

Supplementary Materials:

DIDO: Deep Inertial Quadrotor Dynamical Odometry

Kunyi Zhang^{1,2} - kunyizhang@zju.edu.cn
Chenxing Jiang^{1,2} - chenxingjiang@zju.edu.cn
Jinghang Li² - jhanglee.ro@gmail.com
Sheng Yang³ - richeng.ys@alibaba-inc.com
Teng Ma³ - damon.mt@alibaba-inc.com
Chao Xu^{1,2} - chaoxu@zju.edu.cn
Fei Gao^{1,2} - fgaoaa@zju.edu.cn

CONTENTS

Acknowledgement	1
I IMU Kinematics and Quadrotor Dynamics	2
I-A IMU Kinematics	2
I-B Quadrotor Dynamics	2
II Two-Stage EKF for Inertial Dynamical Fusion	2
II-A Rotation Stage	2
II-A.1 State	2
II-A.2 Process Model	3
II-A.3 Measurement Model	3
II-A.4 Extended Kalman Filter	3
II-B Translation Stage	3
II-B.1 State	3
II-B.2 Process Model	4
II-B.3 Measurement Models	4
II-B.4 Discrete Extended Kalman Filter	6
III One-Stage EKF for Inertial Dynamical Fusion	6
III-A State	6
III-B Process Model	6
III-C Measurement Model	7
III-C.1 gravity alignment measurements	7
III-C.2 dynamics constraint measurements	7
III-C.3 network velocity measurements	8
III-C.4 network displacement measurements	8
III-D Discrete Extended Kalman Filter	9
IV Observability Analysis	9

¹State Key Laboratory of Industrial Control Technology, Institute of Cyber-Systems and Control, Zhejiang University, Hangzhou 310027, China.

²Huzhou Institute, Zhejiang University, Huzhou 313000, China.

³Alibaba DAMO Academy Autonomous Driving Lab, Hangzhou 311121, China.

I. IMU KINEMATICS AND QUADROTOR DYNAMICS

A. IMU Kinematics

IMU measurements include the gyroscope $\tilde{\boldsymbol{\omega}}$ and non-gravitational acceleration $\tilde{\boldsymbol{a}}$, which are measured in the IMU frame (the \mathcal{I} frame) and given by:

$$\begin{aligned}\mathcal{I}\tilde{\boldsymbol{\omega}} &= \mathcal{I}\boldsymbol{\omega} + \mathcal{I}\mathbf{b}_\omega + \mathbf{n}_\omega, \\ \mathcal{I}\tilde{\boldsymbol{a}} &= \mathcal{I}\mathbf{a} + \mathcal{I}\mathbf{b}_a + \frac{\mathcal{G}}{\mathcal{I}}\mathbf{R}^\top \mathbf{g} + \mathbf{n}_a,\end{aligned}\quad (1)$$

where $\mathcal{I}\boldsymbol{\omega}$ and $\mathcal{I}\mathbf{a}$ are the true angular velocity and acceleration, $\mathcal{G}\mathbf{g} = [0, 0, 9.8]$ is the gravity vector in the gravity-aligned frame (the \mathcal{G} frame), $\frac{\mathcal{G}}{\mathcal{I}}\mathbf{R}$ is the rotation matrix from the \mathcal{I} frame to the \mathcal{G} frame, \mathbf{n}_ω and \mathbf{n}_a are the additive Gaussian white noise in gyroscope and acceleration measurements, \mathbf{b}_ω and \mathbf{b}_a are the bias of IMU modeled as random walk:

$$\begin{aligned}\mathbf{n}_\omega &\sim \mathcal{N}(0, \Sigma_\omega^2), \quad \dot{\mathbf{b}}_\omega \sim \mathcal{N}(0, \Sigma_{b_\omega}^2), \\ \mathbf{n}_a &\sim \mathcal{N}(0, \Sigma_a^2), \quad \dot{\mathbf{b}}_a \sim \mathcal{N}(0, \Sigma_{b_a}^2).\end{aligned}\quad (2)$$

B. Quadrotor Dynamics

Because the propagation of kinematic states is driven by multiple propulsion units in a quadrotor system, we model the Newtonian dynamics according to [1]. The total driving force of a quadrotor in the body frame (the \mathcal{B} frame) is the sum of the thrust ${}^{\mathcal{B}}\mathbf{F}_t$ and drag force ${}^{\mathcal{B}}\mathbf{F}_d$ generated by each propulsion unit as follow:

$${}^{\mathcal{B}}\mathbf{F} = \sum_{i=1}^4 ({}^{\mathcal{B}}\mathbf{F}_{t_i} - {}^{\mathcal{B}}\mathbf{F}_{d_i}) = \sum_{i=1}^4 (\tau u_i^2 \mathbf{e}_3 - u_i D {}^{\mathcal{B}}\mathbf{v}_i), \quad (3)$$

where τ is the thrust coefficient for the propellers, $D = \text{diag}(d_x, d_y, d_z)$ is the matrix of effective linear drag coefficients, $\mathbf{e}_3 = [0, 0, 1]^\top$ is the z axis in any frame, and u_i and ${}^{\mathcal{B}}\mathbf{v}_i$ are the rotation speed and velocity of the i -th rotor, respectively. Actually, the velocity of each rotor is:

$${}^{\mathcal{B}}\mathbf{v}_i = {}^{\mathcal{B}}\mathbf{v} + {}^{\mathcal{B}}\boldsymbol{\omega} \times {}^{\mathcal{B}}\mathbf{r}_i^{\mathcal{B}}, \quad (4)$$

where ${}^{\mathcal{B}}\mathbf{v}$ and ${}^{\mathcal{B}}\boldsymbol{\omega}$ are the linear and angular velocity of the quadrotor's center of mass (CoM), ${}^{\mathcal{B}}\mathbf{r}_i^{\mathcal{B}}$ is the position of the i -th rotor relative to the CoM. To simplify the calculation, we ignore the velocity discrepancy of different rotors and express it as:

$${}^{\mathcal{B}}\mathbf{v}_i \approx {}^{\mathcal{B}}\mathbf{v}. \quad (5)$$

The input notations are abbreviated as:

$$U_{ss} = \sum_{i=1}^4 u_i^2, \quad U_s = \sum_{i=1}^4 u_i, \quad (6)$$

so we can obtain the Newtonian equation in the \mathcal{G} frame:

$$m \frac{d}{dt} ({}^{\mathcal{G}}\mathbf{v}) = \frac{\mathcal{G}}{\mathcal{B}}\mathbf{R} (\tau U_{ss} \mathbf{e}_3 - U_s D \frac{\mathcal{G}}{\mathcal{B}}\mathbf{R}^\top \mathbf{g}) - m {}^{\mathcal{G}}\mathbf{g}. \quad (7)$$

II. TWO-STAGE EKF FOR INERTIAL DYNAMICAL FUSION

A. Rotation Stage

1) *State*: In the rotation stage, the rotation of the \mathcal{I} frame in the \mathcal{G} frame is taken as state:

$$\mathbf{x} = \frac{\mathcal{G}}{\mathcal{I}}\mathbf{q}. \quad (8)$$

2) *Process Model*: The rotational equation is given:

$$\dot{\mathbf{x}} = \frac{1}{2} \mathcal{G} \mathbf{q} \otimes \mathcal{I} \boldsymbol{\omega} = \frac{1}{2} \mathcal{G} \mathbf{q} \otimes (\mathcal{I} \hat{\boldsymbol{\omega}} + \mathbf{n}_\omega), \quad (9)$$

where $\mathcal{I} \hat{\boldsymbol{\omega}} = \mathcal{I} \tilde{\boldsymbol{\omega}} - \mathcal{I} \hat{\mathbf{b}}_\omega$, $\mathbf{n}_\omega \sim \mathcal{N}(0, \Sigma_\omega^2)$, and $\mathcal{I} \hat{\mathbf{b}}_\omega$ is the output of the gyroscope *De-Bias Net*.

The above differential equation is discretized as follows:

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{F}(\mathbf{x}_k, \mathbf{u}_k, \mathbf{n}_k) \\ &= \mathbf{F}_x \mathbf{x}_k + \mathbf{F}_n \mathbf{n}_k, \end{aligned} \quad (10)$$

where,

$$\begin{aligned} \mathbf{F}_x &= \frac{\partial \mathbf{F}}{\partial \mathbf{x}_k} = \mathbf{I}_4 + \frac{\Delta t}{2} \begin{bmatrix} 0 & -\mathcal{I} \hat{\boldsymbol{\omega}}^\top \\ \mathcal{I} \hat{\boldsymbol{\omega}} & [\mathcal{I} \hat{\boldsymbol{\omega}}]_\times \end{bmatrix}, \\ \mathbf{F}_n &= \frac{\partial \mathbf{F}}{\partial \mathbf{n}_\omega} = \frac{\Delta t}{2} \begin{bmatrix} -q_x & -q_y & -q_z \\ q_w & -q_z & q_y \\ q_z & q_w & -q_x \\ -q_y & q_x & q_w \end{bmatrix} = \frac{\Delta t}{2} \begin{bmatrix} -\mathbf{q}_v \\ q_w \mathbf{I}_3 + [\mathbf{q}_v]_\times \end{bmatrix}. \end{aligned} \quad (11)$$

Specifically, $\mathbf{q} = [q_w, q_x, q_y, q_z]^\top = [q_w, \mathbf{q}_v^\top]^\top$.

3) *Measurement Model*: The attitude controllers of most flying robots rely on the complementary filters [2, 3] to obtain attitude observations. Similarly, we consider the gravity alignment constraint to obtain the tilt observation:

$$\mathcal{I} \hat{\mathbf{a}} \approx \mathcal{G} \mathbf{R}^\top \mathbf{g} + \mathbf{n}_a. \quad (12)$$

Similarly, $\mathcal{I} \hat{\mathbf{a}} = \mathcal{I} \tilde{\mathbf{a}} - \mathcal{I} \hat{\mathbf{b}}_a$, $\mathbf{n}_a \sim \mathcal{N}(0, \Sigma_a^2)$, and $\mathcal{I} \hat{\mathbf{b}}_a$ is the output of the accelerometer *De-Bias Net*.

The above observation equation is approximated as:

$$\mathbf{z}_k = \mathbf{H}(\mathbf{x}_k) = \mathbf{H}_x \mathbf{x}_k, \quad (13)$$

where,

$$\mathbf{H}_x = 2 \begin{bmatrix} \mathbf{e}_3 \times \mathbf{q}_v & [\mathbf{e}_3 \times \mathbf{q}_v + q_w \mathbf{e}_3]_\times + (\mathbf{q}_v \cdot \mathbf{e}_3) \mathbf{I}_3 - \mathbf{e}_3 \mathbf{q}_v^\top \end{bmatrix}. \quad (14)$$

4) *Extended Kalman Filter*:

$$\begin{aligned} \mathbf{x}_{k+1|k} &= \mathbf{F}(\mathbf{x}_{k|k}, \mathbf{u}_k), \\ \mathbf{P}_{k+1|k} &= \mathbf{F}_x \mathbf{P}_{k|k} \mathbf{F}_x^\top + \mathbf{F}_u \mathbf{Q}_k \mathbf{F}_u^\top, \\ \mathbf{K}_{k+1} &= \mathbf{P}_{k+1|k} \mathbf{H}_x^\top (\mathbf{H}_x \mathbf{P}_{k+1|k} \mathbf{H}_x^\top + \mathbf{R}_{k+1})^{-1}, \\ \mathbf{x}_{k+1|k+1} &= \mathbf{x}_{k+1|k} \oplus (\mathbf{K}_{k+1} (\mathbf{z}_{k+1} - \mathbf{H}(\mathbf{x}_{k+1|k}))), \\ \mathbf{P}_{k+1|k+1} &= (\mathbf{I} - \mathbf{K}_{k+1} \mathbf{H}_x) \mathbf{P}_{k+1|k}, \end{aligned} \quad (15)$$

where,

$$\mathbf{Q} = \Sigma_\omega^2, \quad \mathbf{R} = \Sigma_a^2. \quad (16)$$

B. Translation Stage

1) *State*: The state of the second stage is defined as:

$$\mathbf{x} = (\mathcal{G} \mathbf{p}_B^\mathcal{G}, \mathcal{G} \mathbf{v}_B^\mathcal{G}, \tau, \mathbf{d}, \mathcal{I}_B \mathbf{q}, \mathcal{I} \mathbf{t}_B^\mathcal{I}), \quad (17)$$

where $\mathcal{G} \mathbf{p}_B^\mathcal{G}$ and $\mathcal{G} \mathbf{v}_B^\mathcal{G}$ are respectively the velocity and position of the quadrotor body \mathcal{B} frame expressed in the \mathcal{G} frame, τ is the thrust coefficient, \mathbf{d} is the drag vector of (d_x, d_y, d_z) , and $(\mathcal{I}_B \mathbf{q}, \mathcal{I} \mathbf{t}_B^\mathcal{I})$ is the extrinsic parameter between the \mathcal{B} and the \mathcal{I} frame.

2) *Process Model*: We regard the quadrotor dynamics as the input, and express the complete process model as follows:

$$\begin{aligned}
{}^{\mathcal{G}}\dot{\mathbf{p}}_{\mathcal{B}} &= {}^{\mathcal{G}}\mathbf{v}_{\mathcal{B}}, \\
{}^{\mathcal{G}}\dot{\mathbf{v}}_{\mathcal{B}} &= \frac{1}{m} {}^{\mathcal{G}}\mathbf{R} \left(\tau U_{ss} \mathbf{e}_3 - U_s D {}^{\mathcal{G}}\mathbf{R}^{\top} {}^{\mathcal{G}}\mathbf{v}_{\mathcal{B}} + \hat{\mathbf{f}}_{res} + \mathbf{n}_f \right) - {}^{\mathcal{G}}\mathbf{g}, \\
\dot{\tau} &= 0, \\
\dot{\mathbf{d}} &= \mathbf{0}, \\
\frac{\mathcal{I}}{\mathcal{B}}\dot{\boldsymbol{\theta}} &= \mathbf{0}, \\
\frac{\mathcal{I}}{\mathcal{B}}\dot{\mathbf{t}}_{\mathcal{B}} &= \mathbf{0},
\end{aligned} \tag{18}$$

where ${}^{\mathcal{G}}\mathbf{R} = \frac{\mathcal{G}}{\mathcal{I}}\mathbf{R}_{\mathcal{B}}^{\mathcal{I}}\mathbf{R}$ is the rotation from the quadrotor body \mathcal{B} frame to the \mathcal{G} frame, and $\hat{\mathbf{f}}_{res}$ and $\hat{\Sigma}_f^2$ ($\mathbf{n}_f \sim \mathcal{N}(0, \hat{\Sigma}_f^2)$) are the *Res-Dynamics Net* outputs.

The above differential equation is discretized as follows:

$$\begin{aligned}
\mathbf{x}_{k+1} &= \mathbf{F}(\mathbf{x}_k, \mathbf{u}_k, \mathbf{n}_k) \\
&= \mathbf{F}_x \mathbf{x}_k + \mathbf{F}_n \mathbf{n}_k,
\end{aligned} \tag{19}$$

where,

$$\begin{aligned}
\mathbf{F}_x &= \begin{bmatrix} \mathbf{I}_3 & \frac{\partial \mathbf{p}_{k+1}}{\partial \mathbf{v}_k} & \mathbf{0}_{31} & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \frac{\partial \mathbf{v}_{k+1}}{\partial \mathbf{v}_k} & \frac{\partial \mathbf{v}_{k+1}}{\partial \tau} & \frac{\partial \mathbf{v}_{k+1}}{\partial \mathbf{d}} & \frac{\partial \mathbf{v}_{k+1}}{\partial \boldsymbol{\theta}} & \mathbf{0}_3 \\ \mathbf{0}_{13} & \mathbf{0}_{13} & 1 & \mathbf{0}_{13} & \mathbf{0}_{13} & \mathbf{0}_{13} \\ \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_{31} & \mathbf{I}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_{31} & \mathbf{0}_3 & \mathbf{I}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_{31} & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{I}_3 \end{bmatrix}, \\
\mathbf{F}_n &= \begin{bmatrix} \mathbf{0}_3 & \frac{\partial \mathbf{v}_{k+1}}{\partial \mathbf{n}_f} & \mathbf{0}_{31} & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 \end{bmatrix}^{\top}.
\end{aligned} \tag{20}$$

Specifically,

$$\begin{aligned}
\frac{\partial \mathbf{p}_{k+1}}{\partial \mathbf{v}_k} &= \mathbf{I}_3 \Delta t, \\
\frac{\partial \mathbf{v}_{k+1}}{\partial \mathbf{v}_k} &= \mathbf{I} - \frac{U_s \Delta t}{m} {}^{\mathcal{G}}\mathbf{R} D {}^{\mathcal{G}}\mathbf{R}^{\top}, \\
\frac{\partial \mathbf{v}_{k+1}}{\partial \tau} &= \frac{U_{ss} \Delta t}{m} {}^{\mathcal{G}}\mathbf{R} \mathbf{e}_3, \\
\frac{\partial \mathbf{v}_{k+1}}{\partial \mathbf{d}} &= -\frac{\Delta t}{m} \begin{bmatrix} {}^{\mathcal{G}}\mathbf{r}_1 {}^{\mathcal{G}}\mathbf{r}_1^{\top} {}^{\mathcal{G}}\mathbf{v}_{\mathcal{B}} & {}^{\mathcal{G}}\mathbf{r}_2 {}^{\mathcal{G}}\mathbf{r}_2^{\top} {}^{\mathcal{G}}\mathbf{v}_{\mathcal{B}} & {}^{\mathcal{G}}\mathbf{r}_3 {}^{\mathcal{G}}\mathbf{r}_3^{\top} {}^{\mathcal{G}}\mathbf{v}_{\mathcal{B}} \end{bmatrix}, \\
\frac{\partial \mathbf{v}_{k+1}}{\partial \boldsymbol{\theta}} &= \frac{\Delta t}{m} {}^{\mathcal{G}}\mathbf{R} \left(-\tau U_{ss} [\mathbf{e}_3]_{\times} + U_s \left([D {}^{\mathcal{G}}\mathbf{R}^{\top} {}^{\mathcal{G}}\mathbf{v}_{\mathcal{B}}]_{\times} - D [{}^{\mathcal{G}}\mathbf{R}^{\top} {}^{\mathcal{G}}\mathbf{v}_{\mathcal{B}}]_{\times} \right) - [\hat{\mathbf{f}}_{res}]_{\times} \right), \\
\frac{\partial \mathbf{v}_{k+1}}{\partial \mathbf{n}_f} &= \frac{1}{m} {}^{\mathcal{G}}\mathbf{R} \Delta t,
\end{aligned} \tag{21}$$

and $\mathbf{r}_i = \mathbf{R} \mathbf{e}_i$.

3) *Measurement Models*: There are three measurements corresponding to the translation stage: the IMU acceleration, the network velocity, and the network displacement measurements.

a) *dynamics constraint measurements*: The dynamics constraint measurements are expressed as:

$$\begin{aligned} {}^{\mathcal{I}}\hat{\mathbf{a}} &= \mathbf{H}_d(\mathbf{x}, \mathbf{n}_a) \\ &= \frac{1}{m} {}^{\mathcal{I}}\mathbf{R} \left(\tau U_{ss} \mathbf{e}_3 - U_s D_B^{\mathcal{G}} \mathbf{R}^{\top \mathcal{G}} \mathbf{v}_B^{\mathcal{G}} + \hat{\mathbf{f}}_{res} \right) + {}^{\mathcal{I}}\hat{\boldsymbol{\omega}} \times ({}^{\mathcal{I}}\hat{\boldsymbol{\omega}} \times {}^{\mathcal{I}}\mathbf{t}_B^{\mathcal{I}}) + {}^{\mathcal{I}}\hat{\boldsymbol{\alpha}} \times {}^{\mathcal{I}}\mathbf{t}_B^{\mathcal{I}} + \mathbf{n}_a, \end{aligned} \quad (22)$$

whose jacobian can be given as:

$$\frac{\partial \mathbf{H}_d}{\partial \mathbf{x}_k} = \begin{bmatrix} \mathbf{0}_3 & \frac{\partial \mathbf{H}_d}{\partial \mathbf{v}_k} & \frac{\partial \mathbf{H}_d}{\partial \tau} & \frac{\partial \mathbf{H}_d}{\partial \mathbf{d}} & \frac{\partial \mathbf{H}_d}{\partial {}^{\mathcal{I}}\boldsymbol{\theta}} & \frac{\partial \mathbf{H}_d}{\partial {}^{\mathcal{I}}\mathbf{t}_B^{\mathcal{I}}} \end{bmatrix}, \quad (23)$$

where,

$$\begin{aligned} \frac{\partial \mathbf{H}_d}{\partial \mathbf{v}_k} &= -\frac{U_s}{m} {}^{\mathcal{I}}\mathbf{R} D_B^{\mathcal{G}} \mathbf{R}^{\top}, \\ \frac{\partial \mathbf{H}_d}{\partial \tau} &= \frac{U_{ss}}{m} {}^{\mathcal{I}}\mathbf{R} \mathbf{e}_3, \\ \frac{\partial \mathbf{H}_d}{\partial \mathbf{d}} &= -\frac{U_s}{m} \left[{}^{\mathcal{I}}\mathbf{r}_{1B}^{\mathcal{G}} \mathbf{r}_{1B}^{\top \mathcal{G}} \mathbf{v}_B^{\mathcal{G}} \quad {}^{\mathcal{I}}\mathbf{r}_{2B}^{\mathcal{G}} \mathbf{r}_{2B}^{\top \mathcal{G}} \mathbf{v}_B^{\mathcal{G}} \quad {}^{\mathcal{I}}\mathbf{r}_{3B}^{\mathcal{G}} \mathbf{r}_{3B}^{\top \mathcal{G}} \mathbf{v}_B^{\mathcal{G}} \right], \\ \frac{\partial \mathbf{H}_d}{\partial {}^{\mathcal{I}}\boldsymbol{\theta}} &= \frac{1}{m} {}^{\mathcal{I}}\mathbf{R} \left(-\tau U_{ss} [\mathbf{e}_3]_{\times} + U_s \left([D_B^{\mathcal{G}} \mathbf{R}^{\top \mathcal{G}} \mathbf{v}_B^{\mathcal{G}}]_{\times} - D [{}^{\mathcal{G}}\mathbf{R}^{\top \mathcal{G}} \mathbf{v}_B^{\mathcal{G}}]_{\times} \right) - [\hat{\mathbf{f}}_{res}]_{\times} \right), \\ \frac{\partial \mathbf{H}_d}{\partial {}^{\mathcal{I}}\mathbf{t}_B^{\mathcal{I}}} &= [{}^{\mathcal{I}}\hat{\boldsymbol{\omega}}]_{\times} [{}^{\mathcal{I}}\hat{\boldsymbol{\omega}}]_{\times} + [{}^{\mathcal{I}}\hat{\boldsymbol{\alpha}}]_{\times}. \end{aligned} \quad (24)$$

Specifically, ${}^{\mathcal{I}}\hat{\boldsymbol{\alpha}}$ is the angular acceleration in the \mathcal{I} frame, which is obtained by differentiating the angular velocity, ${}^{\mathcal{I}}\hat{\boldsymbol{\alpha}} = \frac{d}{dt} {}^{\mathcal{I}}\hat{\boldsymbol{\omega}}$. In practice, we low-pass filter the $\hat{\boldsymbol{\omega}}, \hat{\boldsymbol{\alpha}}$ to reduce noise. And, the noise of IMU acceleration measurements is $\mathbf{n}_a \sim \mathcal{N}(0, \Sigma_a^2)$.

b) *network velocity measurements*: The velocity measurements from *V-P Net* are expressed as:

$$\begin{aligned} {}^{\mathcal{G}}\hat{\mathbf{v}}_{\mathcal{I}}^{\mathcal{G}} &= \mathbf{H}_v(\mathbf{x}, \mathbf{n}_v) \\ &= {}^{\mathcal{G}}\mathbf{v}_B^{\mathcal{G}} - {}^{\mathcal{G}}\mathbf{R} ({}^{\mathcal{I}}\hat{\boldsymbol{\omega}} \times {}^{\mathcal{I}}\mathbf{t}_B^{\mathcal{I}}) + \mathbf{n}_v, \end{aligned} \quad (25)$$

whose jacobian can be given as:

$$\frac{\partial \mathbf{H}_v}{\partial \mathbf{x}_k} = \begin{bmatrix} \mathbf{0}_3 & \mathbf{I}_3 & \mathbf{0}_{31} & \mathbf{0}_3 & \mathbf{0}_3 & \frac{\partial \mathbf{r}_v}{\partial {}^{\mathcal{I}}\mathbf{t}_B^{\mathcal{I}}} \end{bmatrix}, \quad (26)$$

where,

$$\frac{\partial \mathbf{r}_v}{\partial {}^{\mathcal{I}}\mathbf{t}_B^{\mathcal{I}}} = -{}^{\mathcal{G}}\mathbf{R} [{}^{\mathcal{I}}\hat{\boldsymbol{\omega}}]_{\times}. \quad (27)$$

Specifically, the noise of network velocity measurements is $\mathbf{n}_v \sim \mathcal{N}(0, \hat{\Sigma}_v^2)$.

c) *network displacement measurements*: The displacement measurements from *V-P Net* are expressed as:

$$\begin{aligned} {}^{\mathcal{G}}\hat{\mathbf{p}}_{\mathcal{I}}^{\mathcal{G}} &= \mathbf{H}_p(\mathbf{x}, \mathbf{n}_p) \\ &= {}^{\mathcal{G}}\mathbf{p}_B^{\mathcal{G}} - {}^{\mathcal{G}}\mathbf{R} {}^{\mathcal{I}}\mathbf{t}_B^{\mathcal{I}} + \mathbf{n}_p, \end{aligned} \quad (28)$$

whose jacobian can be given as:

$$\frac{\partial \mathbf{H}_p}{\partial \mathbf{x}_k} = \begin{bmatrix} \mathbf{I}_3 & \mathbf{0}_3 & \mathbf{0}_{31} & \mathbf{0}_3 & \mathbf{0}_3 & \frac{\partial \mathbf{H}_p}{\partial {}^{\mathcal{I}}\mathbf{t}_B^{\mathcal{I}}} \end{bmatrix}, \quad (29)$$

where,

$$\frac{\partial \mathbf{H}_p}{\partial {}^{\mathcal{I}}\mathbf{t}_B^{\mathcal{I}}} = -{}^{\mathcal{G}}\mathbf{R}. \quad (30)$$

Specifically, the noise of network displacement measurements is $\mathbf{n}_p \sim \mathcal{N}(0, \hat{\Sigma}_p^2)$.

4) *Discrete Extended Kalman Filter*: The complete EKF procedures for both rotation and translation can be written in a discretized form as follows:

$$\begin{aligned}
\mathbf{x}_{k+1|k} &= \mathbf{F}(\mathbf{x}_{k|k}, \mathbf{u}_k), \\
\mathbf{P}_{k+1|k} &= \mathbf{F}_x \mathbf{P}_{k|k} \mathbf{F}_x^T + \mathbf{F}_u \mathbf{Q}_k \mathbf{F}_u^T, \\
\mathbf{K}_{k+1} &= \mathbf{P}_{k+1|k} \mathbf{H}_x^T (\mathbf{H}_x \mathbf{P}_{k+1|k} \mathbf{H}_x^T + \mathbf{R}_{k+1})^{-1}, \\
\mathbf{x}_{k+1|k+1} &= \mathbf{x}_{k+1|k} \oplus (\mathbf{K}_{k+1} (\mathbf{z}_{k+1} - h(\mathbf{x}_{k+1|k}))), \\
\mathbf{P}_{k+1|k+1} &= (\mathbf{I} - \mathbf{K}_{k+1} \mathbf{H}_x) \mathbf{P}_{k+1|k},
\end{aligned} \tag{31}$$

where,

$$\mathbf{Q} = \hat{\Sigma}_f^2, \quad \mathbf{R} = \text{diag}\{\Sigma_a^2, \hat{\Sigma}_v^2, \hat{\Sigma}_p^2\}. \tag{32}$$

III. ONE-STAGE EKF FOR INERTIAL DYNAMICAL FUSION

A. State

The state of the system is defined as:

$$\mathbf{x} = (\mathcal{I} \mathbf{q}, \mathcal{G} \mathbf{p}_B^{\mathcal{G}}, \mathcal{G} \mathbf{v}_B^{\mathcal{G}}, \tau, \mathbf{d}, \mathcal{I}_B \mathbf{q}, \mathcal{I} \mathbf{t}_B^{\mathcal{I}}). \tag{33}$$

B. Process Model

We regard the quadrotor dynamics and gyroscope outputs of *De-Bias Net* as the input, and express the complete process model as follows:

$$\begin{aligned}
\mathcal{I} \dot{\mathbf{q}} &= \frac{1}{2} \mathcal{I} \mathbf{q} \otimes (\mathcal{I} \hat{\boldsymbol{\omega}} + \mathbf{n}_\omega), \\
\mathcal{G} \dot{\mathbf{p}}_B^{\mathcal{G}} &= \mathcal{G} \mathbf{v}_B^{\mathcal{G}}, \\
\mathcal{G} \dot{\mathbf{v}}_B^{\mathcal{G}} &= \frac{1}{m} \mathcal{G} \mathbf{R} \left(\tau U_{ss} \mathbf{e}_3 - U_s D_B^{\mathcal{G}} \mathbf{R}^T \mathcal{G} \mathbf{v}_B^{\mathcal{G}} + \hat{\mathbf{f}}_{res} + \mathbf{n}_f \right) - \mathcal{G} \mathbf{g}, \\
\dot{\tau} &= 0, \\
\dot{\mathbf{d}} &= \mathbf{0}, \\
\mathcal{I}_B \dot{\mathbf{q}} &= \mathbf{0}, \\
\mathcal{I} \dot{\mathbf{t}}_B^{\mathcal{I}} &= \mathbf{0},
\end{aligned} \tag{34}$$

The above differential equation is discretized as follows:

$$\begin{aligned}
\mathbf{x}_{k+1} &= \mathbf{F}(\mathbf{x}_k, \mathbf{u}_k) \\
&= \mathbf{F}_x \mathbf{x}_k + \mathbf{F}_n \mathbf{n}_k,
\end{aligned} \tag{35}$$

where,

$$\begin{aligned}
\mathbf{F}_x &= \begin{bmatrix} \frac{\partial \mathcal{I} \boldsymbol{\theta}_{k+1}}{\partial \mathcal{I} \boldsymbol{\theta}_k} & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_{31} & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{I}_3 & \mathbf{I}_3 \Delta t & \mathbf{0}_{31} & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ \frac{\partial \mathbf{v}_{k+1}}{\partial \mathcal{I} \boldsymbol{\theta}_k} & \mathbf{0}_3 & \frac{\partial \mathbf{v}_{k+1}}{\partial \mathbf{v}_k} & \frac{\partial \mathbf{v}_{k+1}}{\partial \tau} & \frac{\partial \mathbf{v}_{k+1}}{\partial \mathbf{d}} & \frac{\partial \mathbf{v}_{k+1}}{\partial \mathcal{I} \boldsymbol{\theta}} & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_{13} & \mathbf{0}_{13} & 1 & \mathbf{0}_{13} & \mathbf{0}_{13} & \mathbf{0}_{13} \\ \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_{31} & \mathbf{I}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_{31} & \mathbf{0}_3 & \mathbf{I}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_{31} & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{I}_3 \end{bmatrix}, \\
\mathbf{F}_n &= \begin{bmatrix} \frac{\partial \mathcal{I} \boldsymbol{\theta}_{k+1}}{\partial \mathbf{n}_\omega} & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_{31} & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 & \frac{\partial \mathbf{v}_{k+1}}{\partial \mathbf{n}_f} & \mathbf{0}_{31} & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 \end{bmatrix}^\top.
\end{aligned} \tag{36}$$

Specifically,

$$\begin{aligned}
\frac{\partial_{\mathcal{I}}^g \boldsymbol{\theta}_{k+1}}{\partial_{\mathcal{I}}^g \boldsymbol{\theta}_k} &= \text{Exp}(-{}^{\mathcal{I}}\hat{\boldsymbol{\omega}}\Delta t), \\
\frac{\partial \mathbf{v}_{k+1}}{\partial_{\mathcal{I}}^g \boldsymbol{\theta}_k} &= \frac{\Delta t}{m} {}^{\mathcal{I}}\mathbf{R} \left(-\tau U_{ss} [\mathbf{r}_3]_{\times} + U_s \left({}^{\mathcal{G}}\mathbf{R} [\mathbf{r}_3 \mathbf{D}_{\mathcal{B}}^{\mathcal{G}} \mathbf{R}^{\top} \mathbf{v}_{\mathcal{B}}^{\mathcal{G}}]_{\times} - {}^{\mathcal{G}}\mathbf{R} \mathbf{D}_{\mathcal{B}}^{\mathcal{I}} \mathbf{R}^{\top} [{}^{\mathcal{G}}\mathbf{R}^{\top} \mathbf{v}_{\mathcal{B}}^{\mathcal{G}}]_{\times} \right) - [\mathbf{r}_3 \hat{\mathbf{f}}_{res}]_{\times} \right), \\
\frac{\partial \mathbf{v}_{k+1}}{\partial \mathbf{v}_k} &= \mathbf{I} - \frac{U_s \Delta t}{m} {}^{\mathcal{G}}\mathbf{R} \mathbf{D}_{\mathcal{B}}^{\mathcal{G}} \mathbf{R}^{\top}, \\
\frac{\partial \mathbf{v}_{k+1}}{\partial \tau} &= \frac{U_{ss} \Delta t}{m} {}^{\mathcal{G}}\mathbf{r}_3, \\
\frac{\partial \mathbf{v}_{k+1}}{\partial \mathbf{d}} &= -\frac{\Delta t}{m} [{}^{\mathcal{G}}\mathbf{r}_1 {}^{\mathcal{G}}\mathbf{r}_1^{\top} \mathbf{v}_{\mathcal{B}}^{\mathcal{G}} \quad {}^{\mathcal{G}}\mathbf{r}_2 {}^{\mathcal{G}}\mathbf{r}_2^{\top} \mathbf{v}_{\mathcal{B}}^{\mathcal{G}} \quad {}^{\mathcal{G}}\mathbf{r}_3 {}^{\mathcal{G}}\mathbf{r}_3^{\top} \mathbf{v}_{\mathcal{B}}^{\mathcal{G}}], \\
\frac{\partial \mathbf{v}_{k+1}}{\partial_{\mathcal{I}}^g \boldsymbol{\theta}} &= \frac{\Delta t}{m} {}^{\mathcal{G}}\mathbf{R} \left(-\tau U_{ss} [\mathbf{e}_3]_{\times} + U_s \left([\mathbf{D}_{\mathcal{B}}^{\mathcal{G}} \mathbf{R}^{\top} \mathbf{v}_{\mathcal{B}}^{\mathcal{G}}]_{\times} - \mathbf{D} [{}^{\mathcal{G}}\mathbf{R}^{\top} \mathbf{v}_{\mathcal{B}}^{\mathcal{G}}]_{\times} \right) - [\hat{\mathbf{f}}_{res}]_{\times} \right), \\
\frac{\partial_{\mathcal{I}}^g \boldsymbol{\theta}_{k+1}}{\partial \mathbf{n}_{\omega}} &= \mathbf{J}_r ({}^{\mathcal{I}}\hat{\boldsymbol{\omega}}\Delta t) \Delta t, \\
\frac{\partial \mathbf{v}_{k+1}}{\partial \mathbf{n}_f} &= \frac{1}{m} {}^{\mathcal{G}}\mathbf{R} \Delta t,
\end{aligned} \tag{37}$$

where,

$$\mathbf{J}_r(\boldsymbol{\theta}) = \mathbf{I}_3 - \frac{1 - \cos\|\boldsymbol{\theta}\|}{\|\boldsymbol{\theta}\|^2} [\boldsymbol{\theta}]_{\times} + \frac{\|\boldsymbol{\theta}\| - \sin\|\boldsymbol{\theta}\|}{\|\boldsymbol{\theta}\|^3} [\boldsymbol{\theta}]_{\times}^2 \tag{38}$$

C. Measurement Model

1) *gravity alignment measurements*: The attitude controllers of most flying robots rely on the complementary filters [2, 3] to obtain attitude observations. Similarly, we consider the gravity alignment constraint to obtain the tilt observation:

$$\begin{aligned}
{}^{\mathcal{I}}\hat{\mathbf{a}} &= \mathbf{H}_g(\mathbf{x}, \mathbf{n}_g) \\
&\approx {}^{\mathcal{G}}\mathbf{R}^{\top} \mathbf{g} + \mathbf{n}_g,
\end{aligned} \tag{39}$$

whose jacobian can be given as:

$$\frac{\partial \mathbf{H}_g}{\partial \mathbf{x}_k} = \begin{bmatrix} \frac{\partial \mathbf{H}_g}{\partial_{\mathcal{I}}^g \boldsymbol{\theta}_k} & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_{31} & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 \end{bmatrix}, \tag{40}$$

where,

$$\frac{\partial \mathbf{H}_g}{\partial_{\mathcal{I}}^g \boldsymbol{\theta}_k} = [{}^{\mathcal{G}}\mathbf{R}^{\top} \mathbf{g}]_{\times}. \tag{41}$$

2) *dynamics constraint measurements*: The dynamics constraint measurements are expressed as:

$$\begin{aligned}
{}^{\mathcal{I}}\hat{\mathbf{a}} &= \mathbf{H}_d(\mathbf{x}, \mathbf{n}_a) \\
&= \frac{1}{m} {}^{\mathcal{I}}\mathbf{R} \left(\tau U_{ss} \mathbf{e}_3 - U_s \mathbf{D}_{\mathcal{B}}^{\mathcal{G}} \mathbf{R}^{\top} \mathbf{v}_{\mathcal{B}}^{\mathcal{G}} + \hat{\mathbf{f}}_{res} \right) + {}^{\mathcal{I}}\hat{\boldsymbol{\omega}} \times ({}^{\mathcal{I}}\hat{\boldsymbol{\omega}} \times {}^{\mathcal{I}}\mathbf{t}_{\mathcal{B}}^{\mathcal{I}}) + {}^{\mathcal{I}}\hat{\boldsymbol{\alpha}} \times {}^{\mathcal{I}}\mathbf{t}_{\mathcal{B}}^{\mathcal{I}} + \mathbf{n}_a,
\end{aligned} \tag{42}$$

whose jacobian can be given as:

$$\frac{\partial \mathbf{H}_d}{\partial \mathbf{x}_k} = \begin{bmatrix} \frac{\partial \mathbf{H}_d}{\partial_{\mathcal{I}}^g \boldsymbol{\theta}_k} & \mathbf{0}_3 & \frac{\partial \mathbf{H}_d}{\partial \mathbf{v}_k} & \frac{\partial \mathbf{H}_d}{\partial \tau} & \frac{\partial \mathbf{H}_d}{\partial \mathbf{d}} & \frac{\partial \mathbf{H}_d}{\partial_{\mathcal{I}}^g \boldsymbol{\theta}} & \frac{\partial \mathbf{H}_d}{\partial_{\mathcal{I}}^g \mathbf{t}_{\mathcal{B}}^{\mathcal{I}}} \end{bmatrix}, \tag{43}$$

where,

$$\begin{aligned}
\frac{\partial \mathbf{H}_d}{\partial \mathcal{I} \boldsymbol{\theta}_k} &= -\frac{U_s}{m} \mathcal{I} \mathbf{R} D_{\mathcal{B}}^{\mathcal{I}} \mathbf{R}^{\top} \left[\frac{\mathcal{G}}{\mathcal{I}} \mathbf{R}^{\top \mathcal{G}} \mathbf{v}_{\mathcal{B}}^{\mathcal{G}} \right]_{\times}, \\
\frac{\partial \mathbf{H}_d}{\partial \mathbf{v}_k} &= -\frac{U_s}{m} \mathcal{I} \mathbf{R} D_{\mathcal{B}}^{\mathcal{G}} \mathbf{R}^{\top}, \\
\frac{\partial \mathbf{H}_d}{\partial \tau} &= \frac{U_{ss}}{m} \mathcal{I} \mathbf{R} \mathbf{r}_3, \\
\frac{\partial \mathbf{H}_d}{\partial \mathbf{d}} &= -\frac{U_s}{m} \left[\mathcal{I} \mathbf{r}_1^{\mathcal{G}} \mathbf{r}_1^{\top \mathcal{G}} \mathbf{v}_{\mathcal{B}}^{\mathcal{G}} \quad \mathcal{I} \mathbf{r}_2^{\mathcal{G}} \mathbf{r}_2^{\top \mathcal{G}} \mathbf{v}_{\mathcal{B}}^{\mathcal{G}} \quad \mathcal{I} \mathbf{r}_3^{\mathcal{G}} \mathbf{r}_3^{\top \mathcal{G}} \mathbf{v}_{\mathcal{B}}^{\mathcal{G}} \right], \\
\frac{\partial \mathbf{H}_d}{\partial \mathcal{I} \boldsymbol{\theta}} &= \frac{1}{m} \mathcal{I} \mathbf{R} \left(-k U_{ss} [\mathbf{e}_3]_{\times} + U_s \left([D_{\mathcal{B}}^{\mathcal{G}} \mathbf{R}^{\top \mathcal{G}} \mathbf{v}_{\mathcal{B}}^{\mathcal{G}}]_{\times} - D [\frac{\mathcal{G}}{\mathcal{B}} \mathbf{R}^{\top \mathcal{G}} \mathbf{v}_{\mathcal{B}}^{\mathcal{G}}]_{\times} \right) - [\hat{\mathbf{f}}_{res}]_{\times} \right), \\
\frac{\partial \mathbf{H}_d}{\partial \mathcal{I} \mathbf{t}_{\mathcal{B}}^{\mathcal{I}}} &= [\mathcal{I} \hat{\boldsymbol{\omega}}]_{\times} [\mathcal{I} \hat{\boldsymbol{\omega}}]_{\times} + [\mathcal{I} \hat{\boldsymbol{\alpha}}]_{\times}.
\end{aligned} \tag{44}$$

Specifically, $\mathcal{I} \hat{\boldsymbol{\alpha}}$ is the angular acceleration in the \mathcal{I} frame, which is obtained by differentiating the angular velocity, $\mathcal{I} \hat{\boldsymbol{\alpha}} = \frac{d}{dt} \mathcal{I} \hat{\boldsymbol{\omega}}$. In practice, we low-pass filter the $\hat{\boldsymbol{\omega}}, \hat{\boldsymbol{\alpha}}$ to reduce noise. And, the noise of IMU acceleration measurements is $\mathbf{n}_a \sim \mathcal{N}(0, \Sigma_a^2)$.

3) *network velocity measurements*: The velocity measurements from *V-P Net* are expressed as:

$$\begin{aligned}
\mathcal{G} \hat{\mathbf{v}}_{\mathcal{I}}^{\mathcal{G}} &= \mathbf{H}_v(\mathbf{x}, \mathbf{n}_v) \\
&= \mathcal{G} \mathbf{v}_{\mathcal{B}}^{\mathcal{G}} - \frac{\mathcal{G}}{\mathcal{I}} \mathbf{R} (\mathcal{I} \hat{\boldsymbol{\omega}} \times \mathcal{I} \mathbf{t}_{\mathcal{B}}^{\mathcal{I}}) + \mathbf{n}_v,
\end{aligned} \tag{45}$$

whose jacobian can be given as:

$$\frac{\partial \mathbf{H}_v}{\partial \mathbf{x}_k} = \begin{bmatrix} \frac{\partial \mathbf{H}_v}{\partial \mathcal{I} \boldsymbol{\theta}_k} & \mathbf{0}_3 & \mathbf{I}_3 & \mathbf{0}_{31} & \mathbf{0}_3 & \mathbf{0}_3 & \frac{\partial \mathbf{H}_v}{\partial \mathcal{I} \mathbf{t}_{\mathcal{B}}^{\mathcal{I}}} \end{bmatrix}, \tag{46}$$

where,

$$\begin{aligned}
\frac{\partial \mathbf{H}_v}{\partial \mathcal{I} \boldsymbol{\theta}_k} &= \frac{\mathcal{G}}{\mathcal{I}} \mathbf{R} [\mathcal{I} \hat{\boldsymbol{\omega}} \times \mathcal{I} \mathbf{t}_{\mathcal{B}}^{\mathcal{I}}]_{\times}, \\
\frac{\partial \mathbf{H}_v}{\partial \mathcal{I} \mathbf{t}_{\mathcal{B}}^{\mathcal{I}}} &= -\frac{\mathcal{G}}{\mathcal{I}} \mathbf{R} [\mathcal{I} \hat{\boldsymbol{\omega}}]_{\times}.
\end{aligned} \tag{47}$$

Specifically, the noise of network velocity measurements is $\mathbf{n}_v \sim \mathcal{N}(0, \hat{\Sigma}_v^2)$.

4) *network displacement measurements*: The displacement measurements from *V-P Net* are expressed as:

$$\begin{aligned}
\mathcal{G} \hat{\mathbf{p}}_{\mathcal{I}}^{\mathcal{G}} &= \mathbf{H}_p(\mathbf{x}, \mathbf{n}_p) \\
&= \mathcal{G} \mathbf{p}_{\mathcal{B}}^{\mathcal{G}} - \frac{\mathcal{G}}{\mathcal{I}} \mathbf{R} \mathcal{I} \mathbf{t}_{\mathcal{B}}^{\mathcal{I}} + \mathbf{n}_p,
\end{aligned} \tag{48}$$

whose jacobian can be given as:

$$\frac{\partial \mathbf{H}_p}{\partial \mathbf{x}_k} = \begin{bmatrix} \frac{\partial \mathbf{H}_p}{\partial \mathcal{I} \boldsymbol{\theta}_k} & \mathbf{I}_3 & \mathbf{0}_3 & \mathbf{0}_{31} & \mathbf{0}_3 & \mathbf{0}_3 & \frac{\partial \mathbf{H}_p}{\partial \mathcal{I} \mathbf{t}_{\mathcal{B}}^{\mathcal{I}}} \end{bmatrix}, \tag{49}$$

where,

$$\begin{aligned}
\frac{\partial \mathbf{H}_p}{\partial \mathcal{I} \boldsymbol{\theta}_k} &= \frac{\mathcal{G}}{\mathcal{I}} \mathbf{R} [\mathcal{I} \mathbf{t}_{\mathcal{B}}^{\mathcal{I}}]_{\times}, \\
\frac{\partial \mathbf{H}_p}{\partial \mathcal{I} \mathbf{t}_{\mathcal{B}}^{\mathcal{I}}} &= -\frac{\mathcal{G}}{\mathcal{I}} \mathbf{R}.
\end{aligned} \tag{50}$$

Specifically, the noise of network displacement measurements is $\mathbf{n}_p \sim \mathcal{N}(0, \hat{\Sigma}_p^2)$.

D. Discrete Extended Kalman Filter

The complete EKF procedures for both rotation and translation can be written in a discretized form as follows:

$$\begin{aligned}
\mathbf{x}_{k+1|k} &= \mathbf{F}(\mathbf{x}_{k|k}, \mathbf{u}_k), \\
\mathbf{P}_{k+1|k} &= \mathbf{F}_x \mathbf{P}_{k|k} \mathbf{F}_x^T + \mathbf{F}_u \mathbf{Q}_k \mathbf{F}_u^T, \\
\mathbf{K}_{k+1} &= \mathbf{P}_{k+1|k} \mathbf{H}_x^T (\mathbf{H}_x \mathbf{P}_{k+1|k} \mathbf{H}_x^T + \mathbf{R}_{k+1})^{-1}, \\
\mathbf{x}_{k+1|k+1} &= \mathbf{x}_{k+1|k} \oplus (\mathbf{K}_{k+1} (\mathbf{z}_{k+1} - h(\mathbf{x}_{k+1|k}))), \\
\mathbf{P}_{k+1|k+1} &= (\mathbf{I} - \mathbf{K}_{k+1} \mathbf{H}_x) \mathbf{P}_{k+1|k},
\end{aligned} \tag{51}$$

where,

$$\mathbf{Q} = \text{diag}\{\Sigma_\omega^2, \hat{\Sigma}_f^2\}, \quad \mathbf{R} = \text{diag}\{\Sigma_a^2, \hat{\Sigma}_v^2, \hat{\Sigma}_p^2\}. \tag{52}$$

IV. OBSERVABILITY ANALYSIS

According to the observability analysis method developed in [4], we could analyze the observability of a control affine system by checking the observability rank criterion.

For the rotation stage which is driven by Eq. (9) and observed by Eq. (12), the rotation $\frac{g}{I} \mathbf{q}$ is composed of two observable angles (roll and pitch) and an unobservable yaw angle.

For the translation stage, we first write the process model of the system in control affine form:

$$\dot{\mathbf{x}} = \mathbf{f}_0(\mathbf{x}) + \sum_{i=1}^n \mathbf{f}_i(\mathbf{x}) \mathbf{u}_i. \tag{53}$$

Since the four motor speeds of the quadrotor are integrated into the two inputs U_{ss} and U_s in Eq. (18), our system could be presented:

$$\begin{aligned}
\dot{\mathbf{x}} &= \mathbf{f}_0(\mathbf{x}) + \mathbf{f}_1(\mathbf{x}) U_{ss} + \mathbf{f}_2(\mathbf{x}) U_s \\
&= \begin{bmatrix} \mathcal{G} \mathbf{v}_B^g \\ -\mathcal{G} \mathbf{g} \\ \mathbf{0}_{10 \times 1} \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{3 \times 1} \\ \frac{1}{m} \mathcal{G} \mathbf{R} \tau \mathbf{e}_3 \\ \mathbf{0}_{10 \times 1} \end{bmatrix} U_{ss} + \begin{bmatrix} \mathbf{0}_{3 \times 1} \\ -\frac{1}{m} \mathcal{G} \mathbf{R} D_B^g \mathbf{R}^\top \mathcal{G} \mathbf{v}_B^g \\ \mathbf{0}_{10 \times 1} \end{bmatrix} U_s.
\end{aligned} \tag{54}$$

And then, we simplify the three measurement models Eq. (22, 25, 28) as:

$$\begin{aligned}
\mathbf{h}_d &= \frac{1}{m} \mathcal{I}_B \mathbf{R} (\tau U_{ss} \mathbf{e}_3 - U_s D_B^g \mathbf{R}^\top \mathcal{G} \mathbf{v}_B^g) + \mathcal{I} \hat{\boldsymbol{\omega}} \times (\mathcal{I} \hat{\boldsymbol{\omega}} \times \mathcal{I} \mathbf{t}_B^{\mathcal{I}}) + \mathcal{I} \hat{\boldsymbol{\alpha}} \times \mathcal{I} \mathbf{t}_B^{\mathcal{I}}, \\
\mathbf{h}_v &= \mathcal{G} \mathbf{v}_B^g - \frac{g}{I} \mathbf{R} (\mathcal{I} \hat{\boldsymbol{\omega}} \times \mathcal{I} \mathbf{t}_B^{\mathcal{I}}), \\
\mathbf{h}_p &= \mathcal{G} \mathbf{p}_B^g - \frac{g}{I} \mathbf{R} \mathcal{I} \mathbf{t}_B^{\mathcal{I}}.
\end{aligned} \tag{55}$$

We use Lie derivatives to quantify the impact of changes in the control input U_{ss} and U_s on the output functions \mathbf{h}_d , \mathbf{h}_v and \mathbf{h}_p .

$$\begin{aligned}
L^0 \mathbf{h} &= \mathbf{h}, \\
L_{\mathbf{f}_{i_1}, \mathbf{f}_{i_2}, \dots, \mathbf{f}_{i_{k+1}}}^{k+1} \mathbf{h} &= \nabla_{\mathbf{x}} \left(L_{\mathbf{f}_{i_1}, \mathbf{f}_{i_2}, \dots, \mathbf{f}_{i_k}}^k \mathbf{h} \right) \mathbf{f}_{i_{k+1}}
\end{aligned} \tag{56}$$

We stack vertically several Lie derivatives of the unforced vector field \mathbf{f}_0 and the control input vector fields \mathbf{f}_1 and \mathbf{f}_2 into a vector \mathcal{O} :

$$\mathcal{O} = \begin{bmatrix} \mathbf{h}_d \\ \mathbf{h}_v \\ \mathbf{h}_p \\ L_{\mathbf{f}_0} \mathbf{h}_d \\ L_{\mathbf{f}_0} \mathbf{h}_p \\ L_{\mathbf{f}_1} \mathbf{h}_d \\ L_{\mathbf{f}_1} \mathbf{h}_v \\ L_{\mathbf{f}_2} \mathbf{h}_v \end{bmatrix}, \tag{57}$$

and calculate its gradients as observability matrix $\nabla_x \mathcal{O}$. Finally, we can evaluate whether the system is locally observable by checking the observability rank criterion. In general, the rank is:

$$\text{rank}\{\nabla_x \mathcal{O}\} = 11. \quad (58)$$

However, under some certain conditions the matrix has a rank deficiency. Some of these cases are when:

$$\begin{cases} {}^{\mathcal{I}}\hat{\boldsymbol{\omega}} = \mathbf{0}_{3 \times 1}, {}^{\mathcal{I}}\mathbf{t}_{\mathcal{B}}^{\mathcal{I}} \text{ is unobservable;} \\ {}^{\mathcal{G}}\mathbf{v}_{\mathcal{B}}^{\mathcal{G}} = \mathbf{0}_{3 \times 1}, D \text{ and } {}^{\mathcal{I}}\mathbf{R} \text{ are unobservable;} \\ {}^{\mathcal{G}}\mathbf{v}_{\mathcal{B}}^{\mathcal{G}} = [0; 0; *], \text{yaw}({}^{\mathcal{I}}\mathbf{R}) \text{ is unobservable,} \end{cases} \quad (59)$$

which means the corresponding parameters to be estimated may not converge under the above conditions.

REFERENCES

- [1] J. Svacha, J. Paulos, G. Loianno, and V. Kumar, “Imu-based inertia estimation for a quadrotor using newton-euler dynamics,” *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 3861–3867, 2020.
- [2] R. Mahony, T. Hamel, and J.-M. Pflimlin, “Nonlinear complementary filters on the special orthogonal group,” *IEEE Transactions on automatic control*, vol. 53, no. 5, pp. 1203–1218, 2008.
- [3] S. O. Madgwick, A. J. Harrison, and R. Vaidyanathan, “Estimation of imu and marg orientation using a gradient descent algorithm,” in *2011 IEEE international conference on rehabilitation robotics*. IEEE, 2011, pp. 1–7.
- [4] R. Hermann and A. Krener, “Nonlinear controllability and observability,” *IEEE Transactions on automatic control*, vol. 22, no. 5, pp. 728–740, 1977.