\boldsymbol{\xi}_{R_{CP_{k_k}}}) \boldsymbol{\xi}_{p_{CP_{k_k}}} \right] \right. \\
&\quad \left. - {}^C\mathbf{R}_I {}^L\mathbf{R}_{I_{k_k}}^\top ({}^L\widehat{\mathbf{p}}_{f_k} - {}^L\widehat{\mathbf{p}}_{I_{k_k}}) \right] + \mathbf{n}_{f_k} \\
&\approx \mathcal{H}_{k_k} \left[{}^C\mathbf{R}_I {}^L\widehat{\mathbf{R}}_{I_{k_k}}^\top (\boldsymbol{\xi}_{p_{Lf_k}} - \boldsymbol{\xi}_{p_{CP_{k_k}}}) \right] + \mathbf{n}_{f_k}. \tag{51}
\end{aligned}$$

2) Feature points are observed by historical clone pose. By replacing the state in error function Eq. (16) of Section IV-C. Feature Observation in the manuscript with (38), the linearized observation error function is derived as:

$$\begin{aligned}
\mathbf{z}_{f_{j_k}} &= \widetilde{\mathbf{y}}_{f_{j_k}} - \widehat{\mathbf{y}}_{f_{j_k}} \\
&= h \left[{}^C\mathbf{R}_I {}^L\mathbf{R}_{I_{j_k}}^\top ({}^L\mathbf{p}_{f_k} - {}^L\mathbf{p}_{I_{j_k}}) + {}^C\mathbf{p}_I \right] - \widehat{\mathbf{y}}_{f_{j_k}} + \mathbf{n}_{f_k} \\
&\approx \mathcal{H}_{j_k} \left[{}^C\mathbf{R}_I {}^L\mathbf{R}_{I_{j_k}}^\top ({}^L\mathbf{p}_{f_k} - {}^L\mathbf{p}_{I_{j_k}}) + {}^C\mathbf{p}_I \right. \\
&\quad \left. - {}^C\mathbf{R}_I {}^L\mathbf{R}_{I_{j_k}}^\top ({}^L\widehat{\mathbf{p}}_{f_k} - {}^L\widehat{\mathbf{p}}_{I_{j_k}}) - {}^C\mathbf{p}_I \right] + \mathbf{n}_{f_k} \\
&= \mathcal{H}_{j_k} \left[{}^C\mathbf{R}_I (\text{Exp}(\boldsymbol{\xi}_{R_{CP_{j_k}}}) {}^L\widehat{\mathbf{R}}_{I_{j_k}})^\top \left[\text{Exp}(\boldsymbol{\xi}_{R_{CP_{j_k}}}) {}^L\widehat{\mathbf{p}}_{f_k} + \right. \right. \\
&\quad \left. \left. \mathbf{J}_l(\boldsymbol{\xi}_{R_{CP_{j_k}}}) \boldsymbol{\xi}_{p_{Lf_k}} - \text{Exp}(\boldsymbol{\xi}_{R_{CP_{j_k}}}) {}^L\widehat{\mathbf{p}}_{I_{j_k}} - \mathbf{J}_l(\boldsymbol{\xi}_{R_{CP_{j_k}}}) \boldsymbol{\xi}_{p_{CP_{j_k}}} \right] \right. \\
&\quad \left. - {}^C\mathbf{R}_I {}^L\mathbf{R}_{I_{j_k}}^\top ({}^L\widehat{\mathbf{p}}_{f_k} - {}^L\widehat{\mathbf{p}}_{I_{j_k}}) \right] + \mathbf{n}_{f_k} \\
&\approx \mathcal{H}_{j_k} \left[{}^C\mathbf{R}_I {}^L\widehat{\mathbf{R}}_{I_{j_k}}^\top (\boldsymbol{\xi}_{p_{Lf_k}} - \boldsymbol{\xi}_{p_{CP_{j_k}}}) \right] + \mathbf{n}_{f_k}. \tag{52}
\end{aligned}$$

D. Observability Analysis of MSCKF

Before deriving the observability analysis of MSC-InEKF-FC, we first introduce the observability of MSCKF. For simplicity, we do not include the biases in the state vector as they have been proved observable [3] for general motion. In addition, according to [3], both MSCKF and EKF are derived based on the same propagation model and the linearized observation error models with only the observation Jacobians are calculated using different estimates. Therefore, it is equivalent to analyze the observability of EKF, as long as we adjust the linearization points. The final considered state for MSCKF is given by $\mathbf{X}_k = (\mathbf{X}_{I_k}, \mathbf{X}_{G_k}, {}^L\mathbf{p}_{f_k})$. The observability matrix is defined as:

$$\mathcal{O} = [\mathcal{O}_0^\top \ \mathcal{O}_1^\top \ \dots \ \mathcal{O}_k^\top]^\top = \left[\mathbf{H}_0^\top \ \Phi_0^{1\top} \ \mathbf{H}_1^\top \ \dots \ \Phi_0^{k\top} \ \mathbf{H}_k^\top \right]^\top \quad (53)$$

where $\mathbf{H}_k = [\mathbf{H}_{o_k}^\top \ \mathbf{H}_{f_k}^\top \ \mathbf{H}_{s_k}^\top \ \mathbf{H}_{p_k}^\top]^\top \in \mathbb{R}^{9 \times 18}$. For MSCKF, the state transition matrix Φ_{k-1}^k is formulated as:

$$\Phi_{k-1}^k = \begin{bmatrix} \Delta \hat{\mathbf{R}}_{k-1}^k & \mathbf{0}_{3 \times 3} \\ -(\Delta \hat{\mathbf{p}}_{k-1}^k) \times {}^L\hat{\mathbf{R}}_{I_{k-1}} & \mathbf{I}_{3 \times 3} & \mathbf{I}_{3 \times 3} \Delta t & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ -(\Delta \hat{\mathbf{v}}_{k-1}^k) \times {}^L\hat{\mathbf{R}}_{I_{k-1}} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{bmatrix} \quad (54)$$

where $\Delta \hat{\mathbf{R}}_{k-1}^k = {}^L\hat{\mathbf{R}}_{I_k}^\top {}^L\hat{\mathbf{R}}_{I_{k-1}}$, $\Delta \hat{\mathbf{p}}_{k-1}^k = {}^L\hat{\mathbf{p}}_{I_k} - {}^L\hat{\mathbf{p}}_{I_{k-1}} - {}^L\hat{\mathbf{v}}_{I_{k-1}} \Delta t_k + \frac{1}{2}\mathbf{g}\Delta t_k^2$, and $\Delta \hat{\mathbf{v}}_{k-1}^k = {}^L\hat{\mathbf{v}}_{I_k} - {}^L\hat{\mathbf{v}}_{I_{k-1}} + \mathbf{g}\Delta t_k$. The transition matrix Φ_0^k from time t_0 to t_k is derived as:

$$\Phi_0^k = \begin{bmatrix} \Delta \hat{\mathbf{R}}_0^k & \mathbf{0}_{3 \times 3} \\ -(\Delta \hat{\mathbf{p}}_0^k) \times {}^L\hat{\mathbf{R}}_{I_0} & \mathbf{I}_{3 \times 3} & \mathbf{I}_{3 \times 3} \Delta t_k & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ -(\Delta \hat{\mathbf{v}}_0^k) \times {}^L\hat{\mathbf{R}}_{I_0} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{bmatrix} \quad (55)$$

where $\Delta \hat{\mathbf{R}}_0^k = {}^L\hat{\mathbf{R}}_{I_k}^\top {}^L\hat{\mathbf{R}}_{I_0}$, $\Delta \hat{\mathbf{p}}_0^k = {}^L\hat{\mathbf{p}}_{I_k} - {}^L\hat{\mathbf{p}}_{I_0} - {}^L\hat{\mathbf{v}}_{I_0} \Delta t_k + \frac{1}{2}\mathbf{g}\Delta t_k^2$, $\Delta \hat{\mathbf{v}}_0^k = {}^L\hat{\mathbf{v}}_{I_k} - {}^L\hat{\mathbf{v}}_{I_0} + \mathbf{g}\Delta t_k$, and $\Delta t_k = k\Delta t$. The odometry observation Jacobian matrix \mathbf{H}_{o_k} is derived as:

$$\mathbf{H}_{o_k} = \left[\begin{array}{ccccc} \frac{\partial \mathbf{z}_{o_k}}{\partial \xi_{R_{L I_k}}} & \frac{\partial \mathbf{z}_{o_k}}{\partial \xi_{P_{L I_k}}} & \frac{\partial \mathbf{z}_{o_k}}{\partial \xi_{T_{L G_k}}} & \frac{\partial \mathbf{z}_{o_k}}{\partial \xi_{P_{L f_k}}} & \frac{\partial \mathbf{z}_{o_k}}{\partial \xi_{T_{L f_k}}} \end{array} \right] = {}^O\mathbf{R}_I \left[\begin{array}{ccccc} ({}^L\hat{\mathbf{R}}_{I_k}^\top {}^L\hat{\mathbf{v}}_{I_k}) \times & \mathbf{0}_{3 \times 3} & {}^L\hat{\mathbf{R}}_{I_k} & \mathbf{0}_{3 \times 6} & \mathbf{0}_{3 \times 3} \end{array} \right]. \quad (56)$$

The feature point observation Jacobian matrix \mathbf{H}_{f_k} is:

$$\begin{aligned} \mathbf{H}_{f_k} &= \left[\begin{array}{ccc} \frac{\partial \mathbf{z}_{f_k}}{\partial \xi_{I_k}} & \frac{\partial \mathbf{z}_{f_k}}{\partial \xi_{T_{L G_k}}} & \frac{\partial \mathbf{z}_{f_k}}{\partial \xi_{P_{L f_k}}} \end{array} \right] = \left[\mathbf{H}_{f_k}^{\mathbf{X}_I} \ \mathbf{0}_{2 \times 6} \ \mathbf{H}_{f_k}^{\mathbf{p}_{L f}} \right] \\ \mathbf{H}_{f_k}^{\mathbf{X}_I} &= \mathcal{H}_k^C \mathbf{R}_I \left[({}^I\hat{\mathbf{p}}_{f_k}) \times -({}^L\hat{\mathbf{R}}_{I_k}) \times \right] \\ \mathbf{H}_{f_k}^{\mathbf{p}_{L f}} &= \mathcal{H}_{j_k}^C \mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_{j_k}} ({}^L\hat{\mathbf{R}}_{I_k}) \times. \end{aligned} \quad (57)$$

The Jacobian matrix of streetlight matches \mathbf{H}_{s_k} is derived as:

$$\begin{aligned} \mathbf{H}_{s_k} &= \left[\begin{array}{ccc} \frac{\partial \mathbf{z}_{s_k}}{\partial \xi_{I_k}} & \frac{\partial \mathbf{z}_{s_k}}{\partial \xi_{T_{L G_k}}} & \frac{\partial \mathbf{z}_{s_k}}{\partial \xi_{P_{L f_k}}} \end{array} \right] = \left[\mathbf{H}_{s_k}^{\mathbf{X}_I} \ \mathbf{H}_{s_k}^{\mathbf{X}_G} \ \mathbf{0}_{2 \times 3} \right] \\ \mathbf{H}_{s_k}^{\mathbf{X}_I} &= \mathcal{H}_k^C \mathbf{R}_I \left[({}^I\hat{\mathbf{c}}_{s_k}) \times -{}^L\hat{\mathbf{R}}_{I_k} \ \mathbf{0}_{3 \times 3} \right] \\ \mathbf{H}_{s_k}^{\mathbf{X}_G} &= \left[{}^L\hat{\mathbf{R}}_{I_k}^\top {}^L\hat{\mathbf{R}}_{G_k} ({}^G\tilde{\mathbf{c}}_s) \times \ {}^L\hat{\mathbf{R}}_{I_k} \right] \end{aligned} \quad (58)$$

where ${}^I\hat{\mathbf{c}}_{s_k} = {}^L\hat{\mathbf{R}}_{I_k}^\top ({}^L\hat{\mathbf{R}}_{G_k} {}^G\tilde{\mathbf{c}}_s + {}^L\hat{\mathbf{p}}_{G_k} - {}^L\hat{\mathbf{p}}_{I_k})$ is the virtual center of streetlight cluster represented in the body frame. The Jacobian matrix of prior poses \mathbf{H}_{p_k} is formulated as:

$$\begin{aligned} \mathbf{H}_{p_k} &= \left[\begin{array}{ccc} \frac{\partial \mathbf{z}_{p_k}}{\partial \xi_{I_k}} & \frac{\partial \mathbf{z}_{p_k}}{\partial \xi_{T_{L G_k}}} & \frac{\partial \mathbf{z}_{p_k}}{\partial \xi_{P_{L f_k}}} \end{array} \right] = \left[\mathbf{H}_{p_k}^{\mathbf{X}_I} \ \mathbf{H}_{p_k}^{\mathbf{X}_G} \ \mathbf{0}_{2 \times 3} \right] \\ \mathbf{H}_{p_k}^{\mathbf{X}_I} &= \mathbf{e}_3^\top {}^O\mathbf{R}_I {}^G\tilde{\mathbf{R}}_{I'}^\top \begin{bmatrix} \mathbf{0}_{3 \times 3} & -{}^L\hat{\mathbf{R}}_{G_k}^\top \mathbf{0}_{3 \times 3} \\ -{}^L\hat{\mathbf{R}}_{G_k}^\top {}^L\hat{\mathbf{R}}_{I_k} ({}^O\mathbf{R}_I^\top \mathbf{e}_3) \times & \mathbf{0}_{3 \times 3} \end{bmatrix} \\ \mathbf{H}_{p_k}^{\mathbf{X}_G} &= \mathbf{e}_3^\top {}^O\mathbf{R}_I {}^G\tilde{\mathbf{R}}_{I'}^\top \begin{bmatrix} {}^L\hat{\mathbf{R}}_{G_k}^\top ({}^L\hat{\mathbf{p}}_{I_k} - {}^L\hat{\mathbf{p}}_{G_k})) \times & -{}^L\hat{\mathbf{R}}_{G_k}^\top \\ ({}^L\hat{\mathbf{R}}_{G_k}^\top {}^L\hat{\mathbf{R}}_{I_k} {}^O\mathbf{R}_I^\top \mathbf{e}_3) \times & \mathbf{0}_{3 \times 3} \end{bmatrix}. \end{aligned} \quad (59)$$

The k -th block row of \mathcal{O} is formulated as:

$$\mathcal{O}_k = \mathbf{H}_k \Phi_0^k = \begin{bmatrix} \mathbf{O}_{11} & \mathbf{0}_{3 \times 3} & \mathbf{O}_{13} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{O}_{21} & \mathbf{O}_{22} & \mathbf{O}_{23} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{O}_{26} \\ \mathbf{O}_{31} & \mathbf{O}_{32} & \mathbf{O}_{33} & \mathbf{O}_{34} & \mathbf{O}_{35} & \mathbf{0}_{3 \times 3} \\ \mathbf{O}_{41} & \mathbf{O}_{42} & \mathbf{O}_{43} & \mathbf{O}_{44} & \mathbf{O}_{45} & \mathbf{0}_{3 \times 3} \\ \mathbf{O}_{51} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{O}_{54} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \end{bmatrix} \quad (60)$$

where the block matrices are respectively formulated as:

$$\begin{aligned} \mathbf{O}_{11} &= {}^O\mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top ({}^L\hat{\mathbf{v}}_{I_k} - \Delta \hat{\mathbf{v}}_0^k) \times {}^L\hat{\mathbf{R}}_{I_0} \\ \mathbf{O}_{13} &= {}^O\mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top \Delta t_k \\ \mathbf{O}_{21} &= \mathcal{H}_k^C \mathbf{R}_I (({}^I\hat{\mathbf{p}}_{f_k}) \times \Delta \hat{\mathbf{R}}_0^k + {}^L\hat{\mathbf{R}}_{I_k}^\top (\Delta \hat{\mathbf{p}}_0^k) \times {}^L\hat{\mathbf{R}}_{I_0}) \\ \mathbf{O}_{22} &= -\mathcal{H}_k^C \mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top \\ \mathbf{O}_{23} &= -\mathcal{H}_k^C \mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top \Delta t_k \\ \mathbf{O}_{26} &= \mathcal{H}_k^C \mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top \\ \mathbf{O}_{31} &= \mathcal{H}_k^C \mathbf{R}_I (({}^I\hat{\mathbf{c}}_{s_k}) \times \Delta \hat{\mathbf{R}}_0^k + {}^L\hat{\mathbf{R}}_{I_k}^\top (\Delta \hat{\mathbf{p}}_0^k) \times {}^L\hat{\mathbf{R}}_{I_0}) \\ \mathbf{O}_{32} &= -\mathcal{H}_k^C \mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top \\ \mathbf{O}_{33} &= -\mathcal{H}_k^C \mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top \Delta t_k \\ \mathbf{O}_{34} &= -\mathcal{H}_k^C \mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top {}^L\hat{\mathbf{R}}_{G_k}^\top ({}^G\tilde{\mathbf{c}}_s) \times \\ \mathbf{O}_{35} &= \mathcal{H}_k^C \mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top \\ \mathbf{O}_{41} &= \mathbf{e}_3^\top {}^O\mathbf{R}_I {}^G\tilde{\mathbf{R}}_{I'}^\top {}^L\hat{\mathbf{R}}_{G_k} (\Delta \hat{\mathbf{p}}_0^k) \times {}^L\hat{\mathbf{R}}_{I_0} \\ \mathbf{O}_{42} &= -\mathbf{e}_3^\top {}^O\mathbf{R}_I {}^G\tilde{\mathbf{R}}_{I'}^\top {}^L\hat{\mathbf{R}}_{G_k} \\ \mathbf{O}_{43} &= -\mathbf{e}_3^\top {}^O\mathbf{R}_I {}^G\tilde{\mathbf{R}}_{I'}^\top {}^L\hat{\mathbf{R}}_{G_k} \Delta t_k \\ \mathbf{O}_{44} &= -\mathbf{e}_3^\top {}^O\mathbf{R}_I {}^G\tilde{\mathbf{R}}_{I'}^\top ({}^L\hat{\mathbf{R}}_{G_k}^\top ({}^L\hat{\mathbf{p}}_{I_k} - {}^L\hat{\mathbf{p}}_{G_k})) \times \\ \mathbf{O}_{45} &= \mathbf{e}_3^\top {}^O\mathbf{R}_I {}^G\tilde{\mathbf{R}}_{I'}^\top {}^L\hat{\mathbf{R}}_{G_k} \\ \mathbf{O}_{51} &= -\mathbf{e}_3^\top {}^O\mathbf{R}_I {}^G\tilde{\mathbf{R}}_{I'}^\top {}^L\hat{\mathbf{R}}_{G_k}^\top {}^L\hat{\mathbf{R}}_{I_k} ({}^O\mathbf{R}_I^\top \mathbf{e}_3) \times \\ \mathbf{O}_{54} &= \mathbf{e}_3^\top {}^O\mathbf{R}_I {}^G\tilde{\mathbf{R}}_{I'}^\top ({}^L\hat{\mathbf{R}}_{G_k}^\top {}^L\hat{\mathbf{R}}_{I_k} {}^O\mathbf{R}_I^\top \mathbf{e}_3) \times. \end{aligned} \quad (61)$$

Then the right nullspace of \mathcal{O}_k spans as:

$$\mathcal{N}_k = \text{span} \left(\begin{bmatrix} \mathbf{0}_{3 \times 3} & {}^L\hat{\mathbf{R}}_{I_0}^\top \mathbf{g} \\ \mathbf{I}_{3 \times 3} & -({}^L\hat{\mathbf{p}}_{I_0}) \times \mathbf{g} \\ \mathbf{0}_{3 \times 3} & -({}^L\hat{\mathbf{v}}_{I_0}) \times \mathbf{g} \\ \mathbf{0}_{3 \times 3} & {}^L\hat{\mathbf{R}}_{G_k}^\top \mathbf{g} \\ \mathbf{I}_{3 \times 3} & -({}^L\hat{\mathbf{p}}_{G_k}) \times \mathbf{g} \\ \mathbf{I}_{3 \times 3} & -({}^L\hat{\mathbf{p}}_{f_k}) \times \mathbf{g} \end{bmatrix} \right). \quad (62)$$

Ideally, if all Jacobians are calculated using the true states [3], \mathcal{N}_k is constant regardless of the value of k . Hence the observability matrix has a null space of dimension four, indicating the system's observability is not disrupted after

linearization. However, the Jacobians are always evaluated using current state estimates. Consequently, \mathcal{N}_k varies with k due to the different values of ${}^L\hat{\mathbf{R}}_{G_k}$, ${}^L\hat{\mathbf{p}}_{G_k}$, ${}^L\hat{\mathbf{p}}_{f_k}$, leading to the null space of \mathcal{O} collapsing to three dimensions. In other words, one dimension of the unobservable space becomes observable, leading to the inconsistency problem.

E. Observability Analysis of MSC-InEKF-FC

To take the k -th element of \mathcal{O} for analysis, Φ_1^k is formulated as:

$$\begin{aligned}\Phi_1^k &= \left(\begin{bmatrix} \Phi_{II} & \mathbf{0}_{9 \times 9} \\ \mathbf{0}_{9 \times 9} & \mathbf{I}_{9 \times 9} \end{bmatrix} \right)^k \quad (63) \\ &= \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \frac{1}{2}(\mathbf{g}) \times \Delta t_k^2 & \mathbf{I}_{3 \times 3} & \mathbf{I}_{3 \times 3} \Delta t_k & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ (\mathbf{g}) \times \Delta t_k & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{bmatrix}\end{aligned}$$

where $\Delta t_k = k\Delta t$ and Δt is the time interval between two consecutive time steps. We then obtain:

$$\begin{aligned}\mathbf{H}_k \Phi_1^k &= \begin{bmatrix} \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{H}_{13} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{2 \times 3} & \mathbf{H}_{22} & \mathbf{0}_{2 \times 3} & \mathbf{0}_{2 \times 3} & \mathbf{0}_{2 \times 3} & \mathbf{H}_{26} \\ \mathbf{H}_{31} & \mathbf{H}_{32} & \mathbf{0}_{2 \times 3} & \mathbf{H}_{34} & \mathbf{H}_{35} & \mathbf{0}_{2 \times 3} \\ \mathbf{H}_{41} & \mathbf{H}_{42} & \mathbf{0}_{2 \times 3} & \mathbf{H}_{44} & \mathbf{H}_{45} & \mathbf{0}_{2 \times 3} \end{bmatrix} \\ &\quad \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \frac{1}{2}(\mathbf{g}) \times \Delta t_k^2 & \mathbf{I}_{3 \times 3} & \mathbf{I}_{3 \times 3} \Delta t_k & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ (\mathbf{g}) \times \Delta t_k & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{M}_{11} & \mathbf{0}_{3 \times 3} & \mathbf{M}_{13} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{M}_{21} & \mathbf{M}_{22} & \mathbf{M}_{23} & \mathbf{0}_{2 \times 3} & \mathbf{0}_{2 \times 3} & \mathbf{M}_{26} \\ \mathbf{M}_{31} & \mathbf{M}_{32} & \mathbf{M}_{33} & \mathbf{M}_{34} & \mathbf{M}_{35} & \mathbf{0}_{2 \times 3} \\ \mathbf{M}_{41} & \mathbf{M}_{42} & \mathbf{M}_{43} & \mathbf{M}_{44} & \mathbf{M}_{45} & \mathbf{0}_{2 \times 3} \end{bmatrix} \quad (64)\end{aligned}$$

where the matrix elements are respectively formulated as:

$$\begin{aligned}\mathbf{H}_{13} &= {}^O\mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top \\ \mathbf{H}_{22} &= -\mathbf{H}_{26} = -\mathbf{\mathcal{H}}_k {}^C\mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top \\ \mathbf{H}_{31} &= -\mathbf{H}_{34} = \mathbf{\mathcal{H}}_k {}^C\mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top ({}^L\hat{\mathbf{c}}_s)_\times \\ \mathbf{H}_{32} &= -\mathbf{H}_{35} = -\mathbf{\mathcal{H}}_k {}^C\mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top \\ \mathbf{H}_{41} &= -\mathbf{H}_{44} = \begin{bmatrix} -\mathbf{e}_3^\top {}^I\mathbf{R}_O^\top {}^G\tilde{\mathbf{R}}_{I'}^\top {}^L\hat{\mathbf{R}}_{G_k}({}^L\hat{\mathbf{p}}_{I_k})_\times \\ -\mathbf{e}_3^\top {}^I\mathbf{R}_O^\top {}^G\tilde{\mathbf{R}}_{I'}^\top {}^L\hat{\mathbf{R}}_{G_k}^\top ({}^L\hat{\mathbf{R}}_{I_k} {}^O\mathbf{R}_I^\top \mathbf{e}_3)_\times \end{bmatrix} \\ \mathbf{H}_{42} &= -\mathbf{H}_{45} = \begin{bmatrix} \mathbf{e}_3^\top {}^I\mathbf{R}_O^\top {}^G\tilde{\mathbf{R}}_{I'}^\top {}^L\hat{\mathbf{R}}_{G_k} \\ \mathbf{0}_{3 \times 3} \end{bmatrix} \\ \mathbf{M}_{11} &= {}^O\mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top (\mathbf{g})_\times \Delta t_k \quad \mathbf{M}_{13} = {}^O\mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top \\ \mathbf{M}_{21} &= -\frac{1}{2}\mathbf{\mathcal{H}}_k {}^C\mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top (\mathbf{g})_\times \Delta t_k^2 \quad \mathbf{M}_{22} = -\mathbf{\mathcal{H}}_k {}^C\mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top \\ \mathbf{M}_{23} &= -\mathbf{\mathcal{H}}_k {}^C\mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top \Delta t_k \quad \mathbf{M}_{26} = \mathbf{\mathcal{H}}_k {}^C\mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top \\ \mathbf{M}_{31} &= \mathbf{\mathcal{H}}_k {}^C\mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top ({}^L\hat{\mathbf{c}}_s)_\times - \frac{1}{2}\mathbf{\mathcal{H}}_k {}^C\mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top (\mathbf{g})_\times \Delta t_k^2 \\ \mathbf{M}_{32} &= -\mathbf{\mathcal{H}}_k {}^C\mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top \quad \mathbf{M}_{33} = -\mathbf{\mathcal{H}}_k {}^C\mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top \Delta t_k \\ \mathbf{M}_{34} &= -\mathbf{\mathcal{H}}_k {}^C\mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top ({}^L\hat{\mathbf{c}}_s)_\times \quad \mathbf{M}_{35} = \mathbf{\mathcal{H}}_k {}^C\mathbf{R}_I {}^L\hat{\mathbf{R}}_{I_k}^\top\end{aligned}$$

$$\begin{aligned}\mathbf{M}_{41} &= \mathbf{e}_3^\top {}^I\mathbf{R}_O^\top {}^G\tilde{\mathbf{R}}_{I'}^\top \begin{bmatrix} -{}^L\hat{\mathbf{R}}_{G_k}({}^L\hat{\mathbf{p}}_{I_k})_\times + \frac{1}{2}{}^L\hat{\mathbf{R}}_{G_k}(\mathbf{g})_\times \Delta t_k^2 \\ -{}^L\hat{\mathbf{R}}_{G_k}^\top ({}^L\hat{\mathbf{R}}_{I_k} {}^O\mathbf{R}_I^\top \mathbf{e}_3)_\times \end{bmatrix} \\ \mathbf{M}_{42} &= \begin{bmatrix} \mathbf{e}_3^\top {}^I\mathbf{R}_O^\top {}^G\tilde{\mathbf{R}}_{I'}^\top {}^L\hat{\mathbf{R}}_{G_k} \\ \mathbf{0}_{3 \times 3} \end{bmatrix} \quad \mathbf{M}_{43} = \begin{bmatrix} \mathbf{e}_3^\top {}^I\mathbf{R}_O^\top {}^G\tilde{\mathbf{R}}_{I'}^\top {}^L\hat{\mathbf{R}}_{G_k} \Delta t_k \\ \mathbf{0}_{3 \times 3} \end{bmatrix} \\ \mathbf{M}_{44} &= \mathbf{e}_3^\top {}^I\mathbf{R}_O^\top {}^G\tilde{\mathbf{R}}_{I'}^\top \begin{bmatrix} {}^L\hat{\mathbf{R}}_{G_k}({}^L\hat{\mathbf{p}}_{I_k})_\times \\ {}^L\hat{\mathbf{R}}_{G_k}^\top ({}^L\hat{\mathbf{R}}_{I_k} {}^O\mathbf{R}_I^\top \mathbf{e}_3)_\times \end{bmatrix} \\ \mathbf{M}_{45} &= \begin{bmatrix} -\mathbf{e}_3^\top {}^I\mathbf{R}_O^\top {}^G\tilde{\mathbf{R}}_{I'}^\top {}^L\hat{\mathbf{R}}_{G_k} \\ \mathbf{0}_{3 \times 3} \end{bmatrix}. \quad (65)\end{aligned}$$

The right nullspace \mathcal{N}_{FC} of the observability matrix can be deduced by spanning the following unobservable subspace:

$$\mathcal{N}_{FC} = \text{span} \left(\begin{bmatrix} \mathbf{g}^\top & \mathbf{0}_{1 \times 3} & \mathbf{0}_{1 \times 3} & \mathbf{g}^\top & \mathbf{0}_{1 \times 3} & \mathbf{0}_{1 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{bmatrix} \right)^\top. \quad (66)$$

We can see that MSC-InEKF-FC maintains the correct dimension of the nullspace without introducing spurious gain.

APPENDIX E MSC-INEKF-FDN

A. State Representation

For MSC-InEKF-FDN, the features are not maintained in any Lie group. The system state is defined as:

$$\begin{aligned}\mathbf{X}_k &= (\mathbf{X}_{I_k}, \mathbf{B}_k, \mathbf{X}_{C_k}, \mathbf{X}_{G_k}, {}^L\mathbf{p}_{f_k}) \\ \mathbf{X}_{I_k} &= \begin{bmatrix} {}^L\mathbf{R}_{I_k} & {}^L\mathbf{p}_{I_k} & {}^L\mathbf{v}_{I_k} \\ \mathbf{0}_{2 \times 3} & & \mathbf{I}_{2 \times 2} \end{bmatrix} \\ \mathbf{B}_k &= [\mathbf{b}_{g_k}^\top \quad \mathbf{b}_{a_k}^\top]^\top \\ \mathbf{X}_{C_k} &= (\mathbf{X}_{CP_{k_k}}, \dots, \mathbf{X}_{CP_{k-c+1_k}}), \quad \mathbf{X}_{CP_{j_k}} = \begin{bmatrix} {}^L\mathbf{R}_{I_{j_k}} & {}^L\mathbf{p}_{I_{j_k}} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} \\ \mathbf{X}_{G_k} &= \begin{bmatrix} {}^L\mathbf{R}_{G_k} & {}^L\mathbf{p}_{G_k} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix}. \quad (67)\end{aligned}$$

We keep the error states in the same formula with that in MSC-InEKF-FDRC, i.e., $\xi_k = [\xi_{I_k}^\top \quad \xi_{B_k}^\top \quad \xi_{C_k}^\top \quad \xi_{T_{LG_k}}^\top \quad \xi_{p_{Lf_k}}^\top]^\top$. The feature is decoupled from the Lie group and we have:

$${}^L\mathbf{p}_{f_k} = \hat{\mathbf{p}}_{f_k} + \xi_{p_{Lf_k}}. \quad (68)$$

B. State Propagation

The derivation of state propagation is same as that in Section A.

C. State Update

Since only the feature points are not maintained in Lie group, the feature-based state update is different from (19) while the derivation of other state update equations keeps the same. By replacing the state in error function **Eq. (16)** of **Section IV-C. Feature Observation** in the manuscript with (68), the linearized observation error function is derived as:

$$\begin{aligned}\mathbf{z}_{f_k} &= \tilde{\mathbf{y}}_{f_k} - \hat{\mathbf{y}}_{f_k} \\ &= h \left[{}^C\mathbf{R}_I {}^L\mathbf{R}_{I_k}^\top ({}^L\mathbf{p}_{f_k} - {}^L\mathbf{p}_{I_k}) + {}^C\mathbf{p}_I \right] - \hat{\mathbf{y}}_{f_k} + \mathbf{n}_{f_k} \\ &\approx \mathbf{H}_{j_k} \left[{}^C\mathbf{R}_I {}^L\mathbf{R}_{I_k}^\top ({}^L\mathbf{p}_{f_k} - {}^L\mathbf{p}_{I_k}) + {}^C\mathbf{p}_I \right]\end{aligned}$$

$$\begin{aligned}
& -{}^C \mathbf{R}_I {}^L \widehat{\mathbf{R}}_{I_{jk}}^\top ({}^L \widehat{\mathbf{p}}_{f_k} - {}^L \widehat{\mathbf{p}}_{I_{jk}}) - {}^C \mathbf{p}_I \Big] + \mathbf{n}_{f_k} \\
= & \mathcal{H}_{jk} \left[{}^C \mathbf{R}_I (\text{Exp}(\boldsymbol{\xi}_{\mathbf{R}_{LI_k}}) {}^L \widehat{\mathbf{R}}_{I_{jk}})^\top \left[{}^L \widehat{\mathbf{p}}_{f_k} + \boldsymbol{\xi}_{\mathbf{p}_{LI_k}} \right. \right. \\
& \left. \left. - \text{Exp}(\boldsymbol{\xi}_{\mathbf{R}_{LI_k}}) {}^L \widehat{\mathbf{p}}_{I_{jk}} - \mathbf{J}_l(\boldsymbol{\xi}_{\mathbf{R}_{LI_k}}) \boldsymbol{\xi}_{\mathbf{p}_{LI_k}} \right] - {}^C \mathbf{R}_I {}^L \widehat{\mathbf{R}}_{I_{jk}}^\top \right. \\
& \left. ({}^L \widehat{\mathbf{p}}_{f_k} - {}^L \widehat{\mathbf{p}}_{I_{jk}}) \right] + \mathbf{n}_{f_k} \\
\approx & \mathcal{H}_{jk} \left[{}^C \mathbf{R}_I {}^L \widehat{\mathbf{R}}_{I_{jk}}^\top (\mathbf{I}_{3 \times 3} - (\boldsymbol{\xi}_{\mathbf{R}_{LI_k}}) \times) \left[{}^L \widehat{\mathbf{p}}_{f_k} + \boldsymbol{\xi}_{\mathbf{p}_{Lf_k}} \right. \right. \\
& \left. \left. - (\mathbf{I}_{3 \times 3} + (\boldsymbol{\xi}_{\mathbf{R}_{LI_k}}) \times) {}^L \widehat{\mathbf{p}}_{I_{jk}} - \boldsymbol{\xi}_{\mathbf{p}_{LI_k}} \right] - {}^C \mathbf{R}_I {}^L \widehat{\mathbf{R}}_{I_{jk}}^\top \right. \\
& \left. ({}^L \widehat{\mathbf{p}}_{f_k} - {}^L \widehat{\mathbf{p}}_{I_{jk}}) \right] + \mathbf{n}_{f_k} \\
\approx & \mathcal{H}_{jk} {}^C \mathbf{R}_I {}^L \widehat{\mathbf{R}}_{I_{jk}}^\top \left[({}^L \widehat{\mathbf{p}}_{f_k}) \times \boldsymbol{\xi}_{\mathbf{R}_{LI_k}} - \boldsymbol{\xi}_{\mathbf{p}_{LI_k}} + \boldsymbol{\xi}_{\mathbf{p}_{Lf_k}} \right] + \mathbf{n}_{f_k}. \tag{69}
\end{aligned}$$

D. Observability Analysis

In MSC-InEKF-FDN, the formulation of Φ_1^k is same as (63). Therefore, the k -th element of the observability matrix is:

$$\begin{aligned}
\mathbf{H}_k \Phi_1^k = & \begin{bmatrix} \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{H}_{13} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{H}_{21} & \mathbf{H}_{22} & \mathbf{0}_{2 \times 3} & \mathbf{0}_{2 \times 3} & \mathbf{0}_{2 \times 3} & \mathbf{H}_{26} \\ \mathbf{H}_{31} & \mathbf{H}_{32} & \mathbf{0}_{2 \times 3} & \mathbf{H}_{34} & \mathbf{H}_{35} & \mathbf{0}_{2 \times 3} \\ \mathbf{H}_{41} & \mathbf{H}_{42} & \mathbf{0}_{2 \times 3} & \mathbf{H}_{44} & \mathbf{H}_{45} & \mathbf{0}_{2 \times 3} \end{bmatrix} \\
& \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \frac{1}{2}(\mathbf{g}) \times \Delta t_k^2 & \mathbf{I}_{3 \times 3} & \mathbf{I}_{3 \times 3} \Delta t_k & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ (\mathbf{g}) \times \Delta t_k & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{bmatrix} \\
= & \begin{bmatrix} \mathbf{M}_{11} & \mathbf{0}_{3 \times 3} & \mathbf{M}_{13} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{M}_{21} & \mathbf{M}_{22} & \mathbf{M}_{23} & \mathbf{0}_{2 \times 3} & \mathbf{M}_{26} & \mathbf{0}_{2 \times 3} \\ \mathbf{M}_{31} & \mathbf{M}_{32} & \mathbf{M}_{33} & \mathbf{M}_{34} & \mathbf{M}_{35} & \mathbf{0}_{2 \times 3} \\ \mathbf{M}_{41} & \mathbf{M}_{42} & \mathbf{M}_{43} & \mathbf{M}_{44} & \mathbf{M}_{45} & \mathbf{0}_{2 \times 3} \end{bmatrix} \tag{70}
\end{aligned}$$

where $\mathbf{H}_{21} = \mathcal{H}_k {}^C \mathbf{R}_I {}^L \widehat{\mathbf{R}}_{I_k}^\top ({}^L \widehat{\mathbf{p}}_{f_k}) \times$, $\mathbf{M}_{21} = -\frac{1}{2} \mathcal{H}_k {}^C \mathbf{R}_I {}^L \widehat{\mathbf{R}}_{I_k}^\top (\mathbf{g}) \times \Delta t_k^2 + \mathcal{H}_k {}^C \mathbf{R}_I {}^L \widehat{\mathbf{R}}_{I_k}^\top ({}^L \widehat{\mathbf{p}}_{f_k}) \times$ and other elements keep the same formulation as (65). The right nullspace \mathcal{N}_N of the observability matrix can be deduced by spanning the following unobservable subspace:

$$\mathcal{N}_N = \begin{bmatrix} \mathbf{g}^\top & \mathbf{0}_{1 \times 3} & \mathbf{0}_{1 \times 3} & \mathbf{0}_{1 \times 3} & \mathbf{0}_{1 \times 3} & \mathbf{g}^\top ({}^L \widehat{\mathbf{p}}_{f_k}) \times \\ \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{bmatrix}^\top. \tag{71}$$

APPENDIX F MSC-INKEF-FDR

A. State Representation

The state definition in MSC-InEKF-FDR is formulated as:

$$\begin{aligned}
\mathbf{X}_k &= (\mathbf{X}_{L_k}, \mathbf{X}_{G_k}) \\
\mathbf{X}_{L_k} &= (\mathbf{X}_{I_k}, \mathbf{B}_k, \mathbf{X}_{C_k}) \\
\mathbf{X}_{I_k} &= \left[\begin{array}{c|cc} {}^L \mathbf{R}_{I_k} & {}^L \mathbf{p}_{I_k} & {}^L \mathbf{v}_{I_k} \\ \hline \mathbf{0}_{2 \times 3} & \mathbf{I}_{2 \times 2} \end{array} \right] \quad \mathbf{B}_k = [\mathbf{b}_{g_k}^\top \quad \mathbf{b}_{a_k}^\top]^\top \\
\mathbf{X}_{G_k} &= \left[\begin{array}{c|cc} {}^L \mathbf{R}_{G_k} & {}^L \mathbf{p}_{G_k} & {}^L \mathbf{p}_{f_k} \\ \hline \mathbf{0}_{2 \times 3} & \mathbf{I}_{2 \times 2} \end{array} \right] \\
\mathbf{X}_{C_k} &= (\mathbf{X}_{CP_{k_k}}, \dots, \mathbf{X}_{CP_{k-c+1_k}}), \quad \mathbf{X}_{CP_{j_k}} = \begin{bmatrix} {}^L \mathbf{R}_{I_{jk}} & {}^L \mathbf{p}_{I_{jk}} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix}.
\end{aligned}$$

The error state keeps the same formulation as Eq. (5) of Section III-B. Preliminary and State Definition in the manuscript.

B. State Propagation

The state propagation is same as that in Section A.

C. State Update

All the derivation of state update equations keep the same as that in Section B.

D. Observability Analysis

In MSC-InEKF-FDN, the formulation of Φ_1^k is same as (63). Therefore, the k -th element of the observability matrix is:

$$\begin{aligned}
\mathbf{H}_k \Phi_1^k = & \begin{bmatrix} \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{H}_{13} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{H}_{21} & \mathbf{H}_{22} & \mathbf{0}_{2 \times 3} & \mathbf{H}_{24} & \mathbf{0}_{2 \times 3} & \mathbf{H}_{26} \\ \mathbf{H}_{31} & \mathbf{H}_{32} & \mathbf{0}_{2 \times 3} & \mathbf{H}_{34} & \mathbf{H}_{35} & \mathbf{0}_{2 \times 3} \\ \mathbf{H}_{41} & \mathbf{H}_{42} & \mathbf{0}_{2 \times 3} & \mathbf{H}_{44} & \mathbf{H}_{45} & \mathbf{0}_{2 \times 3} \end{bmatrix} \\
& \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \frac{1}{2}(\mathbf{g}) \times \Delta t_k^2 & \mathbf{I}_{3 \times 3} & \mathbf{I}_{3 \times 3} \Delta t_k & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ (\mathbf{g}) \times \Delta t_k & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{bmatrix} \\
= & \begin{bmatrix} \mathbf{M}_{11} & \mathbf{0}_{3 \times 3} & \mathbf{M}_{13} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{M}_{21} & \mathbf{M}_{22} & \mathbf{M}_{23} & \mathbf{0}_{2 \times 3} & \mathbf{M}_{24} & \mathbf{0}_{2 \times 3} \\ \mathbf{M}_{31} & \mathbf{M}_{32} & \mathbf{M}_{33} & \mathbf{M}_{34} & \mathbf{M}_{35} & \mathbf{0}_{2 \times 3} \\ \mathbf{M}_{41} & \mathbf{M}_{42} & \mathbf{M}_{43} & \mathbf{M}_{44} & \mathbf{M}_{45} & \mathbf{0}_{2 \times 3} \end{bmatrix} \tag{72}
\end{aligned}$$

where $\mathbf{H}_{21} = -\mathbf{H}_{24} = \mathcal{H}_k {}^C \mathbf{R}_I {}^L \widehat{\mathbf{R}}_{I_k}^\top ({}^L \widehat{\mathbf{p}}_{f_k}) \times$, $\mathbf{M}_{21} = -\frac{1}{2} \mathcal{H}_k {}^C \mathbf{R}_I {}^L \widehat{\mathbf{R}}_{I_k}^\top (\mathbf{g}) \times \Delta t_k^2 + \mathcal{H}_k {}^C \mathbf{R}_I {}^L \widehat{\mathbf{R}}_{I_k}^\top ({}^L \widehat{\mathbf{p}}_{f_k}) \times$ and other elements keep the same formulation as (65). The right nullspace \mathcal{N}_{R_1} of the observability matrix can be deduced by spanning the following unobservable subspace:

$$\mathcal{N}_{R_1} = \text{span} \left(\begin{bmatrix} \mathbf{g}^\top & \mathbf{0}_{1 \times 3} & \mathbf{0}_{1 \times 3} & \mathbf{g}^\top & \mathbf{0}_{1 \times 3} & \mathbf{0}_{1 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{bmatrix} \right)^\top. \tag{73}$$

We can see that when the streetlight-based observations are available, MSC-InEKF-FDN maintains the correct dimension of the nullspace without introducing spurious gain. When the streetlight-based observations and prior pose-based observations are not available, we have:

$$\begin{aligned}
\mathbf{H}_k \Phi_1^k = & \begin{bmatrix} \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{H}_{13} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{H}_{21} & \mathbf{H}_{22} & \mathbf{0}_{2 \times 3} & \mathbf{H}_{24} & \mathbf{0}_{2 \times 3} & \mathbf{H}_{26} \end{bmatrix} \\
& \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \frac{1}{2}(\mathbf{g}) \times \Delta t_k^2 & \mathbf{I}_{3 \times 3} & \mathbf{I}_{3 \times 3} \Delta t_k & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ (\mathbf{g}) \times \Delta t_k & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{bmatrix} \\
= & \begin{bmatrix} \mathbf{M}_{11} & \mathbf{0}_{3 \times 3} & \mathbf{M}_{13} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{M}_{21} & \mathbf{M}_{22} & \mathbf{M}_{23} & \mathbf{M}_{24} & \mathbf{0}_{2 \times 3} & \mathbf{M}_{26} \end{bmatrix}. \tag{74}
\end{aligned}$$

All the elements keep the same formulation as (72). The right nullspace \mathcal{N}_{R_2} of the observability matrix can be deduced by spanning the following unobservable subspace:

$$\mathcal{N}_{R_2} = \begin{bmatrix} \mathbf{g}^\top & \mathbf{0}_{1 \times 3} & \mathbf{0}_{1 \times 3} & \mathbf{g}^\top & \mathbf{0}_{1 \times 3} & \mathbf{0}_{1 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & {}^L\widehat{\mathbf{p}}_{f_k} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} \end{bmatrix}^\top. \quad (75)$$

The lack of streetlight observations and prior poses results in the insufficient constraints on the relative pose state, thus the dimension of unobservable subspace should be 10. However, the feature estimate ${}^L\widehat{\mathbf{p}}_{f_k}$ in \mathcal{N}_{R_2} reduces the dimension to 7 and leads to the inconsistency problem.

APPENDIX G MSC-INEKF-FDRC

A. State Representation

The state definition in MSC-InEKF-FDRC is formulated as:

$$\begin{aligned} \mathbf{X}_k &= (\mathbf{X}_{L_k}, \mathbf{X}_{G_k}) \\ \mathbf{X}_{L_k} &= (\mathbf{X}_{I_k}, \mathbf{B}_k, \mathbf{X}_{C_k}) \\ \mathbf{X}_{I_k} &= \left[\begin{array}{c|cc} {}^L\mathbf{R}_{I_k} & {}^L\mathbf{p}_{I_k} & {}^L\mathbf{v}_{I_k} \\ \hline \mathbf{0}_{2 \times 3} & & \mathbf{I}_{2 \times 2} \end{array} \right] \quad \mathbf{B}_k = [\mathbf{b}_{g_k}^\top \quad \mathbf{b}_{a_k}^\top]^\top \\ \mathbf{X}_{C_k} &= (\mathbf{X}_{CP_{k_k}}, \dots, \mathbf{X}_{CP_{k-c+1_k}}), \quad \mathbf{X}_{CP_{j_k}} = \left[\begin{array}{cc} {}^L\mathbf{R}_{I_{j_k}} & {}^L\mathbf{p}_{I_{j_k}} \\ \mathbf{0}_{1 \times 3} & 1 \end{array} \right] \\ \mathbf{X}_{G_k} &= \left[\begin{array}{c|cc} {}^L\mathbf{R}_{G_k} & {}^L\mathbf{p}_{G_k} & {}^L\mathbf{p}_{f_k} \\ \hline \mathbf{0}_{2 \times 3} & & \mathbf{I}_{2 \times 2} \end{array} \right]. \end{aligned} \quad (76)$$

The error state keeps the same formulation as **Eq. (5)** of **Section III-B. Preliminary and State Definition**. When the feature is associated with the pose clone, the state definition becomes:

$$\begin{aligned} \mathbf{X}_k &= (\mathbf{X}_{I_k}, \mathbf{B}_k, \mathbf{X}_{C_k}, \mathbf{X}_{G_k}) \\ \mathbf{X}_{I_k} &= \left[\begin{array}{c|cc} {}^L\mathbf{R}_{I_k} & {}^L\mathbf{p}_{I_k} & {}^L\mathbf{v}_{I_k} \\ \hline \mathbf{0}_{2 \times 3} & & \mathbf{I}_{2 \times 2} \end{array} \right] \quad \mathbf{B}_k = [\mathbf{b}_{g_k}^\top \quad \mathbf{b}_{a_k}^\top]^\top \\ \mathbf{X}_{C_k} &= (\mathbf{X}_{CP_{k_k}}, \dots, \mathbf{X}_{CP_{k-c+1_k}}), \quad \mathbf{X}_{CP_{k_k}} = \left[\begin{array}{c|cc} {}^L\mathbf{R}_{I_{k_k}} & {}^L\mathbf{p}_{I_{k_k}} & {}^L\mathbf{p}_{f_k} \\ \hline \mathbf{0}_{2 \times 3} & & \mathbf{I}_{2 \times 2} \end{array} \right] \\ \mathbf{X}_{G_k} &= \left[\begin{array}{cc} {}^L\mathbf{R}_{G_k} & {}^L\mathbf{p}_{G_k} \\ \mathbf{0}_{1 \times 3} & 1 \end{array} \right]. \end{aligned} \quad (77)$$

And the feature error state is formulated as:

$${}^L\mathbf{p}_{f_k} = \text{Exp}(\boldsymbol{\xi}_{R_{CP_{k_k}}}) {}^L\widehat{\mathbf{p}}_{f_k} + \mathbf{J}_l(\boldsymbol{\xi}_{p_{CP_{k_k}}}) \boldsymbol{\xi}_{p_{L_{f_k}}}. \quad (78)$$

B. State Propagation

The derivation of state propagation is same as that in **Section A** regardless of whether the feature point is associated with the relative transformation or the pose clone.

C. State Update

When the feature point is associated with the relative transformation, all the state update equations are the same as that in MSC-InEKF-FDR. When the feature point is associated with the pose clone, all the state update equations are the same as that in MSC-InEKF-FC.

D. Observability Analysis

When the feature point is associated with the relative transformation, the unobservable space is same as in (73). When the feature point is associated with the pose clone, the unobservable space is same as in (66). In this way, the dimension of the unobservable subspace is well reserved and the computation cost is significantly reduced compared with MSC-InEKF-FC.

APPENDIX H

EVALUATION OF THE EFFECT OF OUTLIER LIGHT SOURCES

In our system, it is not necessary to distinguish between streetlights and other light sources. Several proposed methods can accurately and efficiently detect and exclude unstable or non-map elements of light sources, thereby mitigating their negative impact on state estimation:

1. As illustrated in **Section V-B. Data Association and Streetlight Observation**, the set of the streetlight clusters is expanded by (-1) to indicate the absence of a matching streetlight cluster for any given detection. In other words, the set is represented as $\{\mathcal{L}_1 = ({}^G\widetilde{\mathbf{c}}_{s_1}, {}^G\widetilde{\mathbf{p}}_{s_1}^1, \dots, {}^G\widetilde{\mathbf{p}}_{s_1}^Q), \dots, \mathcal{L}_{m_k} = ({}^G\widetilde{\mathbf{c}}_{s_{m_k}}, {}^G\widetilde{\mathbf{p}}_{s_{m_k}}^1, \dots, {}^G\widetilde{\mathbf{p}}_{s_{m_k}}^Q), (-1)\}$. The matching score of each detection with (-1) is calculated by $1 - \sum_{g=1}^{m_k} s_{dg}$ where s_{dg} is the matching score of the detection with other streetlight clusters, which is computed using **Eq. (23)** in **Section V-B. Data Association and Streetlight Observation** in the manuscript. In this way, the non-streetlight sources such as car headlights and reflective objects can be detected during the matching process and be assigned to (-1) , indicating there is no correspondence of the false detections.

2. Another way of reducing the effect of false detections is the Chi-square test. When the system state and the observation follow the Gaussian distributions, the square of the Mahalanobis distance from the observation to the corresponding prediction should satisfy a Chi-squared distribution [4]. Therefore, before utilizing the identified streetlight matches for state update, we perform the Chi-squared test to remove the incorrect matches.

To demonstrate that our system is not affected by the outlier light sources, we select the scenarios containing outlier light sources. In addition, for adequate verification, we collect extra sequences where two high-powered flashlights held by a walking person are utilized to simulate the outlier light sources. The results are displayed in Fig. 1 and Fig. 2. It can be observed that although the outlier light sources are falsely detected, the matches of detected streetlights are still correctly identified and no correspondence of false detections is determined.

APPENDIX I

COMPARISON OF DIFFERENT MAP-BASED VISUAL LOCALIZATION METHODS

In this experiment, we compare Night-Voyager with several popular visual localization algorithms which leverage different map types, including methods based on feature point maps (Hloc [5]), 3D line maps (VL-Line [6]), and point cloud maps

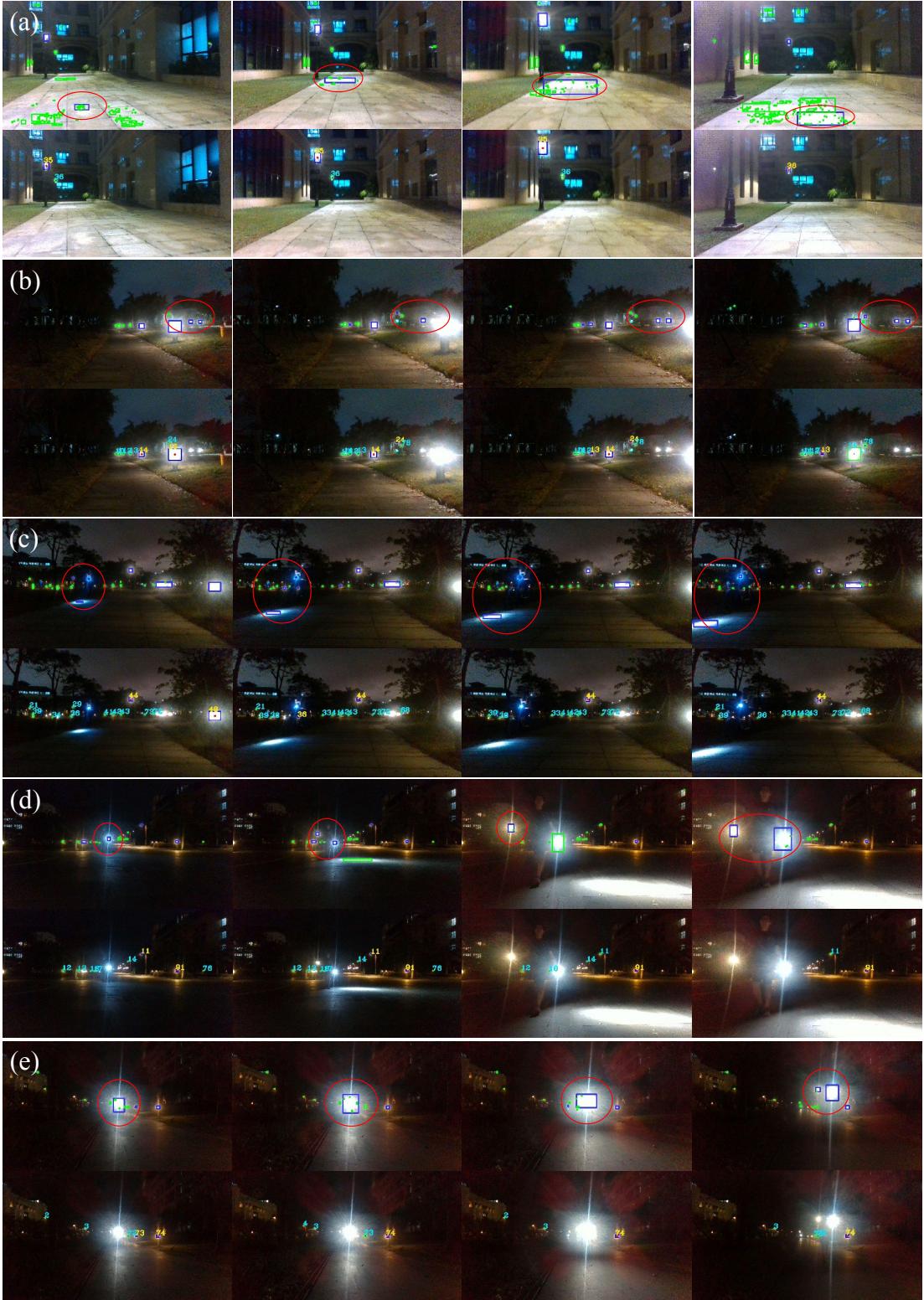


Fig. 1: The matching results using the proposed two strategies. The top figures of each scene are original streetlight detections and the bottom ones represent the final matches. Blue bounding boxes are detected using the learning-based method while the green ones are detected using the binary-based method. The numbers indicate the indices of matched streetlight clusters. For each scenario, we can see that the outliers (red circles) are falsely detected (top figures) while they can be accurately identified and excluded (bottom figures). (a) *Scene_02*. The reflective ground is falsely detected. (b) *Scene_03*. The truck lights are falsely detected. (c) *Scene_04*. A flashlight held by a walking person and the illuminated ground are falsely detected. (d) *Newly Collected Sequence 1*. The flashlights held by a walking person are falsely detected. (e) *Newly Collected Sequence 2*. The flashlights held by a walking person are falsely detected.

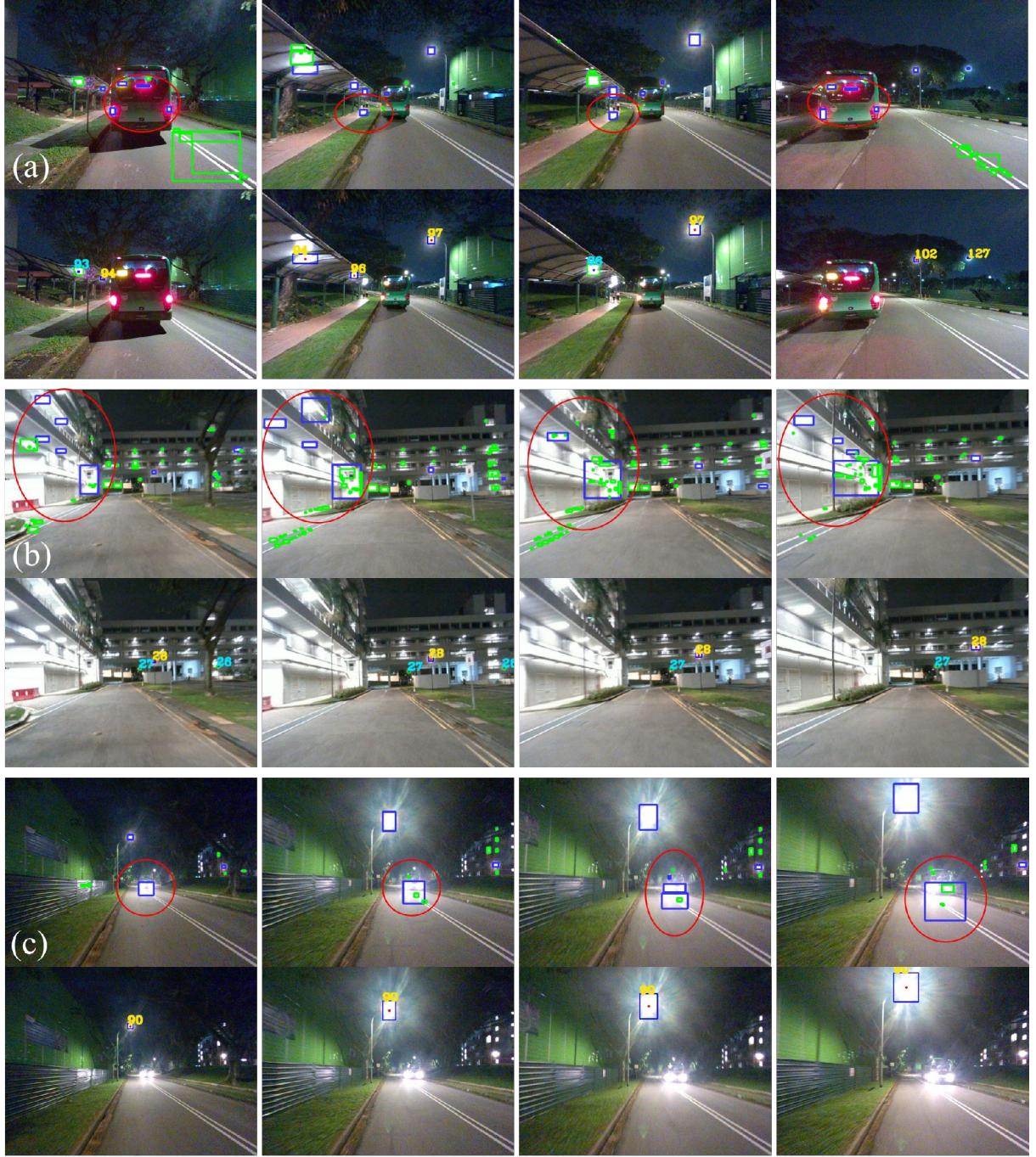


Fig. 2: The matching results in MCD dataset with outlier detections. The top figures of each scene are original streetlight detections and the bottom ones represent the final matches. Blue bounding boxes are detected using the learning-based method while the green ones are detected using the binary-based method. The numbers indicate the indices of matched streetlight clusters. For each scenario, we can see that the outliers (red circles) are falsely detected (top figures) while they can be accurately identified and excluded (bottom figures). (a) *ntu_night_04*. The tail lights of a bus and the reflective ground are falsely detected. (b) *ntu_night_08*. The lights emitted from the buildings are falsely detected. (c) *ntu_night_13*. The head lights of a car and the illuminated ground are falsely detected.

(CMRNet [7] and LHMap [8]). Details of these methods are listed as follows:

1. Hloc [5]. With the development of deep learning, many designed networks [9]–[13] are verified to extract more reliable features and descriptors than manually designed features. Hloc further integrates these feature extractors and matchers

into a framework, achieving coarse-to-fine localization in the feature maps constructed by the Structure-from-Motion (SfM) method, COLMAP [14], [15]. The popular SuperPoint [10] and SuperGlue [9] are utilized for feature map reconstruction and visual localization. The final constructed feature maps are

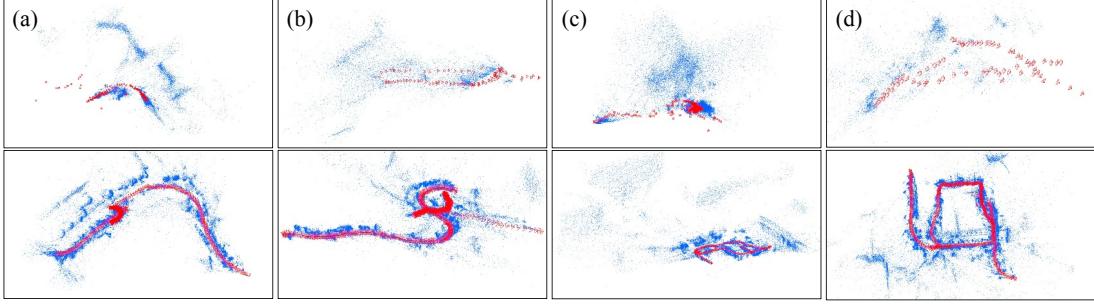


Fig. 3: The constructed feature maps using the nighttime sequences (top row) and the corresponding daytime sequences (bottom row). The sequences in the same column are collected in the identical scene. The 3D features are represented by the blue points and the trajectories are in red. (a) *Scene_02*. (b) *Scene_03*. (c) *Scene_04*. (d) *Scene_08*.

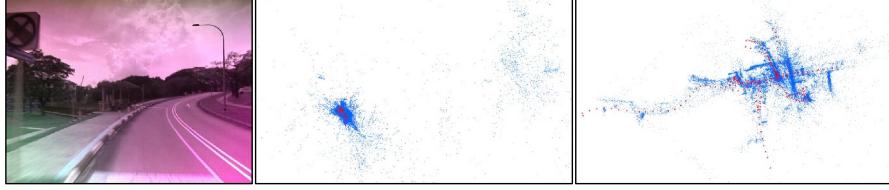


Fig. 4: Failure cases of feature map construction using the daytime or nighttime sequences from MCD dataset. From the left to right: the daytime image with purple blooming, incorrect feature map construction using the daytime sequence, incorrect feature map construction using the nighttime sequence.

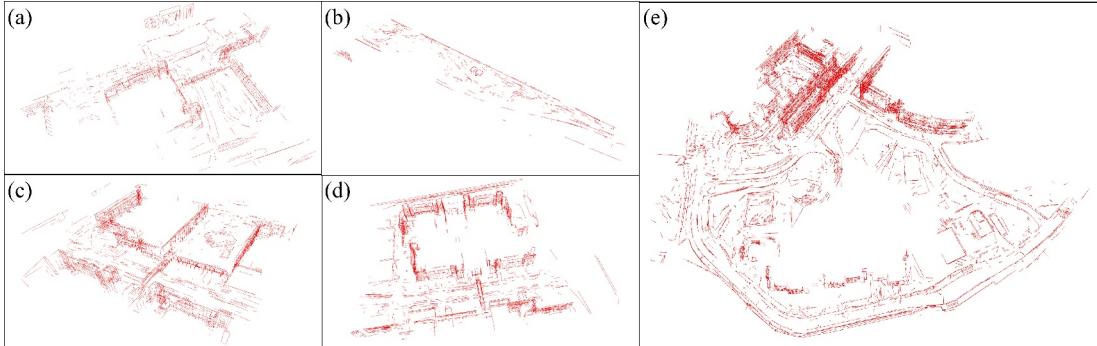


Fig. 5: The constructed line maps for VL-Line using the nighttime sequences. (a) *Scene_01*. (b) *Scene_05*. (c) *Scene_07*. (d) *Scene_09*. (e) Line map for all three sequences of MCD dataset.

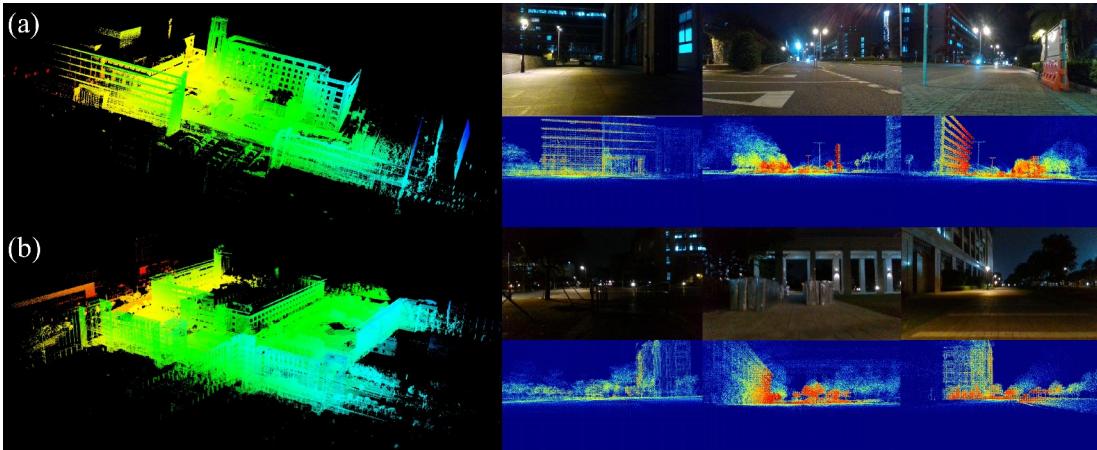


Fig. 6: The constructed point cloud maps using the nighttime sequences. The right figures display the captured images and the corresponding depth images generated by the projection of LiDAR point cloud maps. (a) *Scene_04*. (b) *Scene_07*.

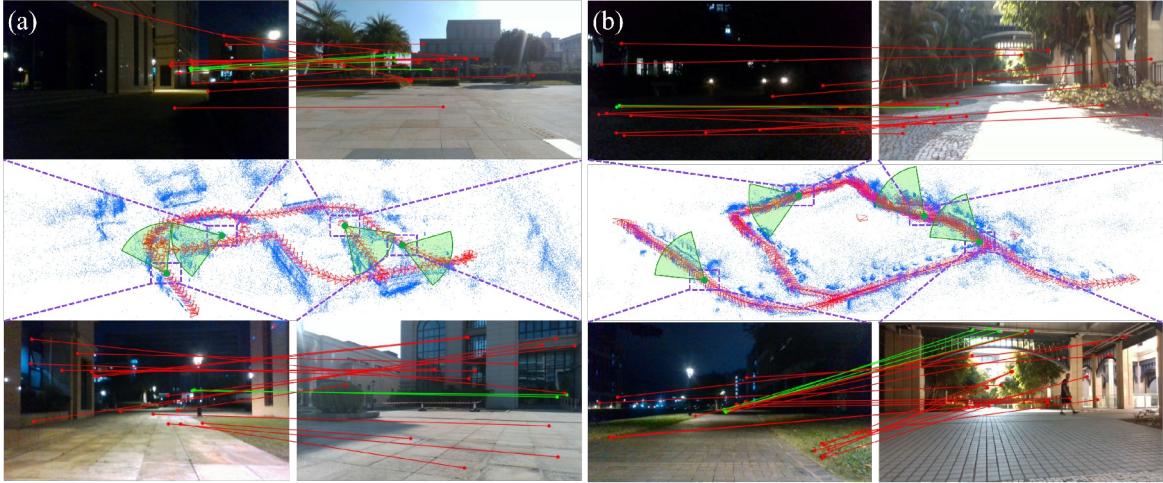


Fig. 7: Hloc failure cases. The middle images show the constructed feature maps and mapping trajectories. In each sequence, the left figures are the captured images during localization and the right ones are the queried images from the mapping sequence. It can be found that there are significantly large localization errors between the localization images and the queried images. (a) *Scene_04*. (b) *Scene_08*.



Fig. 8: VL-Line failure cases. The red lines in the images denote the detected 2D lines and the green ones represent the matched projected lines. Despite VL-Line is based on structural line features, it cannot fundamentally address the *insufficiency* issue inherent in nocturnal visual problems. (a) *Scene_01*. (b) *Scene_06*. (c) *Scene_10*.

presented in Fig. 3. Note that the constructed feature maps using the data from nighttime sequences are rambling due to the insufficient and inconsistent feature points. Therefore, we additionally collect daytime sequences for constructing

consistent feature maps. As for the public MCD dataset, the constructed feature map using the daytime sequence is not satisfactory as shown in Fig. 4 because the images are affected by a purple cast. Furthermore, the feature map constructed

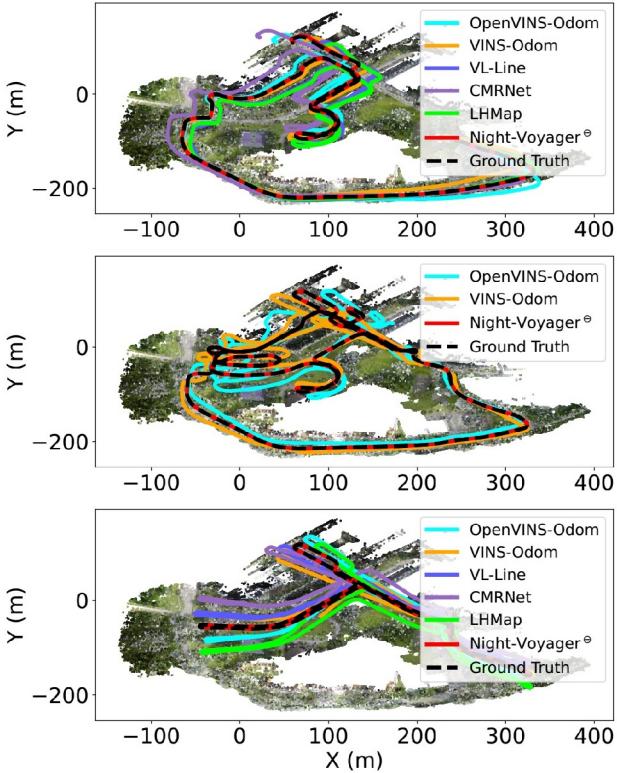


Fig. 9: Estimated trajectories of different algorithms on the prior map. Top: ntu_night_04. Middle: ntu_night_08. Bottom: ntu_night_13.

from the nighttime sequence is also fragmented and chaotic due to the drastically changed illumination. Furthermore, we also try to feed the system even with the ground-truth mapping trajectory for map construction, but the performance still remains significantly poor. Consequently, the comparison for Hloc is conducted only on the collected dataset, not on the MCD dataset.

2. VL-Line [6]. This algorithm leverages both geometric point and line features to achieve accurate visual localization in line maps. Compared to VINS-Mono, strengthened with a loop closure detection module, it still demonstrates superior performance in the EuRoC dataset [16]. After constructing the Light Detection And Ranging (LiDAR) point cloud map, VL-Line utilizes a 3D line detection method [17] to extract the structure lines. During localization, AFM [18] is utilized for online 2D line detection. With the prior poses provided by VINS-Mono [19], line matches are identified for pose refinement. We replace the VINS-Mono with the version fused with the odometer, i.e., VINS-Odom [20], for a fair comparison. We employ the LiDAR data in the collected mapping sequences to construct the line maps. The survey-grade prior maps in MCD dataset are directly used for 3D line map construction. The final line maps are shown in Fig. 5.

3. CMRNet [21]. CMRNet is a popular state-of-the-art (SOTA) learning-based visual localization method that employs a point cloud map as prior. Taking a color image and a sparse depth image obtained by projecting the LiDAR map



Fig. 10: Initialization cases. The images with red bounds indicate the wrong initialization. The images with green bounds represent the correct initialization after using the provided coarse position. The green numbers represent the indices of the matched streetlight clusters in the map. -1 represents no matches are identified for the detection. It is seen that the incorrect initialization caused by overlapped streetlight detections and repetitive streetlight distribution patterns (as mentioned in **Section IX-E. Evaluation of Initialization** in the manuscript) is corrected thanks to the coarse position.

according to the coarse pose as input, the designed networks predict a refined pose for localization. For a fair comparison, we design a system which utilizes MSC-InEKF-FDRC as the estimator to fuse IMU and odometer measurements, feature points, and CMRNet pose prediction to achieve continuous and robust state estimation. We present the LiDAR point cloud maps and the projected depth images used for CMRNet in Fig. 6.

4. LHMap [8]. Similar to CMRNet, LHMap also takes a pair of color and depth images as the input of the neural network and predicts a refined pose. Based on optical flow estimation and spatial attention, it achieves SOTA performance in both KITTI [22] and Argoverse [23] datasets. To facilitate a fair comparison, we also design a system utilizing MSC-InEKF-FDRC to fuse the LHMap predicted poses with IMU, odometer measurements, and feature points.

To test their performance on MCD dataset, we utilize the data of *ntu_night_08* to finetune the networks (except for Hloc due to the feature map construction failure). When testing the algorithms on the collected datasets, we finetune the networks with the mapping sequences. The final localization results on the two datasets are reported in **Table IV** and **Table VI** in the manuscript, respectively. We can see that Night-Voyager significantly outperforms all the map-based localization methods. As shown in Fig. 7, the performance of Hloc is significantly affected by the severe illumination differences between daytime and nighttime images. 3D line maps have limited effectiveness in visual localization because a structured scene is required for VL-Line. However, as shown in Fig. 8, the textural lines of traffic lanes and floor tile gaps, along with spurious line detections caused by uneven lighting, influence the line association. Essentially, the poor performance of Hloc and VL-Line stems from the *insufficiency*

TABLE I: SR (%), ART (s), and AAPE (m/ $^{\circ}$) of the initialization module on different sequences.

		Scene_01	Scene_02	Scene_03	Scene_04	Scene_05	Scene_06	Scene_07	Scene_08	Scene_09	Scene_10
w/o Coarse Position	SR	81.3	79.2	100.0	100.0	58.9	60.9	75.0	100.0	100.0	81.6
	ART	2.47	5.38	0.76	0.69	4.82	1.76	6.49	5.72	2.11	6.11
	AAPE	0.25/0.40	0.18/0.51	0.32/1.66	0.31/0.28	0.24/0.74	0.21/0.63	0.25/0.34	0.35/0.74	0.34/0.45	0.25/0.57
w/ Coarse Position (10 m)	SR	88.9	94.2	100.0	100.0	85.7	87.8	81.8	100.0	100.0	90.0
	ART	0.27	0.50	0.09	0.11	0.51	0.18	0.35	0.47	0.13	0.41
	AAPE	0.25/0.48	0.20/0.64	0.32/1.66	0.31/0.28	0.27/0.70	0.22/0.57	0.23/0.31	0.35/0.74	0.34/0.45	0.25/0.55
Hloc [5]	SR	0.00	0.12	0.00	0.19	0.00	0.00	0.00	0.09	0.00	0.00
	ART	21.88	24.77	19.56	15.21	26.47	29.02	25.35	23.26	29.56	31.18
	AAPE	—	0.29/3.33	—	0.45/1.69	—	—	—	0.36/2.13	—	—
CMRNet [7] (10 m, 20°)	SR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	ART	0.37	0.26	0.31	0.73	0.56	0.39	0.37	0.48	0.29	0.18
	AAPE	—	—	—	—	—	—	—	—	—	—
CMRNet [7] (2 m, 10°)	SR	0.00	0.02	0.00	0.00	0.00	0.00	0.09	0.00	0.00	0.05
	ART	0.35	0.29	0.26	0.55	0.47	0.39	0.33	0.53	0.32	0.27
	AAPE	—	0.49/6.40	—	—	—	—	0.35/4.17	—	—	0.25/6.49
LHMap [8] (10 m, 20°)	SR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	ART	0.38	0.37	0.36	0.76	0.52	0.38	0.37	0.40	0.35	0.26
	AAPE	—	—	—	—	—	—	—	—	—	—
LHMap [8] (2 m, 10°)	SR	0.00	0.03	0.00	0.00	0.00	0.00	0.02	0.04	0.06	0.05
	ART	0.27	0.45	0.32	0.79	0.52	0.33	0.30	0.49	0.50	0.19
	AAPE	—	0.46/4.77	—	—	—	—	0.31/5.95	0.48/4.92	0.41/6.05	0.36/3.89

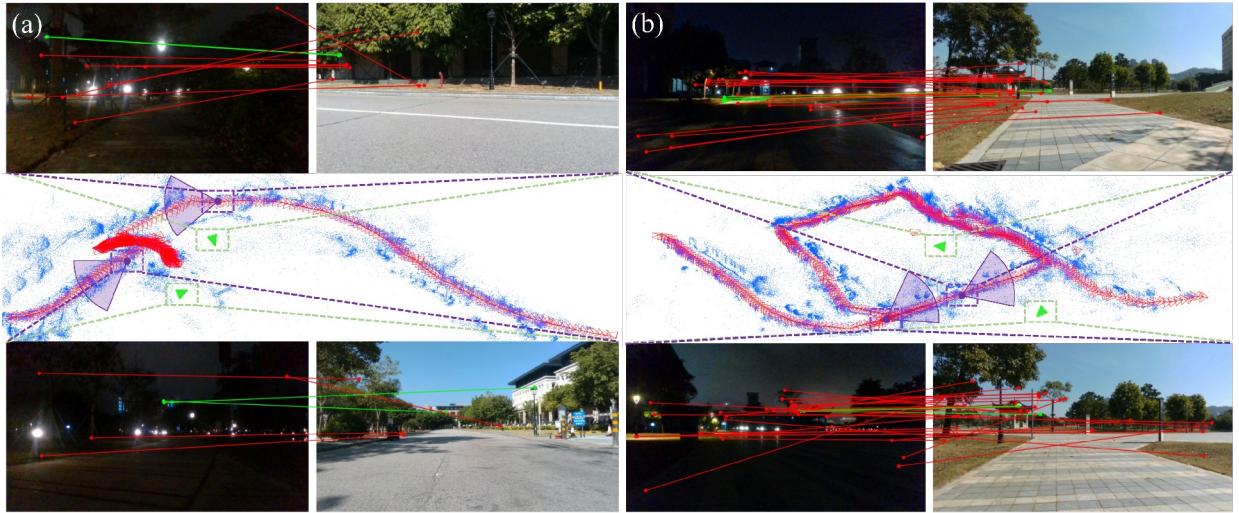


Fig. 11: Failure cases of Hloc. In each sequence, the left figures are the tested images and the right ones are the queried images from the mapping sequence. The middle figure shows the constructed feature map and mapping trajectory. The green triangulations indicate the predicted pose in the feature map while the purple circles with sectors indicate the ground truth pose. Note that all the labels denote the poses during the localization process (i.e., poses corresponding to the left figures in each scene.). The poses corresponding to the queried images of are not annotated, which is different from **Fig. 16 of Section IX-D. Evaluation on the Collected Dataset** in the revised manuscript. For the poses of queried mapping images, **Fig. 16** in the revised manuscript provides a similar annotation style for reference, indicating large errors between the localization images and queried images. (a) Scene_02. (b) Scene_08.

and *inconsistency* issues of pixel-level features. Due to the limited information at night, the improvement of CMRNet and LHMap is much inferior to that of Night-Voyager, which demonstrates the substantial benefits brought by the object-level map and data association. We also demonstrate the trajectories generated by different algorithms in Fig. 9. The almost perfect alignment of the Night-Voyager trajectory with the ground truth proves the effectiveness of our method.

APPENDIX J COMPARISON OF DIFFERENT INITIALIZATION METHODS

When the coarse pose is provided for initialization, the success rate can be significantly improved. The failure initialization caused by the overlapped streetlight detection and repetitive streetlight distribution patterns that are mentioned in **Section IX-E. Evaluation of Initialization** in the manuscript is corrected as shown in Fig. 10.

To the best of our knowledge, no prior work has explored global localization in sparse object maps. Additionally, most map-based localization methods [6], [8], [21], [24]–

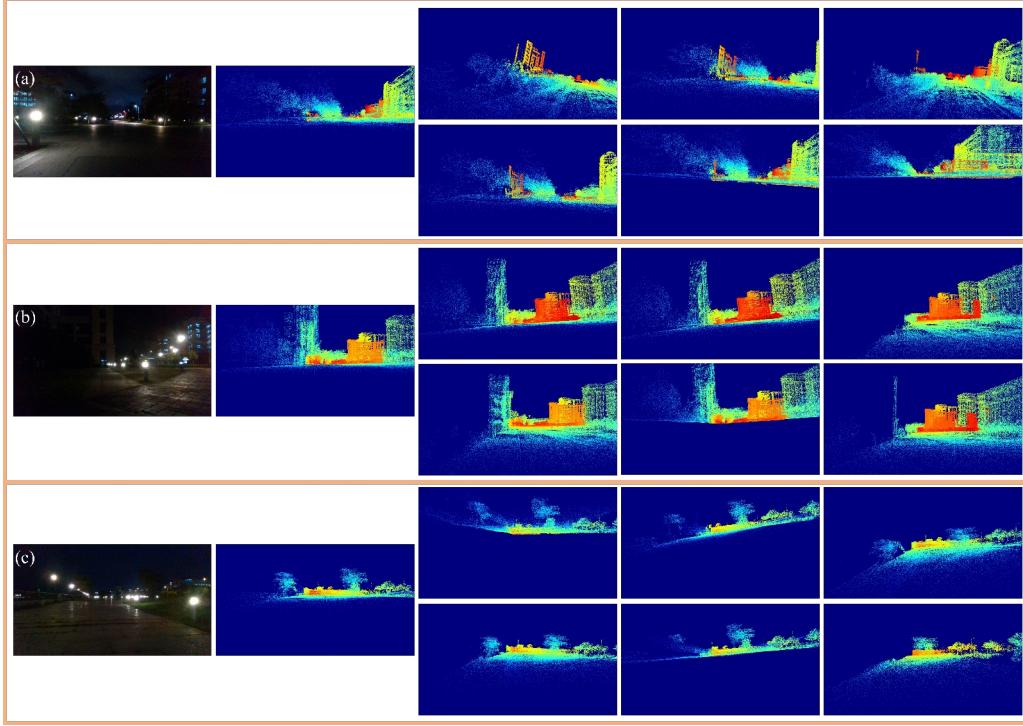


Fig. 12: Failure cases of CMRNet and LHMap. From the left to right: tested images, projections of point cloud maps using the ground truth poses, projections of point cloud maps using the perturbed poses, projections of point cloud maps using the poses refined by CMRNet, projections of point cloud maps using the poses refined by LHMap. For each scene, the top rows show results with 10 m translation and 20° rotation perturbations, while the bottom rows show results with 2 m translation and 10° rotation perturbations. (a) *Scene_01*. (b) *Scene_03*. (c) *Scene_05*

[27] assume that an accurate initial pose can be provided, not focusing on solving the global localization problem. Therefore, we select two representative kinds of methods to perform comparison experiments. The first one is Hloc [5]. We utilize finetuned SuperPoint [10] and SuperGlue [9] for feature map reconstruction and visual localization. Due to insufficient feature points at night, building a feature point map is infeasible. Therefore, we utilize the additionally collected daytime sequences for feature map construction. Additionally, we also compare our proposed initialization module with the two open-sourced approaches, CMRNet [7] and LHMap [8]. We provide the coarse pose by corrupting the ground truth with two additive Gaussian noises (10 m in the translation direction and 20° in the rotation, as well as 2 m in the translation direction and 10° in the rotation).

The results are shown in Table I. Compared with the proposed initialization module, Hloc, CMRNet, and LHMap almost fail across all sequences. Fig. 11 shows the initialization performance of Hloc. It is observed that the initialized poses are far from the true values. The essential reason lies in the poor information in nighttime images. The qualitative results of CMRNet and LHMap are presented in Fig. 12. Even the poses provided to CMRNet and LHMap are corrupted by a perturbation of 2 m translation and 10° rotation, they fail to refine the poses to the correct values. In contrast, utilizing streetlight information, our initialization module is able to calculate most initial poses correctly.

REFERENCES

- [1] Y. Yang, B. P. W. Babu, C. Chen, G. Huang, and L. Ren, “Analytic combined imu integration (aci 2) for visual inertial navigation,” in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 4680–4686. [1](#)
- [2] Y. Yang, C. Chen, W. Lee, and G. Huang, “Decoupled right invariant error states for consistent visual-inertial navigation,” *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 1627–1634, 2022. [1](#)
- [3] M. Li and A. I. Mourikis, “High-precision, consistent ekf-based visual-inertial odometry,” *The International Journal of Robotics Research*, vol. 32, no. 6, pp. 690–711, 2013. [6](#)
- [4] G. Chang, “Robust kalman filtering based on mahalanobis distance as outlier judging criterion,” *Journal of Geodesy*, vol. 88, no. 4, pp. 391–401, 2014. [9](#)
- [5] P.-E. Sarlin, C. Cadena, R. Siegwart, and M. Dymczyk, “From coarse to fine: Robust hierarchical localization at large scale,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 12 716–12 725. [9, 11, 15, 16](#)
- [6] H. Yu, W. Zhen, W. Yang, J. Zhang, and S. Scherer, “Monocular camera localization in prior lidar maps with 2d-3d line correspondences,” in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 4588–4594. [9, 14, 15](#)
- [7] D. Cattaneo, M. Vaghi, A. L. Ballardini, S. Fontana, D. G. Sorrenti, and W. Burgard, “Cmrnet: Camera to lidar-map registration,” in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, Oct 2019, pp. 1283–1289. [11, 15, 16](#)
- [8] X. Wu, J. Xu, P. Hu, G. Wang, and H. Wang, “Lhmap-loc: Cross-modal monocular localization using lidar point cloud heat map,” in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 2024, pp. 8500–8506. [11, 14, 15, 16](#)
- [9] P.-E. Sarlin, D. DeTone, T. Malisiewicz, and A. Rabinovich, “Superglue: Learning feature matching with graph neural networks,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 4938–4947. [11, 16](#)
- [10] D. DeTone, T. Malisiewicz, and A. Rabinovich, “Superpoint: Self-supervised interest point detection and description,” in *Proceedings*

- of the IEEE conference on computer vision and pattern recognition workshops*, 2018, pp. 224–236. 11, 16
- [11] J. Revaud, C. De Souza, M. Humenberger, and P. Weinzaepfel, “R2d2: Reliable and repeatable detector and descriptor,” *Advances in neural information processing systems*, vol. 32, 2019. 11
- [12] M. Tyszkiewicz, P. Fua, and E. Trulls, “Disk: Learning local features with policy gradient,” *Advances in Neural Information Processing Systems*, vol. 33, pp. 14 254–14 265, 2020. 11
- [13] P. Lindenberger, P.-E. Sarlin, and M. Pollefeys, “Lightglue: Local feature matching at light speed,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 17 627–17 638. 11
- [14] J. L. Schönberger and J.-M. Frahm, “Structure-from-motion revisited,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 4104–4113. 11
- [15] J. L. Schönberger, E. Zheng, J.-M. Frahm, and M. Pollefeys, “Pixelwise view selection for unstructured multi-view stereo,” in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part III* 14. Springer, 2016, pp. 501–518. 11
- [16] M. Burri, J. Nikolic, P. Gohl, T. Schneider, J. Rehder, S. Omari, M. W. Achtelik, and R. Siegwart, “The euroc micro aerial vehicle datasets,” *The International Journal of Robotics Research*, vol. 35, no. 10, pp. 1157–1163, 2016. 14
- [17] X. Lu, Y. Liu, and K. Li, “Fast 3d line segment detection from unorganized point cloud,” *arXiv preprint arXiv:1901.02532*, 2019. 14
- [18] N. Xue, S. Bai, F. Wang, G.-S. Xia, T. Wu, and L. Zhang, “Learning attraction field representation for robust line segment detection,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 1595–1603. 14
- [19] T. Qin, P. Li, and S. Shen, “Vins-mono: A robust and versatile monocular visual-inertial state estimator,” *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 1004–1020, 2018. 14
- [20] J. Liu, W. Gao, and Z. Hu, “Visual-inertial odometry tightly coupled with wheel encoder adopting robust initialization and online extrinsic calibration,” in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 5391–5397. 14
- [21] D. Sorrenti, C. Daniele, A. Valada *et al.*, “Cmnet++: Map and camera agnostic monocular visual localization in lidar maps,” in *Proceeding of ICRA 2020 Workshop on Emerging Learning and Algorithmic Methods for Data Association in Robotics* <https://sites.google.com/view/edai/home> also available as arxiv report <https://arxiv.org/abs/2004.13795>, 2020. 14, 15
- [22] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, “Vision meets robotics: The kitti dataset,” *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1231–1237, 2013. 14
- [23] M.-F. Chang, J. Lambert, P. Sangkloy, J. Singh, S. Bak, A. Hartnett, D. Wang, P. Carr, S. Lucey, D. Ramanan *et al.*, “Argoverse: 3d tracking and forecasting with rich maps,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 8748–8757. 14
- [24] H. Wang, C. Xue, Y. Zhou, F. Wen, and H. Zhang, “Visual semantic localization based on hd map for autonomous vehicles in urban scenarios,” in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 11 255–11 261. 15
- [25] Z. Liao, J. Shi, X. Qi, X. Zhang, W. Wang, Y. He, X. Liu, and R. Wei, “Coarse-to-fine visual localization using semantic compact map,” in *2020 3rd International Conference on Control and Robots (ICCR)*. IEEE, 2020, pp. 30–37. 15
- [26] C. Zhang, H. Zhao, C. Wang, X. Tang, and M. Yang, “Cross-modal monocular localization in prior lidar maps utilizing semantic consistency,” in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 4004–4010. 15
- [27] X. Zuo, P. Geneva, Y. Yang, W. Ye, Y. Liu, and G. Huang, “Visual-inertial localization with prior lidar map constraints,” *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3394–3401, 2019. 15