
Supplementary Material:
T-ESKF: Transformed Error-State Kalman Filter for
Consistent Visual-Inertial Navigation

Chungeng Tian
Ning Hao
Fenghua He

2024-07

Contents

1 Structure of This Supplementary Material	3
2 Preliminaries: Special Orthogonal Group	3
2.1 Lie algebra	3
2.2 Exponential and logarithmic map	3
2.3 Adjoint operation	4
2.4 Jacobians	4
3 IMU model and error-state kinematics	5
3.1 IMU model	5
3.2 IMU state estimate and error-state	5
3.3 IMU error-state kinematics	6
3.3.1 Orientation	7
3.3.2 Position	8
3.3.3 Velocity	8
3.3.4 Bias	8
4 Visual-Inertial Navigation System	8
4.1 System model	8
4.2 Linearized error-state system	9
5 Transformed Linearized Error-State System	11
5.1 Linear time-varying transformation on error-state	11
5.2 Observability of the transformed system	12
6 Transformed Error-State Kalman Filter	14
6.1 Propagation	14
6.2 Update	16
6.3 T-ESKF properties	18
7 Real-World Experiments	19
A Derivation for (28) in the manuscript	22
B Derivation for (82) and (83)	23
C T-ESKF state update equation	24
C.1 Update with (104): $\hat{\mathbf{x}}_{k+1 k+1} = \hat{\mathbf{x}}_{k+1 k} \oplus (\mathbf{T}(\hat{\mathbf{x}}_{k+1 k})^{-1} \delta \mathbf{x}^*)$	24
C.2 Update with (103): $\hat{\mathbf{x}}_{k+1 k+1} = \hat{\mathbf{x}}_{k+1 k} \oplus (\mathbf{T}(\hat{\mathbf{x}}_{k+1 k+1})^{-1} \delta \mathbf{x}^*)$	25

1 Structure of This Supplementary Material

In this supplementary material, Section 2 provides fundamental knowledge of $\mathbf{SO}(3)$. Sections 3-6 present the model and method encompassing the consideration of IMU bias. The real-world experiments on our customized platform are detailed in Section 7. Appendix A contains the derivation of (28), which is omitted from the manuscript. Appendix A provides the state update equations of T-ESKF.

2 Preliminaries: Special Orthogonal Group

2.1 Lie algebra

The Lie algebra corresponding to $\mathbf{SO}(3)$ is denoted by $\mathfrak{so}(3)$. It consists of all 3×3 skew-symmetric matrices, i.e.,

$$\mathfrak{so}(3) = \{\boldsymbol{\theta}^\wedge | \boldsymbol{\theta} \in \mathbb{R}^3\}, \quad (1)$$

where a general element of $\mathfrak{so}(3)$ can be written as $\boldsymbol{\theta}^\wedge = [\boldsymbol{\theta}]_\times$ (the operator $(\cdot)^\wedge$ converts a vector to its corresponding skew-symmetric matrix form) with

$$[\boldsymbol{\theta}]_\times = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix}_\times = \begin{bmatrix} 0 & -\theta_3 & \theta_2 \\ \theta_3 & 0 & -\theta_1 \\ -\theta_2 & \theta_1 & 0 \end{bmatrix}. \quad (2)$$

2.2 Exponential and logarithmic map

Let $\boldsymbol{\theta}$ represent a rotation vector, and $\theta \triangleq |\boldsymbol{\theta}|$ be the rotation angle and $\mathbf{u} \triangleq \frac{\boldsymbol{\theta}}{|\boldsymbol{\theta}|}$ the rotation axis. The exponential map $\exp : \mathfrak{so}(3) \rightarrow \mathbf{SO}(3)$ transforms the space of $\mathfrak{so}(3)$ to the space of the rotations represented by rotation matrices

$$\exp([\boldsymbol{\theta}]_\times) = \mathbf{I}_3 + \frac{1}{1!} [\boldsymbol{\theta}]_\times + \frac{1}{2!} [\boldsymbol{\theta}]_\times^2 + \frac{1}{3!} [\boldsymbol{\theta}]_\times^3 + \dots \quad (3a)$$

$$= \mathbf{I}_3 + \left(\frac{\theta}{1!} - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \dots \right) [\mathbf{u}]_\times + \left(\frac{\theta^2}{2!} - \frac{\theta^4}{4!} + \frac{\theta^6}{6!} - \dots \right) [\mathbf{u}]_\times^2 \quad (3b)$$

$$= \mathbf{I}_3 + \sin\theta [\mathbf{u}]_\times + (1 - \cos\theta) [\mathbf{u}]_\times^2 \quad (3c)$$

$$= \mathbf{I}_3 \cos\theta + [\mathbf{u}]_\times \sin\theta + \mathbf{u} \mathbf{u}^\top (1 - \cos\theta) \quad (3d)$$

$$\triangleq \mathbf{R}. \quad (3e)$$

The exponential map $\text{Exp} : \mathbb{R}^3 \rightarrow \mathbf{SO}(3)$ is defined as

$$\text{Exp}(\boldsymbol{\theta}) = \exp([\boldsymbol{\theta}]_\times). \quad (4)$$

The logarithmic map is defined as the inverse of the exponential map. Specifically, the map $\log : \mathbf{SO}(3) \rightarrow \mathfrak{so}(3)$ is given by

$$\log(\mathbf{R}) = [\mathbf{u}\theta]_\times \quad (5)$$

and the map $\text{Log} : \mathbf{SO}(3) \rightarrow \mathbb{R}^3$ is given by

$$\text{Log}(\mathbf{R}) = \mathbf{u}\theta. \quad (6)$$

For a given rotation matrix $\mathbf{R} \in \mathbf{SO}(3)$, the logarithmic map can be expressed as:

$$\theta = \arccos\left(\frac{\text{trace}(\mathbf{R}) - 1}{2}\right) \quad (7)$$

$$\mathbf{u} = \frac{(\mathbf{R} - \mathbf{R}^\top)^\vee}{2\sin\theta} \quad (8)$$

where the operator $(\cdot)^\vee$ is the inverse of the operator $(\cdot)^\wedge$.

2.3 Adjoint operation

For $\mathbf{R} \in \mathbf{SO}(3)$, the adjoint operation $Ad_{\mathbf{R}}$ is defined as follows:

$$Ad_{\mathbf{R}}(\phi) = \mathbf{R}\phi, \quad (9)$$

where $\phi \in \mathbb{R}^3$ is a vector, and $Ad_{\mathbf{R}} = \mathbf{R}$ denotes the operation which applies the rotation matrix \mathbf{R} to a vector.

For completeness, the adjoint operation can also be expressed in terms of the matrix representation:

$$[Ad_{\mathbf{R}}\phi]_\times = \mathbf{R}[\phi]_\times \mathbf{R}^\top. \quad (10)$$

Then the exponential map satisfies the following relationship:

$$\text{Exp}(Ad_{\mathbf{R}}\phi) = \mathbf{R}\text{Exp}(\phi)\mathbf{R}^\top. \quad (11)$$

2.4 Jacobians

Assuming that ϕ is small enough, the exponential map can be approximated by

$$\text{Exp}(\boldsymbol{\theta} + \phi) \approx \text{Exp}(\boldsymbol{\theta})\text{Exp}(\mathbf{J}_r(\boldsymbol{\theta})\phi) \quad (12)$$

$$\text{Exp}(\boldsymbol{\theta} + \phi) \approx \text{Exp}(\mathbf{J}_l(\boldsymbol{\theta})\phi)\text{Exp}(\boldsymbol{\theta}) \quad (13)$$

where

$$\mathbf{J}_r(\boldsymbol{\theta}) = \frac{\sin\theta}{\theta}\mathbf{I}_3 + \left(1 - \frac{\sin\theta}{\theta}\right)\mathbf{u}\mathbf{u}^\top - \frac{1 - \cos\theta}{\theta}[\mathbf{u}]_\times \quad (14)$$

and $\mathbf{J}_l(\boldsymbol{\theta}) = \mathbf{J}_r(-\boldsymbol{\theta})$. \mathbf{J}_r and \mathbf{J}_l are called right and left Jacobians of $\mathbf{SO}(3)$, respectively.

3 IMU model and error-state kinematics

This section presents the IMU model and its error-state kinematics. The main result is equivalent to what is described in *The ESKF using global angular errors* [1, Chapter 7]. If you are familiar with ESKF, you can skip this section.

3.1 IMU model

The IMU state vector is defined as

$$\mathbf{x}_I = (\mathbf{R}, \mathbf{p}, \mathbf{v}, \mathbf{b}_g, \mathbf{b}_a), \quad (15)$$

where $\mathbf{R} \in \mathbf{SO}(3)$, $\mathbf{p} \in \mathbb{R}^3$ and $\mathbf{v} \in \mathbb{R}^3$ are the orientation, position, and velocity of IMU in the global frame, $\mathbf{b}_g \in \mathbb{R}^3$ and $\mathbf{b}_a \in \mathbb{R}^3$ are the gyroscope bias and accelerometer bias, respectively.

The continuous motion model for the IMU state vector is given by the following differential equations:

$$\dot{\mathbf{R}} = \mathbf{R} [\boldsymbol{\omega}]_{\times} \quad (16a)$$

$$\dot{\mathbf{p}} = \mathbf{v} \quad (16b)$$

$$\dot{\mathbf{v}} = \mathbf{a} \quad (16c)$$

$$\dot{\mathbf{b}}_g = \mathbf{n}_{gw} \quad (16d)$$

$$\dot{\mathbf{b}}_a = \mathbf{n}_{aw} \quad (16e)$$

where

$$\boldsymbol{\omega} = \boldsymbol{\omega}_m - \mathbf{b}_g - \mathbf{n}_g \quad (17)$$

$$\mathbf{a} = \mathbf{R}(\mathbf{a}_m - \mathbf{b}_a - \mathbf{n}_a) + \mathbf{g} \quad (18)$$

$\boldsymbol{\omega}_m \in \mathbb{R}^3$ and $\mathbf{a}_m \in \mathbb{R}^3$ are the gyroscope and accelerometer measurements in the IMU frame, $\mathbf{n}_{gw} \sim \mathcal{N}(0, \sigma_{gw}^2 \mathbf{I}_3)$, $\mathbf{n}_{aw} \sim \mathcal{N}(0, \sigma_{aw}^2 \mathbf{I}_3)$, $\mathbf{n}_g \sim \mathcal{N}(0, \sigma_g^2 \mathbf{I}_3)$, and $\mathbf{n}_a \sim \mathcal{N}(0, \sigma_a^2 \mathbf{I}_3)$ are assumed to be white Gaussian noises.

3.2 IMU state estimate and error-state

Let $\hat{\mathbf{x}}_I$ denote the IMU state estimate, which is propagated following (16) by setting the process noise to zero:

$$\dot{\hat{\mathbf{R}}} = \hat{\mathbf{R}} [\hat{\boldsymbol{\omega}}]_{\times} \quad (19a)$$

$$\dot{\hat{\mathbf{p}}} = \hat{\mathbf{v}} \quad (19b)$$

$$\dot{\hat{\mathbf{v}}} = \hat{\mathbf{a}} \quad (19c)$$

$$\dot{\hat{\mathbf{b}}}_g = \mathbf{0} \quad (19d)$$

$$\dot{\hat{\mathbf{b}}}_a = \mathbf{0} \quad (19e)$$

with

$$\hat{\omega} = \omega_m - \hat{\mathbf{b}}_g \quad (20)$$

$$\hat{\mathbf{a}} = \hat{\mathbf{R}}(\mathbf{a}_m - \hat{\mathbf{b}}_a) + \mathbf{g}. \quad (21)$$

In the classical VINS estimator, the error-state Kalman filter (ESKF), the error-state representation is used to handle uncertainties in the state estimates:

$$\tilde{\mathbf{x}}_I = \mathbf{x}_I \ominus \hat{\mathbf{x}}_I \quad (22)$$

where the error-state of orientation is defined using the logarithm of $\mathbf{SO}(3)$ while the errors of other variables are defined on vector space, the general minus \ominus is defined as

$$\tilde{\boldsymbol{\theta}} = \text{Log}(\mathbf{R}\hat{\mathbf{R}}^{-1}) \quad (23a)$$

$$\tilde{\mathbf{p}} = \mathbf{p} - \hat{\mathbf{p}} \quad (23b)$$

$$\tilde{\mathbf{v}} = \mathbf{v} - \hat{\mathbf{v}} \quad (23c)$$

$$\tilde{\mathbf{b}}_g = \mathbf{b}_g - \hat{\mathbf{b}}_g \quad (23d)$$

$$\tilde{\mathbf{b}}_a = \mathbf{b}_a - \hat{\mathbf{b}}_a. \quad (23e)$$

This error-state formulation allows for the use of standard Kalman filter techniques to estimate the uncertainties and correct the state estimates.

3.3 IMU error-state kinematics

The IMU error-state kinematics is

$$\dot{\tilde{\boldsymbol{\theta}}} = -\hat{\mathbf{R}}\tilde{\mathbf{b}}_g - \hat{\mathbf{R}}\mathbf{n}_g \quad (24a)$$

$$\dot{\tilde{\mathbf{p}}} = \tilde{\mathbf{v}} \quad (24b)$$

$$\dot{\tilde{\mathbf{v}}} = -[\hat{\mathbf{R}}(\mathbf{a}_m - \hat{\mathbf{b}}_a)]_{\times} \tilde{\boldsymbol{\theta}} - \hat{\mathbf{R}}\tilde{\mathbf{b}}_a - \hat{\mathbf{R}}\mathbf{n}_a \quad (24c)$$

$$\dot{\tilde{\mathbf{b}}}_g = \mathbf{n}_{gw} \quad (24d)$$

$$\dot{\tilde{\mathbf{b}}}_a = \mathbf{n}_{aw}. \quad (24e)$$

The derivation of the equations will be given in Section 3.3.1-3.3.4. The above equations can be written in matrix form:

$$\dot{\tilde{\mathbf{x}}}_I = \mathbf{F}_I \tilde{\mathbf{x}}_I + \mathbf{G}_I \mathbf{n} \quad (25)$$

where

$$\mathbf{F}_I = \begin{bmatrix} 0 & 0 & -\hat{\mathbf{R}} & 0 \\ 0 & 0 & \mathbf{I}_3 & 0 \\ -[\hat{\mathbf{R}}(\mathbf{a}_m - \hat{\mathbf{b}}_a)]_\times & 0 & 0 & -\hat{\mathbf{R}} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad (26)$$

$$\mathbf{G}_I = \begin{bmatrix} -\hat{\mathbf{R}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -\hat{\mathbf{R}} & 0 & 0 \\ 0 & 0 & \mathbf{I} & 0 \\ 0 & 0 & 0 & \mathbf{I} \end{bmatrix}, \quad (27)$$

$$\mathbf{n} = \begin{bmatrix} \mathbf{n}_g \\ \mathbf{n}_a \\ \mathbf{n}_{gw} \\ \mathbf{n}_{aw} \end{bmatrix}, \quad (28)$$

with

$$\mathbf{E}(\mathbf{n}\mathbf{n}^\top) = \begin{bmatrix} \sigma_g^2 \mathbf{I}_3 & 0 & 0 & 0 \\ 0 & \sigma_a^2 \mathbf{I}_3 & 0 & 0 \\ 0 & 0 & \sigma_{gw}^2 \mathbf{I}_3 & 0 \\ 0 & 0 & 0 & \sigma_{aw}^2 \mathbf{I}_3 \end{bmatrix} \triangleq \mathbf{Q}. \quad (29)$$

3.3.1 Orientation

We proceed by computing $\dot{\mathbf{R}}$ by two different means (left and right developments)

$$\begin{aligned} \left(\text{Exp}(\tilde{\boldsymbol{\theta}}) \hat{\mathbf{R}} \right) &= \dot{\mathbf{R}} = \mathbf{R} [\boldsymbol{\omega}]_\times \\ \text{Exp}(\tilde{\boldsymbol{\theta}}) \hat{\mathbf{R}} + \text{Exp}(\tilde{\boldsymbol{\theta}}) \dot{\mathbf{R}} &= \left(\text{Exp}(\tilde{\boldsymbol{\theta}}) \hat{\mathbf{R}} \right) [\boldsymbol{\omega}]_\times \\ \text{Exp}(\tilde{\boldsymbol{\theta}}) \left[\dot{\tilde{\boldsymbol{\theta}}} \right]_\times \hat{\mathbf{R}} + \text{Exp}(\tilde{\boldsymbol{\theta}}) \hat{\mathbf{R}} [\boldsymbol{\omega}]_\times &= \text{Exp}(\tilde{\boldsymbol{\theta}}) \hat{\mathbf{R}} [\boldsymbol{\omega}]_\times \end{aligned} \quad (30)$$

Having $\boldsymbol{\omega} - \hat{\boldsymbol{\omega}} = -\tilde{\mathbf{b}}_g - \mathbf{n}_g$, this reduces to

$$\left[\dot{\tilde{\boldsymbol{\theta}}} \right]_\times \hat{\mathbf{R}} = \hat{\mathbf{R}} \left[-\tilde{\mathbf{b}}_g - \mathbf{n}_g \right]_\times \quad (31)$$

Right-multiplying left and right terms by $\hat{\mathbf{R}}^{-1}$, and recalling the adjoint operation on $\mathbf{SO}(3)$, i.e., (10), we have

$$\left[\dot{\tilde{\boldsymbol{\theta}}} \right]_\times = \hat{\mathbf{R}} \left[-\tilde{\mathbf{b}}_g - \mathbf{n}_g \right]_\times \hat{\mathbf{R}}^{-1} \quad (32)$$

$$= \left[-\hat{\mathbf{R}} \tilde{\mathbf{b}}_g - \hat{\mathbf{R}} \mathbf{n}_g \right]_\times \quad (33)$$

Finally, the error-state kinematic equation for orientation is given by

$$\dot{\tilde{\theta}} = -\hat{\mathbf{R}}\tilde{\mathbf{b}}_g - \hat{\mathbf{R}}\mathbf{n}_g \quad (34)$$

3.3.2 Position

$$\begin{aligned} (\hat{\mathbf{p}} + \tilde{\mathbf{p}}) &= \dot{\mathbf{p}} = \mathbf{v} \\ \hat{\mathbf{v}} + \dot{\tilde{\mathbf{p}}} &= \dot{\mathbf{v}} + \tilde{\mathbf{v}}. \end{aligned} \quad (35)$$

The position error-state kinematics is :

$$\dot{\tilde{\mathbf{p}}} = \tilde{\mathbf{v}} \quad (36)$$

3.3.3 Velocity

$$\begin{aligned} (\hat{\mathbf{v}} + \tilde{\mathbf{v}}) &= \dot{\mathbf{v}} = \mathbf{a} \\ \hat{\mathbf{a}} + \dot{\tilde{\mathbf{v}}} &= \dot{\mathbf{v}} = \mathbf{R}(\mathbf{a}_m - \mathbf{b}_a - \mathbf{n}_a) + \mathbf{g} \\ \hat{\mathbf{R}}(\mathbf{a}_m - \hat{\mathbf{b}}_a) + \mathbf{g} + \dot{\tilde{\mathbf{v}}} &= \dot{\mathbf{v}} = \text{Exp}(\tilde{\theta})\hat{\mathbf{R}}(\mathbf{a}_m - \mathbf{b}_a - \mathbf{n}_a) + \mathbf{g} \end{aligned} \quad (37)$$

Simplifying this equation yields the velocity error-state kinematics:

$$\dot{\tilde{\mathbf{v}}} = \text{Exp}(\tilde{\theta})\hat{\mathbf{R}}(\mathbf{a}_m - \mathbf{b}_a - \mathbf{n}_a) - \hat{\mathbf{R}}(\mathbf{a}_m - \hat{\mathbf{b}}_a) \quad (38a)$$

$$\simeq \left(\mathbf{I}_3 + [\tilde{\theta}]_{\times} \right) \hat{\mathbf{R}}(\mathbf{a}_m - \mathbf{b}_a - \mathbf{n}_a) - \hat{\mathbf{R}}(\mathbf{a}_m - \hat{\mathbf{b}}_a) \quad (38b)$$

$$= -\hat{\mathbf{R}}\tilde{\mathbf{b}}_a - \hat{\mathbf{R}}\mathbf{n}_a + [\tilde{\theta}]_{\times} \hat{\mathbf{R}}(\mathbf{a}_m - \mathbf{b}_a - \mathbf{n}_a) \quad (38c)$$

$$\simeq -[\hat{\mathbf{R}}(\mathbf{a}_m - \hat{\mathbf{b}}_a)]_{\times} \tilde{\theta} - \hat{\mathbf{R}}\tilde{\mathbf{b}}_a - \hat{\mathbf{R}}\mathbf{n}_a \quad (38d)$$

3.3.4 Bias

$$\dot{\tilde{\mathbf{b}}}_g = \dot{\mathbf{b}}_g - \dot{\hat{\mathbf{b}}}_g = \mathbf{n}_{gw} - \mathbf{0} = \mathbf{n}_{\text{gw}} \quad (39)$$

$$\dot{\tilde{\mathbf{b}}}_a = \dot{\mathbf{b}}_a - \dot{\hat{\mathbf{b}}}_a = \mathbf{n}_{aw} - \mathbf{0} = \mathbf{n}_{\text{aw}} \quad (40)$$

4 Visual-Inertial Navigation System

4.1 System model

The system state is defined as

$$\mathbf{x} = (\mathbf{x}_I, \boldsymbol{\ell}) \quad (41)$$

where $\boldsymbol{\ell} \in \mathbb{R}^3$ is the landmark position in the global frame.

The system kinematic equations are

$$\dot{\mathbf{R}} = \mathbf{R} [\boldsymbol{\omega}]_{\times} \quad (42a)$$

$$\dot{\mathbf{p}} = \mathbf{v} \quad (42b)$$

$$\dot{\mathbf{v}} = \mathbf{a} \quad (42c)$$

$$\dot{\mathbf{b}}_g = \mathbf{n}_{gw} \quad (42d)$$

$$\dot{\mathbf{b}}_a = \mathbf{n}_{aw} \quad (42e)$$

$$\dot{\boldsymbol{\ell}} = \mathbf{0} \quad (42f)$$

Let ${}^I\mathbf{p}_L$ denote the landmark position in the IMU frame, expressed as

$${}^I\mathbf{p}_L = \mathbf{R}^T(\boldsymbol{\ell} - \mathbf{p}). \quad (43)$$

As the camera explores the environment, the visual measurement of the landmark is available after the data association and rectification, formulated as

$$\mathbf{y} = \mathbf{h}({}^I\mathbf{p}_L) + \boldsymbol{\epsilon} \quad (44)$$

where $\mathbf{h} = \pi \circ \Upsilon$, $\Upsilon : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ transforms points from the IMU frame to the camera frame, and $\pi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ is the camera perspective projection function, $\boldsymbol{\epsilon} \in \mathbb{R}^2$ is the zero-mean Gaussian noise with $\mathbf{E}(\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T) = \mathbf{V}$.

4.2 Linearized error-state system

Let $\hat{\mathbf{x}}$ denote the VINS state estimate, which is propagated following (42) by setting the process noise to zero:

$$\dot{\hat{\mathbf{R}}} = \hat{\mathbf{R}} [\hat{\boldsymbol{\omega}}]_{\times} \quad (45a)$$

$$\dot{\hat{\mathbf{p}}} = \hat{\mathbf{v}} \quad (45b)$$

$$\dot{\hat{\mathbf{v}}} = \hat{\mathbf{a}} \quad (45c)$$

$$\dot{\hat{\mathbf{b}}}_g = \mathbf{0} \quad (45d)$$

$$\dot{\hat{\mathbf{b}}}_a = \mathbf{0} \quad (45e)$$

$$\dot{\hat{\boldsymbol{\ell}}} = \mathbf{0} \quad (45f)$$

The error-state of VINS is

$$\tilde{\mathbf{x}} = \mathbf{x} \ominus \hat{\mathbf{x}} = \begin{bmatrix} \mathbf{x}_I \ominus \hat{\mathbf{x}}_I \\ \boldsymbol{\ell} - \hat{\boldsymbol{\ell}} \end{bmatrix} \triangleq \begin{bmatrix} \tilde{\mathbf{x}}_I \\ \tilde{\boldsymbol{\ell}} \end{bmatrix} \quad (46)$$

According to (25), the VINS error-state kinematics can be written as

$$\dot{\tilde{\mathbf{x}}} = \mathbf{F}\tilde{\mathbf{x}} + \mathbf{G}\mathbf{n} \quad (47)$$

where

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_I & \mathbf{0}_{15 \times 3} \\ \mathbf{0}_{3 \times 15} & \mathbf{0}_{3 \times 3} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & -\hat{\mathbf{R}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_3 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -[\hat{\mathbf{R}}(\mathbf{a}_m - \hat{\mathbf{b}}_a)]_{\times} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -\hat{\mathbf{R}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (48)$$

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_I \\ \mathbf{0}_{3 \times 12} \end{bmatrix} = \begin{bmatrix} -\hat{\mathbf{R}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\hat{\mathbf{R}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (49)$$

The measurement residual is :

$$\tilde{\mathbf{y}} = \mathbf{y} - \mathbf{h}({}^I\hat{\mathbf{p}}_L) \quad (50a)$$

$$= \mathbf{h}({}^I\mathbf{p}_L) - \mathbf{h}({}^I\hat{\mathbf{p}}_L) + \epsilon \quad (50b)$$

$$= \frac{\partial \mathbf{h}}{\partial {}^I\mathbf{p}_L} d{}^I\mathbf{p}_L + \epsilon \quad (50c)$$

$$= \frac{\partial \mathbf{h}}{\partial {}^I\mathbf{p}_L} \left(\mathbf{R}^T (\boldsymbol{\ell} - \mathbf{p}) - \hat{\mathbf{R}}^T (\hat{\boldsymbol{\ell}} - \hat{\mathbf{p}}) \right) + \epsilon \quad (50d)$$

$$\simeq \frac{\partial \mathbf{h}}{\partial {}^I\mathbf{p}_L} \left(\hat{\mathbf{R}}^T (\mathbf{I}_3 - [\tilde{\boldsymbol{\theta}}]_{\times}) (\boldsymbol{\ell} - \mathbf{p}) - \hat{\mathbf{R}}^T (\hat{\boldsymbol{\ell}} - \hat{\mathbf{p}}) \right) + \epsilon \quad (50e)$$

$$= \frac{\partial \mathbf{h}}{\partial {}^I\mathbf{p}_L} \hat{\mathbf{R}}^T ([\boldsymbol{\ell} - \mathbf{p}]_{\times} \tilde{\boldsymbol{\theta}} + \tilde{\boldsymbol{\ell}} - \tilde{\mathbf{p}}) + \epsilon \quad (50f)$$

$$\simeq \frac{\partial \mathbf{h}}{\partial {}^I\mathbf{p}_L} \hat{\mathbf{R}}^T ([\hat{\boldsymbol{\ell}} - \hat{\mathbf{p}}]_{\times} \tilde{\boldsymbol{\theta}} + \tilde{\boldsymbol{\ell}} - \tilde{\mathbf{p}}) + \epsilon \quad (50g)$$

$$= \mathbf{H}\tilde{\mathbf{x}} + \epsilon \quad (50h)$$

where

$$\mathbf{H} = \mathbf{\Pi} \mathbf{H}_e \quad (51)$$

$$\mathbf{\Pi} = \frac{\partial \mathbf{h}}{\partial {}^I\mathbf{p}_L} \mathbf{R}^T \quad (52)$$

$$\mathbf{H}_e = \left[[\hat{\boldsymbol{\ell}} - \hat{\mathbf{p}}]_{\times} \quad -\mathbf{I}_3 \quad \mathbf{0} \quad \mathbf{0} \quad \mathbf{0} \quad \mathbf{I}_3 \right]. \quad (53)$$

We call \mathbf{H}_e the *essential measurement Jacobian*.

By combining (47) and (50), we obtain the linearized error-state system of VINS:

$$\begin{cases} \dot{\tilde{\mathbf{x}}} = \mathbf{F}\tilde{\mathbf{x}} + \mathbf{G}\mathbf{n} \\ \tilde{\mathbf{y}} = \mathbf{H}\tilde{\mathbf{x}} + \epsilon \end{cases} \quad (54)$$

5 Transformed Linearized Error-State System

5.1 Linear time-varying transformation on error-state

Denote the linear time-varying transformation by $\mathbf{T}(\hat{\mathbf{x}})$, where $\mathbf{T}(\cdot)$ is a 18×18 nonsingular matrix function and remains to be chosen. The transformed error-state is obtained by multiplying the transformation with the original error-state:

$$\tilde{\mathbf{x}}^* = \mathbf{T}(\hat{\mathbf{x}})\hat{\mathbf{x}}. \quad (55)$$

For simplicity, the variable $\hat{\mathbf{x}}$ of $\mathbf{T}(\hat{\mathbf{x}})$ is omitted in the following if there is no ambiguity. Taking the derivative of both sides of (55) with respect to time, we have

$$\dot{\tilde{\mathbf{x}}}^* = \dot{\mathbf{T}}\tilde{\mathbf{x}} + \mathbf{T}\dot{\tilde{\mathbf{x}}}. \quad (56)$$

Substituting (55) and (56) into the linearized error-state system (54) yields the transformed linearized error-state system as follows

$$\begin{cases} \dot{\tilde{\mathbf{x}}}^* = \mathbf{F}^*\tilde{\mathbf{x}}^* + \mathbf{G}^*\mathbf{n} \\ \tilde{\mathbf{y}} = \mathbf{H}^*\tilde{\mathbf{x}}^* + \epsilon \end{cases} \quad (57)$$

where

$$\mathbf{F}^* = \dot{\mathbf{T}}\mathbf{T}^{-1} + \mathbf{T}\mathbf{F}\mathbf{T}^{-1} \quad (58)$$

$$\mathbf{G}^* = \mathbf{T}\mathbf{G} \quad (59)$$

$$\mathbf{H}^* = \mathbf{H}\mathbf{T}^{-1}. \quad (60)$$

According to (51) and (60), the essential measurement Jacobian \mathbf{H}_e is transformed into

$$\mathbf{H}_e^* = \mathbf{H}_e\mathbf{T}^{-1}. \quad (61)$$

The transformation is designed as follows:

$$\mathbf{T} = \begin{bmatrix} \mathbf{I}_3 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ [\hat{\mathbf{p}}]_x & \mathbf{I}_3 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ [\hat{\mathbf{v}}]_x & \mathbf{0} & \mathbf{I}_3 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_3 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_3 & \mathbf{0} \\ [\hat{\boldsymbol{\epsilon}}]_x & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_3 \end{bmatrix} \quad (62)$$

Correspondingly, its inverse and derivative are

$$\mathbf{T}^{-1} = \begin{bmatrix} \mathbf{I}_3 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -[\hat{\mathbf{p}}]_{\times} & \mathbf{I}_3 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -[\hat{\mathbf{v}}]_{\times} & \mathbf{0} & \mathbf{I}_3 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_3 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_3 & \mathbf{0} \\ -[\hat{\boldsymbol{\theta}}]_{\times} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_3 \end{bmatrix} \quad (63)$$

$$\dot{\mathbf{T}} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ [\hat{\mathbf{v}}]_{\times} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ [\hat{\mathbf{a}}]_{\times} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}. \quad (64)$$

The transformed Jacobians are

$$\mathbf{F}^* = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & -\hat{\mathbf{R}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_3 & -[\hat{\mathbf{p}}]_{\times} \hat{\mathbf{R}} & \mathbf{0} & \mathbf{0} \\ [\mathbf{g}]_{\times} & \mathbf{0} & \mathbf{0} & -[\hat{\mathbf{v}}]_{\times} \hat{\mathbf{R}} & -\hat{\mathbf{R}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & -[\hat{\boldsymbol{\theta}}]_{\times} \hat{\mathbf{R}} & \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (65)$$

$$\mathbf{G}^* = \begin{bmatrix} -\hat{\mathbf{R}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -[\hat{\mathbf{p}}]_{\times} \hat{\mathbf{R}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -[\hat{\mathbf{v}}]_{\times} \hat{\mathbf{R}} & -\hat{\mathbf{R}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \\ -[\hat{\boldsymbol{\theta}}]_{\times} \hat{\mathbf{R}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (66)$$

$$\mathbf{H}^* = \Pi \mathbf{H}_v^* \quad (67)$$

$$\mathbf{H}_e^* = \begin{bmatrix} \mathbf{0} & -\mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \quad (68)$$

5.2 Observability of the transformed system

Observability refers to the ability of a system to recover its initial states using all available measurements. The set of states that cannot be recovered from measurements constitutes the unobservable subspace of the system. For the system described in (57), we can utilize the local observability matrix [2, P. 180] to conduct

the observability analysis. Let \mathbf{M}^* denote the local observability matrix of system (54), then

$$\mathbf{M}^* = \begin{bmatrix} \mathbf{M}_0^* \\ \mathbf{M}_1^* \\ \vdots \\ \mathbf{M}_{n-1}^* \end{bmatrix} \quad (69)$$

where $\mathbf{M}_0 = \mathbf{\Pi H}_e^*$ and

$$\mathbf{M}_{k+1}^* = \mathbf{M}_k^* \mathbf{F}^* + \dot{\mathbf{M}}_k^* \quad k = 0, 1, \dots, n-1. \quad (70)$$

According to (70), we can calculate

$$\mathbf{M}_1^* = \mathbf{\Pi H}_e^* \mathbf{F}^* + \mathbf{\Pi}^{(1)} \mathbf{H}_e^* \quad (71)$$

$$\mathbf{M}_2^* = \mathbf{\Pi H}_e^* \mathbf{F}^{*2} + \mathbf{\Pi}^{(1)} \mathbf{H}_e^* \mathbf{F}^* + \mathbf{\Pi}^{(1)} \mathbf{H}_e^* \mathbf{F}^* + \mathbf{\Pi H}_e^* \mathbf{F}^{*(1)} + \mathbf{\Pi}^{(2)} \mathbf{H}_e^* \quad (72a)$$

$$= \mathbf{\Pi}(\mathbf{H}_e^* \mathbf{F}^{*2} + \mathbf{H}_e^* \mathbf{F}^{*(1)}) + 2\mathbf{\Pi}^{(1)} \mathbf{H}_e^* \mathbf{F}^* + \mathbf{\Pi}^{(2)} \mathbf{H}_e^* \quad (72b)$$

The observability matrix is written as

$$\mathbf{M}^* = \begin{bmatrix} \mathbf{\Pi} & \mathbf{0} & \mathbf{0} \\ \mathbf{\Pi}^{(1)} & \mathbf{\Pi} & \mathbf{0} \\ \mathbf{\Pi}^{(2)} & 2\mathbf{\Pi}^{(1)} & \mathbf{\Pi} \end{bmatrix} \begin{bmatrix} \mathbf{H}_e^* \\ \mathbf{H}_e^* \mathbf{F}^* \\ \mathbf{H}_e^* \mathbf{F}^{*2} + \mathbf{H}_e^* \mathbf{F}^{*(1)} \end{bmatrix} \quad (73)$$

where

$$\begin{bmatrix} \mathbf{H}_e^* \\ \mathbf{H}_e^* \mathbf{F}^* \\ \mathbf{H}_e^* \mathbf{F}^{*2} + \mathbf{H}_e^* \mathbf{F}^{*(1)} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & -\mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} & -\mathbf{I} & \left[\hat{\mathbf{p}} - \hat{\boldsymbol{\ell}} \right] \times \hat{\mathbf{R}} & \mathbf{0} \\ -[\mathbf{g}]_\times & \mathbf{0} & \mathbf{0} & [\hat{\mathbf{v}}]_\times \hat{\mathbf{R}} + [\hat{\mathbf{v}}]_\times \hat{\mathbf{R}} + \left[\hat{\mathbf{p}} - \hat{\boldsymbol{\ell}} \right]_\times \hat{\mathbf{R}} [\hat{\omega}]_\times \hat{\mathbf{R}} & \mathbf{0} \end{bmatrix} \quad (74)$$

There exist four unobservable dimensions for the transformed linearized error-state system, s.t., $\mathbf{M}^* \mathbf{N}^* = \mathbf{0}$ with unobservable subspace

$$\mathbf{N}^* = \begin{bmatrix} \mathbf{0}_{3 \times 3} & \mathbf{g} \\ \mathbf{I}_3 & \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 1} \\ \mathbf{I}_3 & \mathbf{0}_{3 \times 1} \end{bmatrix}. \quad (75)$$

The unobservable subspace of the transformed system is independent of the state. When performing state estimation based on this transformed system, the erroneous reduction of unobservable dimensions is prevented, thus preserving consistency.

Remark 1. For $k > 2$, \mathbf{M}_k^* is complicated but does not affect our result about the observability. This is because only \mathbf{H}_e^* , $\mathbf{H}_e^* \mathbf{F}^*$, and $\mathbf{H}_e^* \mathbf{F}^{*2}$ contribute to observability analysis. If we continue to calculate \mathbf{M}_k^* when $k > 2$, the items such as $\mathbf{H}_e^* \mathbf{F}^{*m}$ and $\mathbf{H}_e^* \mathbf{F}^{*m} \mathbf{F}^{*(n)}$ will also be included in the left side of (74).

However, these terms only affect the blue-highlighted columns of (74). The nullspace blocks related to the blue-highlighted columns are always zeros, as shown in (75). Therefore, there is no need to calculate \mathbf{M}_k^* for $k > 2$.

6 Transformed Error-State Kalman Filter

In this section, we present the T-ESKF, a consistent VINS estimator based on the transformed linearized error-state system (57). The pipeline of T-ESKF is shown in Fig. 1.

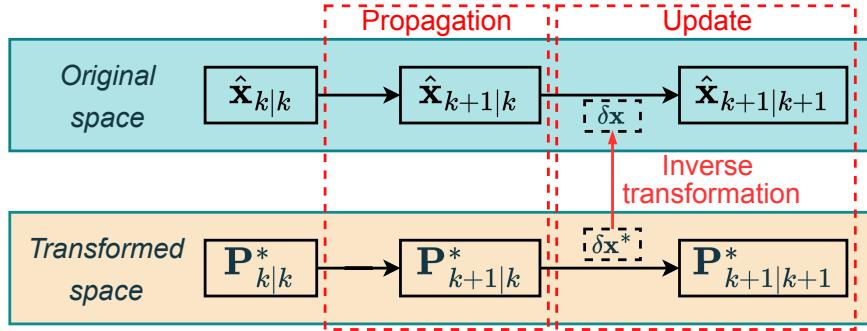


Figure 1: Pipeline of T-ESKF. It propagates and updates the covariance estimates in the transformed space. The state estimate is propagated by integrating (45) and updated using the state correction derived from the transformed system. The correction is obtained in the transformed space, and then transformed back to the original space.

The transformed error-state $\tilde{\mathbf{x}}^*$ matches \mathbf{P}^* and the error-state $\tilde{\mathbf{x}}$ matches \mathbf{P} . Since \mathbf{P} can be always obtained through the inverse transformation:

$$\mathbf{P} = \mathbf{T}(\hat{\mathbf{x}})^{-1} \mathbf{P} \mathbf{T}(\hat{\mathbf{x}})^{-\top} \quad (76)$$

and it is not involved in the estimation process, we do not maintain it in T-ESKF.

6.1 Propagation

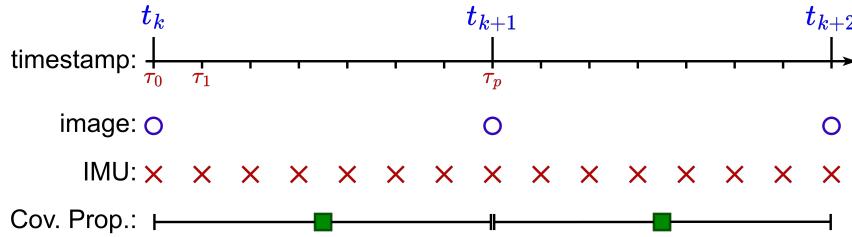


Figure 2: Different frequencies for camera and IMU measurements, along with covariance propagation. τ_0, τ_1, \dots , are the IMU measurement timestamp and t_k, t_{k+1}, \dots are the camera measurement timestamp. For simplify, τ_0 and τ_q coincide with t_k and t_{k+1} , respectively.

Usually, the frequency of IMU measurement is much higher than that of the camera. Before performing update using visual measurement at timestamp t_{k+1} , the covariance matrix need to be propagated from t_k to t_{k+1} , as follows:

$$\mathbf{P}_{k+1|k}^* = \Phi_k^* \mathbf{P}_{k|k}^* \Phi_k^{*\top} + \mathbf{Q}_k^* \quad (77)$$

where the transition matrix $\Phi_k^* \triangleq \Phi^*(\tau_p, \tau_0)$ is defined by the differential equation:

$$\frac{d}{d\tau} \Phi^*(\tau, \tau_0) = \mathbf{F}^* \Phi^*(\tau, \tau_0) \quad (78)$$

with the initial condition $\Phi^*(\tau_0, \tau_0) = \mathbf{I}_{15+3m}$. The accumulated noise matrix $\mathbf{Q}_k^* \triangleq \mathbf{Q}^*(\tau_p, \tau_0)$ is defined as

$$\mathbf{Q}_k^* = \int_{\tau_0}^{\tau_p} \Phi^*(\tau_p, \tau) \mathbf{G}_\tau^* \mathbf{Q} \mathbf{G}_\tau^{*\top} \Phi^*(\tau_p, \tau)^\top d\tau. \quad (79)$$

In practice, the transition matrix Φ_k^* and the accumulated noise \mathbf{Q}_k^* matrix are computed iteratively:

$$\Phi^*(\tau_{i+1}, \tau_0) = \Phi^*(\tau_{i+1}, \tau_i) \Phi^*(\tau_i, \tau_0), \quad i \in \{0, 1, \dots, p-1\} \quad (80)$$

$$\mathbf{Q}^*(\tau_{i+1}, \tau_0) = \Phi^*(\tau_{i+1}, \tau_i) \mathbf{Q}^*(\tau_i, \tau_0) \Phi^*(\tau_{i+1}, \tau_i)^\top + \mathbf{G}^*(\tau_{i+1}, \tau_i) \mathbf{Q}_d \mathbf{G}^*(\tau_{i+1}, \tau_i)^\top. \quad (81)$$

When IMU bias and a large number of landmarks are included in the state vector, T-ESKF will encounter the same computational bottleneck in covariance propagation as RI-EKF [3]. To solve this computational bottleneck, instead of directly integrating Φ_k^* and \mathbf{Q}_k^* using (80) and (81) directly, we develop an efficient propagation technique to compute them. Specifically, by leveraging (58) and (59), we can decompose Φ_k^* and \mathbf{Q}_k^* into following computationally efficient form (derivations please refer to Appendix B):

$$\Phi_k^* = \mathbf{T}(\hat{\mathbf{x}}_{k+1|k}) \begin{bmatrix} \Phi_k^I & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \mathbf{T}(\hat{\mathbf{x}}_{k|k})^{-1}, \quad (82)$$

$$\mathbf{Q}_k^* = \mathbf{T}(\hat{\mathbf{x}}_{k+1|k}) \begin{bmatrix} \mathbf{Q}_k^I & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{T}(\hat{\mathbf{x}}_{k+1|k})^\top. \quad (83)$$

where the IMU transition matrix $\Phi_k^I \triangleq \Phi^I(\tau_p, \tau_0) \in \mathbb{R}^{15 \times 15}$ and the IMU accumulated noise matrix $\mathbf{Q}_k^I \triangleq \mathbf{Q}^I(\tau_p, \tau_0) \in \mathbb{R}^{15 \times 15}$ are computed iteratively:

$$\Phi^I(\tau_{i+1}, \tau_0) = \Phi^I(\tau_{i+1}, \tau_i) \Phi^I(\tau_i, \tau_0), \quad i \in \{0, 1, \dots, p-1\} \quad (84)$$

$$\mathbf{Q}^I(\tau_{i+1}, \tau_0) = \Phi^I(\tau_{i+1}, \tau_i) \mathbf{Q}^I(\tau_i, \tau_0) \Phi^I(\tau_{i+1}, \tau_i)^\top + \mathbf{G}^I(\tau_{i+1}, \tau_i) \mathbf{Q}_d \mathbf{G}^I(\tau_{i+1}, \tau_i)^\top, \quad (85)$$

with

$$\Phi^I(\tau_{i+1}, \tau_i) = \begin{bmatrix} \mathbf{I}_3 & \mathbf{0} & \mathbf{0} & -\hat{\mathbf{R}}_{i+1}\mathbf{J}_r(\hat{\omega}_k\Delta t)\Delta t & \mathbf{0} & \mathbf{0} \\ [\hat{\mathbf{p}}_i - \hat{\mathbf{p}}_{i+1} + \hat{\mathbf{v}}_i\Delta t + \frac{1}{2}\mathbf{g}\Delta t^2]_\times & \mathbf{I}_3 & \mathbf{I}_3\Delta t & \hat{\mathbf{R}}_i\Xi_4 & -\hat{\mathbf{R}}_i\Xi_2 & \mathbf{0} \\ [\hat{\mathbf{v}}_i - \hat{\mathbf{v}}_{i+1} + \mathbf{g}\Delta t]_\times & \mathbf{0} & \mathbf{I}_3 & \hat{\mathbf{R}}_i\Xi_3 & -\hat{\mathbf{R}}_i\Xi_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_3 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_3 \end{bmatrix}, \quad (86)$$

$$\mathbf{G}^I(\tau_{i+1}, \tau_i) = \begin{bmatrix} -\hat{\mathbf{R}}_{i+1}\mathbf{J}_r(\hat{\omega}_k\Delta t)\Delta t & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hat{\mathbf{R}}_i\Xi_4 & -\hat{\mathbf{R}}_i\Xi_2 & \mathbf{0} & \mathbf{0} \\ \hat{\mathbf{R}}_i\Xi_3 & -\hat{\mathbf{R}}_i\Xi_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_3\Delta t & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_3\Delta t \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}. \quad (87)$$

The computations of Ξ_1 , Ξ_2 , Ξ_3 , and Ξ_4 are detailed in [Openvins: Analytical Integration Components](#).

Regardless of the number of landmarks included in the state vector, Φ_k^I and \mathbf{Q}_k^I are always small and fixed-size matrices. Although (84) and (85) need to be iterated p times, the matrices are much smaller than those in (80) and (81). Thus, the computation for Φ_k^I and \mathbf{Q}_k^I can be finished within a short timeframe. Additionally, the sparsity of $\mathbf{T}(\cdot)$ is also fully exploited to accelerate the matrix multiplications in (82) and (83). Consequently, the covariance can be propagated efficiently in T-ESKF.

Remark 2. (82) and (83) are derived in a continuous form and does not presuppose any assumptions regarding IMU model, such as the piecewise constant acceleration. Therefore, we can directly apply existing IMU integration theories, such as ACI² [4], to compute Φ_k^I and \mathbf{Q}_k^I .

6.2 Update

After propagation, we have the prior estimation $\hat{\mathbf{x}}_{k+1|k}$ and the covariance $\mathbf{P}_{k+1|k}^*$ corresponding to the transformed error-state. The covariance corresponding to the original error-state can be calculated through the inverse transformation:

$$\mathbf{P}_{k+1|k} = \mathbf{T}(\hat{\mathbf{x}}_{k+1|k})^{-1}\mathbf{P}_{k+1|k}^*\mathbf{T}(\hat{\mathbf{x}}_{k+1|k})^{-\top}. \quad (88)$$

During the update steps, we start with deriving the state correction from the transformed space, then inversely transform it back into the original space. Let $\delta\mathbf{x}^* \triangleq \mathbf{K}^*\tilde{\mathbf{y}}$ denote the Kalman state correction for the transformed error-state, the transformed error-state is corrected as

$$\tilde{\mathbf{x}}_{k+1|k+1}^* = \tilde{\mathbf{x}}_{k+1|k}^* - \delta\mathbf{x}^* \quad (89)$$

$$= \tilde{\mathbf{x}}_{k+1|k}^* - \mathbf{K}^*\tilde{\mathbf{y}} \quad (90)$$

$$= \tilde{\mathbf{x}}_{k+1|k}^* - \mathbf{K}^*(\mathbf{H}^*\tilde{\mathbf{x}}_{k+1|k}^* + \boldsymbol{\epsilon}) \quad (91)$$

$$= (\mathbf{I} - \mathbf{K}^*\mathbf{H}^*)\tilde{\mathbf{x}}_{k+1|k}^* - \mathbf{K}^*\boldsymbol{\epsilon} \quad (92)$$

with

$$\mathbf{P}_{k+1|k+1}^* = \mathbf{E} \left((\tilde{\mathbf{x}}_{k+1|k+1}^*) (\tilde{\mathbf{x}}_{k+1|k+1}^*)^\top \right) \quad (93)$$

$$= (\mathbf{I} - \mathbf{K}^* \mathbf{H}^*) \mathbf{P}_{k+1|k}^* (\mathbf{I} - \mathbf{K}^* \mathbf{H}^*)^\top + \mathbf{K}^* \mathbf{V} \mathbf{K}^{*\top}. \quad (94)$$

$$\mathbf{H}^* = \mathbf{\Pi}|_{\hat{\mathbf{x}}=\tilde{\mathbf{x}}_{k+1|k}} \mathbf{H}_e^* \quad (95)$$

Note that $\mathbf{P}_{k+1|k+1}^*$ is the covariance for $\tilde{\mathbf{x}}_{k+1|k+1}^* = \tilde{\mathbf{x}}_{k+1|k}^* - \delta \mathbf{x}^*$. The Kalman gain \mathbf{K}^* is obtained by minimizing the trace of $\mathbf{P}_{k+1|k+1}^*$, and it has the same form as the standard Kalman gain:

$$\mathbf{K}^* = \mathbf{P}_{k+1|k}^* \mathbf{H}_{k+1}^{*\top} (\mathbf{H}_{k+1}^* \mathbf{P}_{k+1|k}^* \mathbf{H}_{k+1}^{*\top} + \mathbf{V})^{-1}. \quad (96)$$

Substituting (96) into (94) yields

$$\mathbf{P}_{k+1|k+1}^* = \mathbf{P}_{k+1|k}^* - \mathbf{K}^* \mathbf{H}_{k+1}^* \mathbf{P}_{k+1|k}^*. \quad (97)$$

Subsequently, we derive the state correction in the original space. Let $\delta \mathbf{x}$ denote the state correction in the original space, then we have

$$\hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} \oplus \delta \mathbf{x}. \quad (98)$$

Correspondingly, the error-state is updated as

$$\tilde{\mathbf{x}}_{k+1|k+1} = \tilde{\mathbf{x}}_{k+1|k} - \delta \mathbf{x}. \quad (99)$$

The covariance for $\tilde{\mathbf{x}}_{k+1|k+1}$ is computed by the inverse transformation:

$$\mathbf{P}_{k+1|k+1} = \mathbf{T}(\hat{\mathbf{x}}_{k+1|k+1})^{-1} \mathbf{P}_{k+1|k+1}^* \mathbf{T}(\hat{\mathbf{x}}_{k+1|k+1})^{-\top} \quad (100)$$

To ensure that $\tilde{\mathbf{x}}_{k+1|k+1}$ in (99) matches $\mathbf{P}_{k+1|k+1}$ in (100), $\delta \mathbf{x}$ must satisfy

$$\tilde{\mathbf{x}}_{k+1|k} - \delta \mathbf{x} = \mathbf{T}(\hat{\mathbf{x}}_{k+1|k+1})^{-1} (\tilde{\mathbf{x}}_{k+1|k}^* - \delta \mathbf{x}^*). \quad (101)$$

Then we have

$$\delta \mathbf{x} = \left(\mathbf{I} - \mathbf{T}(\hat{\mathbf{x}}_{k+1|k+1})^{-1} \mathbf{T}(\hat{\mathbf{x}}_{k+1|k}) \right) \tilde{\mathbf{x}}_{k+1|k} + \mathbf{T}(\hat{\mathbf{x}}_{k+1|k+1})^{-1} \delta \mathbf{x}^*. \quad (102)$$

Note that $(\mathbf{I} - \mathbf{T}(\hat{\mathbf{x}}_{k+1|k+1})^{-1} \mathbf{T}(\hat{\mathbf{x}}_{k+1|k})) \approx \mathbf{0}$, thus (102) can be rewritten as

$$\delta \mathbf{x} = \mathbf{T}(\hat{\mathbf{x}}_{k+1|k+1})^{-1} \delta \mathbf{x}^*. \quad (103)$$

The detailed derivation of the updating using (103) is provided in Appendix C.1, which outlines the calculation of $\hat{\mathbf{x}}_{k+1|k+1}$ through an analytical solution.

Another updating method is using $\hat{\mathbf{x}}_{k+1|k}$ instead of $\hat{\mathbf{x}}_{k+1|k+1}$ in (103), that is

$$\delta \mathbf{x} = \mathbf{T}(\hat{\mathbf{x}}_{k+1|k})^{-1} \delta \mathbf{x}^*. \quad (104)$$

We also provide the derivation of updating with (104) in Appendix C.2.

We verified these two update equations. The results indicate that updating through (103) or (104) yields consistent performance, both showing nearly identical outcomes, despite being evaluated at different estimates. In the manuscript, we adopt the updating equation in (104) as it is easy to follow.

6.3 T-ESKF properties

We now show that T-ESKF has the optimal Jacobians and correct observability. Since the Jacobians in T-ESKF, Φ^* and \mathbf{H}^* , are evaluated at the current best estimates, the optimality of Jacobians is automatically preserved. The observability matrix for T-ESKF is [5]:

$$\mathbf{M}_{\text{T-ESKF}} = \begin{bmatrix} \mathbf{H}_0^* \\ \mathbf{H}_1^* \Phi^*(t_1, t_0) \\ \vdots \\ \mathbf{H}_k^* \Phi^*(t_k, t_0) \end{bmatrix} \quad (105)$$

According (95) and (82), we have

$$\mathbf{H}_i^* \Phi^*(t_i, t_0) = \Pi_{\hat{\mathbf{x}}=\hat{\mathbf{x}}_{i|i+1}} \begin{bmatrix} -(t_i - t_0)^2 [\mathbf{g}]_\times & -\mathbf{I}_3 & -\mathbf{I}_3(t_i - t_0) & \boldsymbol{\Gamma}_{i,1} & \boldsymbol{\Gamma}_{i,2} & \mathbf{I}_3 \end{bmatrix} \quad (106)$$

where $\boldsymbol{\Gamma}_{i,1}$ and $\boldsymbol{\Gamma}_{i,2} \in \mathbb{R}^{3 \times 3}$. We can directly find that the unobservable subspace for T-ESKF with $\mathbf{M}_{\text{T-ESKF}} \mathbf{N}_{\text{T-ESKF}} = \mathbf{0}$ is

$$\mathbf{N}_{\text{T-ESKF}} = \begin{bmatrix} \mathbf{0}_{3 \times 3} & \mathbf{g} \\ \mathbf{I}_3 & \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 1} \\ \mathbf{I}_3 & \mathbf{0}_{3 \times 1} \end{bmatrix}. \quad (107)$$

(75) and (107) indicate that erroneous reductions of unobservable dimensions are prevented and T-ESKF has the correct observability. Consequently, T-ESKF does not encounter estimation inconsistencies arising from observability mismatches.

7 Real-World Experiments

In addition to the dataset experiments, we further compare T-ESKF with T-ESKF with ESKF [1], FEJ-ESKF [6], and RI-EKF [7] using our customized sensor platform mounted on an aerial robot, as depicted in Figure 3. This platform provides stereo images at 30Hz with a resolution of 848×480 and IMU data at 200Hz, with the main parameters outlined in Table 1. The camera intrinsic and extrinsic parameters are calibrated using the offline calibration toolbox Kalibr. For robust performance, camera intrinsics and time offset calibrations are enabled.

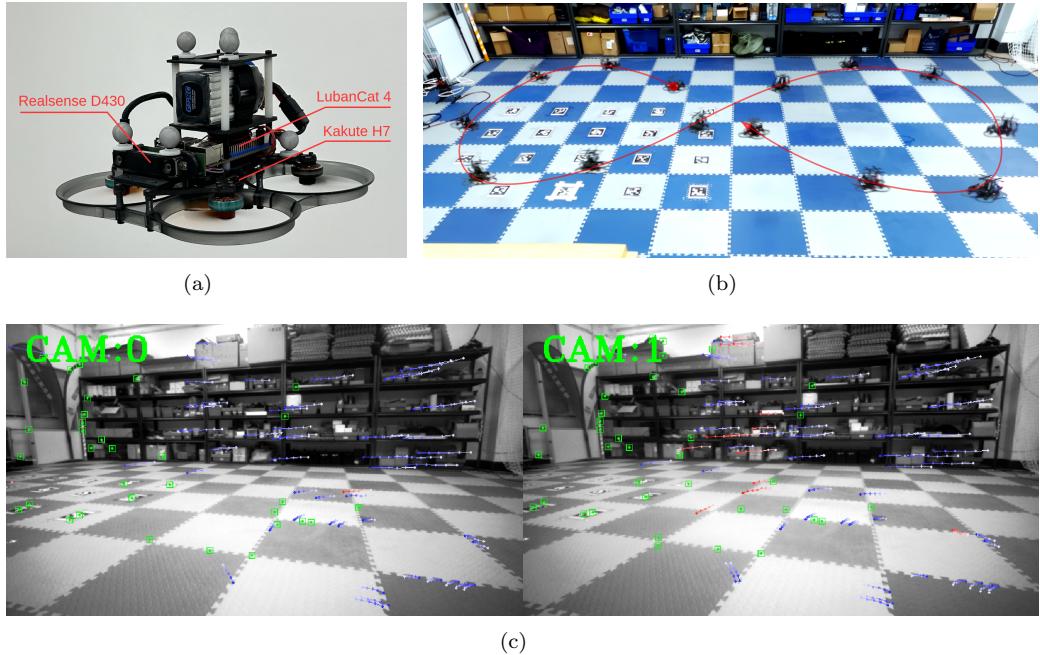


Figure 3: (a) Aerial robot with a Realsense D430 stereo camera, a Kakute H7 flight controller (MPU6000), and an onboard computer, LubanCat 4. (b) Figure-eight flight trajectory with unfixed yaw. (c) A sample frame with tracked features in the experiment.

Table 1: Real-world experiment configuration

Parameter	Value	Parameter	Value
Accel. White Noise	6.33e-03	Gyro White Noise	8.71e-04
Accel. Random Walk	2.87e-03	Gyro Random Walk	3.34e-05
Pixel Noise	1	IMU Freq.	200
Max Cam Pts/Frame	200	Cam Freq.	30
Max Feats/Frame	60	Cam Number	Stereo
Max Clone Size	11	Feat. Rep.	Global XYZ

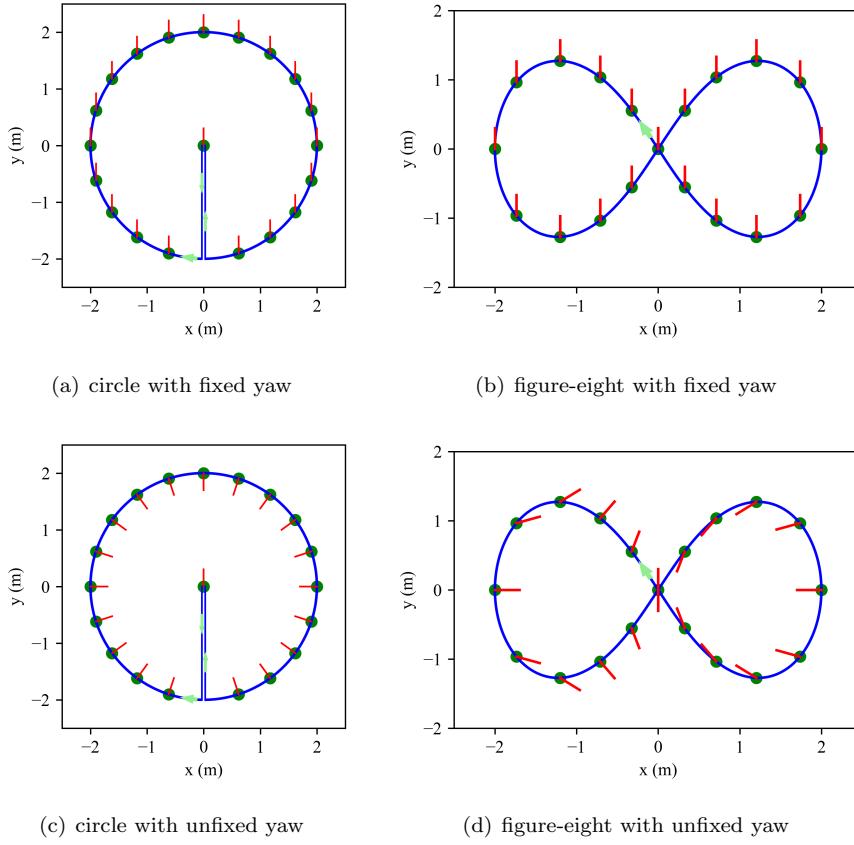


Figure 4: Designed trajectories in real-world experiments. The circle or the figure-eight pattern is repeated 6 times in each trajectory. The red lines represent the orientation of the aerial robot. The lightgreen arrows indicate the direction of the trajectories.

The aerial robot is commanded to follow either a circular or a figure-eight trajectory with fixed or unfixed yaw, as illustrated in Figure 4. The time taken to track a circle is 6.28 seconds, and for a figure-eight, it is 10.47 seconds. Each trajectory includes six repetitions of the circular or figure-eight patterns. Three datasets are gathered for each scenario, consisting of IMU and stereo camera measurements, in addition to groundtruth acquired from the motion capture system.

To evaluate the estimated results, we align the estimated trajectories and the ground truth based on the initial frame, as depicted in Figure 5. The RMSE of these estimators is detailed in Table 2. As seen, the performance of these estimators is close when the yaw is fixed. Nonetheless, T-ESKF surpasses ESKF, particularly in orientation estimation. In the case of unfixed yaw, T-ESKF exhibits superior performance compared to ESKF and FEJ-ESKF.

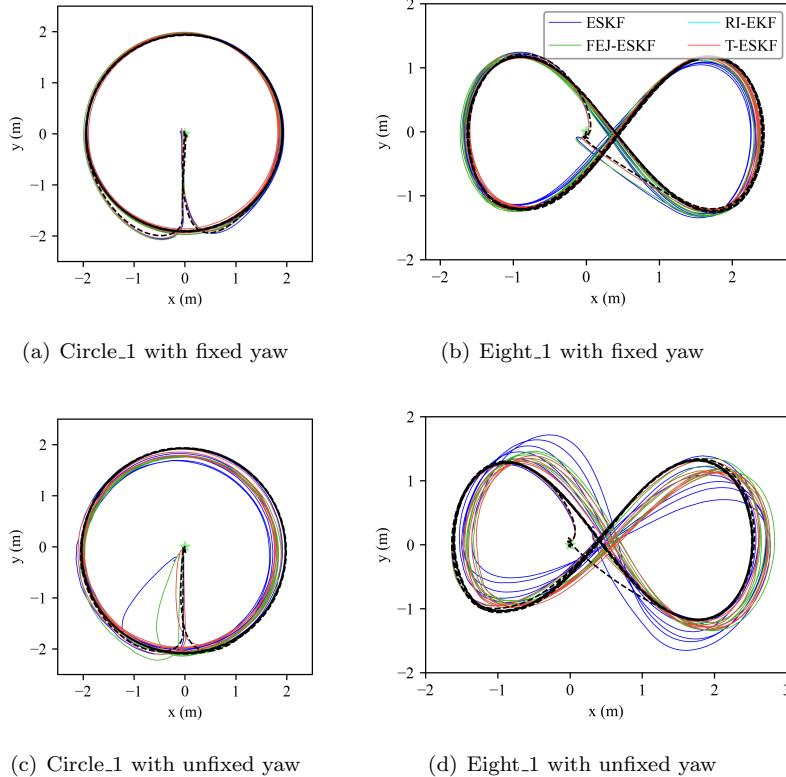


Figure 5: Estimated trajectories on the first dataset in each case. The black dashed line is the groundtruth trajectory with a green star marking the starting point. The estimated trajectories are aligned to the groundtruth trajectory (black dashed line) using the beginning frame.

Table 2: Average orientation (deg) and position (meter) RMSE of the real-world experiments

	Set	ESKF	FEJ-ESKF	RI-EKF	T-ESKF
Fixed yaw	Circle_1	1.432 / 0.066	1.261 / 0.054	1.270 / 0.061	1.270 / 0.061
	Circle_2	1.236 / 0.057	1.492 / 0.059	1.114 / 0.047	1.113 / 0.047
	Circle_3	5.309 / 0.165	1.469 / 0.076	1.216 / 0.068	1.198 / 0.075
	Eight_1	2.333 / 0.114	1.017 / 0.085	0.878 / 0.058	0.878 / 0.058
	Eight_2	3.882 / 0.142	1.884 / 0.133	1.539 / 0.135	1.434 / 0.126
	Eight_3	2.885 / 0.105	1.373 / 0.100	1.148 / 0.096	0.940 / 0.103
	Average	2.846 / 0.108	1.416 / 0.084	1.194 / 0.078	1.139 / 0.078
Unfixed yaw	Circle_1	26.617 / 0.665	11.655 / 0.402	2.734 / 0.128	2.747 / 0.128
	Circle_2	25.674 / 0.727	2.334 / 0.104	2.180 / 0.137	2.180 / 0.137
	Circle_3	25.023 / 0.689	2.056 / 0.134	2.633 / 0.124	1.555 / 0.125
	Eight_1	12.833 / 0.366	4.712 / 0.245	3.702 / 0.216	3.716 / 0.216
	Eight_2	12.447 / 0.397	6.360 / 0.349	6.456 / 0.277	6.138 / 0.273
	Eight_3	15.979 / 0.440	6.534 / 0.355	6.440 / 0.368	6.438 / 0.368
	average	19.762 / 0.547	5.609 / 0.265	4.024 / 0.208	3.796 / 0.208

A Derivation for (28) in the manuscript

To make the transformed Jacobians \mathbf{F}^* and \mathbf{H}_e^* independent of states, an intuitive idea is to transform these green-highlighted sub-matrices exclusively while leaving the remaining sub-matrices unchanged. Thus, the transformation can be designed as a lower triangular matrix with the diagonal being the identity:

$$\mathbf{T} = \begin{bmatrix} \mathbf{I}_3 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{T}_p & \mathbf{I}_3 & \mathbf{0} & \mathbf{0} \\ \mathbf{T}_v & \mathbf{0} & \mathbf{I}_3 & \mathbf{0} \\ \mathbf{T}_\epsilon & \mathbf{0} & \mathbf{0} & \mathbf{I}_3 \end{bmatrix}, \quad (108)$$

where $\mathbf{T}_p, \mathbf{T}_v, \mathbf{T}_\epsilon$ are sub-matrices to be determined such that \mathbf{F}^* and \mathbf{H}_e^* are independent of the states.

$$\mathbf{F}^* = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \dot{\mathbf{T}}_p - \mathbf{T}_v & \mathbf{0} & \mathbf{I}_3 \\ \dot{\mathbf{T}}_v - [\hat{\mathbf{R}}\mathbf{a}_m]_\times & \mathbf{0} & \mathbf{0} \\ \dot{\mathbf{T}}_\epsilon & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad (109)$$

$$\mathbf{H}_e^* = \left[\begin{bmatrix} \hat{\mathbf{t}} - \hat{\mathbf{p}} \end{bmatrix}_\times + \mathbf{T}_p - \mathbf{T}_\epsilon \right] \begin{bmatrix} -\mathbf{I}_3 & \mathbf{0} & \mathbf{I}_3 \end{bmatrix}. \quad (110)$$

To ensure \mathbf{F}^* and \mathbf{H}_e^* independent of states, the green-highlighted blocks in (109) and (110) are required to be constant, i.e.,

$$\mathbf{C}_1 = \dot{\mathbf{T}}_p - \mathbf{T}_v \quad (111)$$

$$\mathbf{C}_2 = \dot{\mathbf{T}}_v - [\hat{\mathbf{R}}\mathbf{a}_m]_\times \quad (112)$$

$$\mathbf{C}_3 = \dot{\mathbf{T}}_\epsilon \quad (113)$$

$$\mathbf{C}_4 = \left[\hat{\mathbf{t}} - \hat{\mathbf{p}} \right]_\times + \mathbf{T}_p - \mathbf{T}_\epsilon \quad (114)$$

Note that $\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3$, and \mathbf{C}_4 are constant matrices. Taking the derivative of (111) and (113) and the second order derivative of (114), we have

$$\mathbf{0} = \ddot{\mathbf{T}}_p - \dot{\mathbf{T}}_v \quad (115)$$

$$\mathbf{0} = \ddot{\mathbf{T}}_\epsilon \quad (116)$$

$$\mathbf{0} = -[\hat{\mathbf{a}}]_\times + \ddot{\mathbf{T}}_p - \ddot{\mathbf{T}}_\epsilon \quad (117)$$

By combining these equations with (112)+(115)-(116)-(117), we get \mathbf{C}_2 :

$$\mathbf{C}_2 = \left[\hat{\mathbf{a}} - \hat{\mathbf{R}}\mathbf{a}_m \right]_\times = [\mathbf{g}]_\times. \quad (118)$$

Back substituting (118) into (111) - (114), we can obtain the solution as follows:

$$\begin{cases} \mathbf{T}_p = [\hat{\mathbf{p}}]_{\times} + t\mathbf{C}_3 + \mathbf{C}_5 \\ \mathbf{T}_v = [\hat{\mathbf{v}}]_{\times} + \mathbf{C}_3 - \mathbf{C}_1 \\ \mathbf{T}_\epsilon = [\hat{\boldsymbol{\epsilon}}]_{\times} + t\mathbf{C}_3 - \mathbf{C}_4 + \mathbf{C}_5 \end{cases} \quad (119)$$

where \mathbf{C}_5 is an integral constant. Among these constant sub-matrices, $\mathbf{C}_1, \mathbf{C}_3, \mathbf{C}_4, \mathbf{C}_5$ can be arbitrarily chosen. To simplify the transformation matrix, we choose

$$\mathbf{C}_1 = \mathbf{0}, \quad \mathbf{C}_3 = \mathbf{0}, \quad \mathbf{C}_4 = \mathbf{0}, \quad \mathbf{C}_5 = \mathbf{0}. \quad (120)$$

Substituting (120) into (119), we get:

$$\mathbf{T}_p = [\hat{\mathbf{p}}]_{\times}, \quad \mathbf{T}_v = [\hat{\mathbf{v}}]_{\times}, \quad \mathbf{T}_\epsilon = [\hat{\boldsymbol{\epsilon}}]_{\times}. \quad (121)$$

B Derivation for (82) and (83)

Let Φ_k and \mathbf{Q}_k be the transition matrix and accumulated noise matrix of ESKF. According to the definitions of Φ and Φ^* and the relationship between \mathbf{F} and \mathbf{F}^* , i.e.,

$$\begin{cases} \frac{d}{d\tau} \Phi(\tau, \tau_0) = \mathbf{F}_\tau \Phi(\tau, \tau_0) \\ \frac{d}{d\tau} \Phi^*(\tau, \tau_0) = \mathbf{F}_\tau^* \Phi^*(\tau, \tau_0) \\ \mathbf{F}_\tau^* = \dot{\mathbf{T}}_\tau \mathbf{T}_\tau^{-1} + \mathbf{T}_\tau \mathbf{F}_\tau \mathbf{T}_\tau^{-1} \end{cases} \quad (122)$$

we can verify that

$$\frac{d}{d\tau} (\mathbf{T}_\tau \Phi(\tau, \tau_0) \mathbf{T}_{\tau_0}^{-1}) = \dot{\mathbf{T}}_\tau \Phi(\tau, \tau_0) \mathbf{T}_\tau^{-1} + \mathbf{T}_\tau \mathbf{F}_\tau \Phi(\tau, \tau_0) \mathbf{T}_{\tau_0}^{-1} \quad (123a)$$

$$= (\dot{\mathbf{T}}_\tau \mathbf{T}_\tau^{-1} + \mathbf{T}_\tau \mathbf{F}_\tau \mathbf{T}_\tau^{-1}) \mathbf{T}_\tau \Phi(\tau, \tau_0) \mathbf{T}_{\tau_0}^{-1} \quad (123b)$$

$$= \mathbf{F}_\tau^* \Phi^*(\tau, \tau_0) \quad (123c)$$

$$= \frac{d}{d\tau} \Phi^*(\tau, \tau_0). \quad (123d)$$

According to (123) and the initial condition $\Phi^*(\tau_0, \tau_0) = \mathbf{T}_{\tau_0} \Phi(\tau_0, \tau_0) \mathbf{T}_{\tau_0}^{-1} = \mathbf{I}_{15+3m}$, we have:

$$\Phi^*(\tau, \tau_0) = \mathbf{T}_\tau \Phi(\tau, \tau_0) \mathbf{T}_{\tau_0}^{-1} \quad (124)$$

Similarly, from

$$\begin{cases} \mathbf{Q}_k = \int_{\tau_0}^{\tau_p} \Phi(\tau_p, \tau) \mathbf{G}_\tau \mathbf{Q} \mathbf{G}_\tau^\top \Phi(\tau_p, \tau)^\top d\tau \\ \mathbf{Q}_k^* = \int_{\tau_0}^{\tau_p} \Phi^*(\tau_p, \tau) \mathbf{G}_\tau^* \mathbf{Q} \mathbf{G}_\tau^{*\top} \Phi^*(\tau_p, \tau)^\top d\tau \\ \mathbf{G}_\tau^* = \mathbf{T}_\tau \mathbf{G}_\tau \end{cases} \quad (125)$$

we can imply that

$$\mathbf{Q}_k^* = \int_{\tau_0}^{\tau_p} \Phi^*(\tau_p, \tau) \mathbf{G}_\tau^* \mathbf{Q} \mathbf{G}_\tau^{*\top} \Phi^*(\tau_p, \tau)^\top d\tau \quad (126a)$$

$$= \int_{\tau_0}^{\tau_p} \mathbf{T}_{\tau_p} \Phi(\tau_p, \tau) \mathbf{T}_\tau^{-1} \mathbf{T}_\tau \mathbf{G}_\tau \mathbf{Q} \mathbf{G}_\tau^\top \mathbf{T}_\tau^\top \mathbf{T}_\tau^{-\top} \Phi(\tau_p, \tau)^\top \mathbf{T}_{\tau_p}^\top d\tau \quad (126b)$$

$$= \mathbf{T}_{\tau_p} \left(\int_{\tau_0}^{\tau_p} \Phi(\tau_p, \tau) \mathbf{G}_\tau \mathbf{Q} \mathbf{G}_\tau^\top \Phi(\tau_p, \tau)^\top d\tau \right) \mathbf{T}_{\tau_p}^\top \quad (126c)$$

$$= \mathbf{T}_{\tau_p} \mathbf{Q}_k \mathbf{T}_{\tau_p}^\top \quad (126d)$$

Moreover, Φ_k and \mathbf{Q}_k are diagonal block matrices,

$$\Phi_k = \begin{bmatrix} \Phi_k^I & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{3m} \end{bmatrix} \quad (127)$$

$$\mathbf{Q}_k = \begin{bmatrix} \mathbf{Q}_k^I & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{3m \times 3m} \end{bmatrix} \quad (128)$$

where $\mathbf{Q}_k^I \in \mathbb{R}^{15 \times 15}$ and $\Phi_k^I \in \mathbb{R}^{15 \times 15}$ the IMU transition matrix and the IMU accumulated noise matrix. Substituting above two equations to (124) and (126) and replacing $(\mathbf{T}_{t_0}, \mathbf{T}_{t_p})$ with $\mathbf{T}(\hat{\mathbf{x}}_{k|k}, \mathbf{T}(\hat{\mathbf{x}}_{k+1|k}))$ yield (82) and (83).

C T-ESKF state update equation

C.1 Update with (104): $\hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} \oplus (\mathbf{T}(\hat{\mathbf{x}}_{k+1|k})^{-1} \delta \mathbf{x}^*)$

$$\hat{\mathbf{R}}_{k+1|k+1} = \text{Exp}(\delta \boldsymbol{\theta}^*) \hat{\mathbf{R}}_{k+1|k} \quad (129)$$

$$\hat{\mathbf{p}}_{k+1|k+1} = (\mathbf{I}_3 + [\delta \boldsymbol{\theta}^*]_\times) \hat{\mathbf{p}}_{k+1|k} + \delta \mathbf{p}^* \quad (130)$$

$$\hat{\mathbf{v}}_{k+1|k+1} = (\mathbf{I}_3 + [\delta \boldsymbol{\theta}^*]_\times) \hat{\mathbf{v}}_{k+1|k} + \delta \mathbf{v}^* \quad (131)$$

$$\hat{\mathbf{b}}_{g,k+1|k+1} = \hat{\mathbf{b}}_{g,k+1|k} + \delta \mathbf{b}_g^* \quad (132)$$

$$\hat{\mathbf{b}}_{a,k+1|k+1} = \hat{\mathbf{b}}_{a,k+1|k} + \delta \mathbf{b}_a^* \quad (133)$$

$$\hat{\boldsymbol{\ell}}_{k+1|k+1} = (\mathbf{I}_3 + [\delta \boldsymbol{\theta}^*]_\times) \hat{\boldsymbol{\ell}}_{k+1|k} + \delta \boldsymbol{\ell}^*. \quad (134)$$

C.2 Update with (103): $\hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} \oplus (\mathbf{T}(\hat{\mathbf{x}}_{k+1|k+1})^{-1} \delta \mathbf{x}^*)$

$$\hat{\mathbf{R}}_{k+1|k+1} = \text{Exp}(\delta \boldsymbol{\theta}^*) \hat{\mathbf{R}}_{k+1|k} \quad (135)$$

$$\hat{\mathbf{p}}_{k+1|k+1} = \mathbf{A}(\hat{\mathbf{p}}_{k+1|k} + \delta \mathbf{p}^*) \quad (136)$$

$$\hat{\mathbf{v}}_{k+1|k+1} = \mathbf{A}(\hat{\mathbf{v}}_{k+1|k} + \delta \mathbf{v}^*) \quad (137)$$

$$\hat{\mathbf{b}}_{g,k+1|k+1} = \hat{\mathbf{b}}_{g,k+1|k} + \delta \mathbf{b}_g^* \quad (138)$$

$$\hat{\mathbf{b}}_{a,k+1|k+1} = \hat{\mathbf{b}}_{a,k+1|k} + \delta \mathbf{b}_a^* \quad (139)$$

$$\hat{\boldsymbol{\theta}}_{k+1|k+1} = \mathbf{A}(\hat{\boldsymbol{\theta}}_{k+1|k} + \delta \boldsymbol{\theta}^*) \quad (140)$$

with

$$\mathbf{A} = \frac{\mathbf{I}_3 + [\delta \boldsymbol{\theta}^*]_\times + \delta \boldsymbol{\theta}^* \delta \boldsymbol{\theta}^{*\top}}{1 + \delta \boldsymbol{\theta}^{*\top} \delta \boldsymbol{\theta}^*}. \quad (141)$$

References

- [1] J. Sola, “Quaternion kinematics for the error-state kalman filter,” *arXiv preprint arXiv:1711.02508*, 2017.
- [2] C.-T. Chen, *Linear System Theory and Design*. The Oxford Series in Electrical and Computer Engineering, New York Oxford: Oxford University Press, 3. ed., 3. print ed., 1999.
- [3] 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, pp. 1627–1634, Apr. 2022.
- [4] Y. Yang, B. P. Wisely Babu, C. Chen, G. Huang, and L. Ren, “Analytic combined imu integration (aci2) for visual inertial navigation,” in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 4680–4686, 2020.
- [5] Z. Chen, K. Jiang, and J. Hung, “Local observability matrix and its application to observability analyses,” in *[Proceedings] IECON '90: 16th Annual Conference of IEEE Industrial Electronics Society*, pp. 100–103 vol.1, 1990.
- [6] G. P. Huang, A. I. Mourikis, and S. I. Roumeliotis, “Analysis and improvement of the consistency of extended Kalman filter based SLAM,” in *2008 IEEE International Conference on Robotics and Automation*, pp. 473–479, May 2008.
- [7] T. Zhang, K. Wu, J. Song, S. Huang, and G. Dissanayake, “Convergence and consistency analysis for a 3-D Invariant-EKF SLAM,” *IEEE Robotics and Automation Letters*, vol. 2, pp. 733–740, Apr. 2017.