GPS-aided Visual-Inertial Navigation in Large-scale Environments

Woosik Lee - woosik@udel.edu
Kevin Eckenhoff - keck@udel.edu
Patrick Geneva - pgeneva@udel.edu
Guoquan Huang - ghuang@udel.edu

Department of Mechanical Engineering
University of Delaware, Delaware, USA

RPNG

Robot Perception and Navigation Group (RPNG)
Tech Report - RPNG-2019-GPSVIO
Last Updated - September 20, 2019

Contents

1 Preliminaries

1.1 Frames

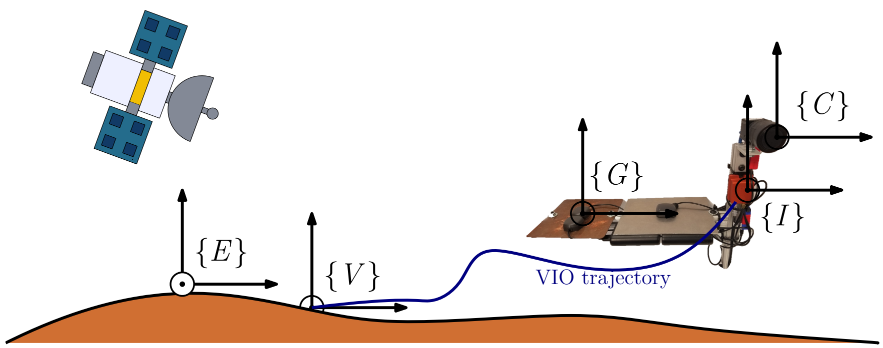

Figure 1: Frames used in this report

Our integrated sensor system is composed of five different frames: East-North-Up (ENU) frame $\{E\}$, VIO frame $\{V\}$, IMU sensor frame $\{I\}$, camera sensor frame $\{C\}$, and GPS sensor frame $\{G\}$. Please see Fig. 1 for a diagram of all frames. $\{E\}$ is the frame of the reference of the GPS measurements that its position is commonly chosen by setting the first GPS measurement as the datum, and $\{V\}$ is the local frame set up by the visual inertial odometry (VIO) system. Both the orientations of the frames are aligned with gravity.

1.2 MSCKF based VIO

The standard VIO state $\mathbf{x}_k$ at timestep k consists of the current IMU state $\mathbf{x}_{I_k}$ and n historical IMU pose clones $\mathbf{x}_{C_k}$ [1]. All states are represented in the arbitrarily chosen gravity aligned frame of reference, $\{V\}$:

$$ {}^V\mathbf{x}_k = \begin{bmatrix} {}^V\mathbf{x}_{I_k}^\top & {}^V\mathbf{x}_{C_k}^\top \end{bmatrix}^\top \tag{1} $$

$$ {}^V\mathbf{x}_{I_k} = \begin{bmatrix} {}^{I_k}_{V}\bar{q}^\top & {}^V\mathbf{p}_{I_k}^\top & {}^V\mathbf{v}_{I_k}^\top & \mathbf{b}_{\omega_k}^\top & \mathbf{b}_{a_k}^\top \end{bmatrix}^\top \tag{2} $$

$$ {}^V\mathbf{x}_{C_k} = \begin{bmatrix} {}^{I_{k-1}}_{V}\bar{q}^\top & {}^V\mathbf{p}_{I_{k-1}}^\top & \cdots & {}^{I_{k-n}}_{V}\bar{q}^\top & {}^V\mathbf{p}_{I_{k-n}}^\top \end{bmatrix}^\top \tag{3} $$

where ${}^{I_k}_{V}\bar{q}$ is the JPL unit quaternion [2] corresponding to the rotation from $\{V\}$ to $\{I\}$ (i.e., rotation matrix ${}^{I_k}_{V}\mathbf{R}$), ${}^V\mathbf{p}_{I_k}$ and ${}^V\mathbf{v}_{I_k}$ are the position and velocity of $\{I\}$ in $\{V\}$, and $\mathbf{b}_{\omega_k}$ and $\mathbf{b}_{a_k}$ are the gyroscope and accelerometer biases, respectively. We define $\mathbf{x} = \hat{\mathbf{x}} \boxplus \tilde{\mathbf{x}}$, where $\mathbf{x}$ is the true state, $\hat{\mathbf{x}}$ is its estimate, $\tilde{\mathbf{x}}$ is the error state, and the operation $\boxplus$ which maps a manifold element and its correction vector to an updated element on the same manifold [3].

1.2.1 IMU State Propagation

The linear acceleration $\mathbf{a}_m$ and angular velocity $\boldsymbol{\omega}_m$ measurements of the IMU are used for ${}^V\mathbf{x}_{I_k}$ propagation:

$$ \mathbf{a}_m = \mathbf{a} + {}^I_V\mathbf{R}\mathbf{g} + \mathbf{b}_a + \mathbf{n}_a \tag{4} $$

$$ \boldsymbol{\omega}_m = \boldsymbol{\omega} + \mathbf{b}_\omega + \mathbf{n}_\omega \tag{5} $$

where $\mathbf{a}$ and $\boldsymbol{\omega}$ are true acceleration and angular velocity, $\mathbf{g} \approx [0\ 0\ 9.81]^\top$ is the global gravity, and $\mathbf{n}_a$ and $\mathbf{n}_\omega$ are zero mean Gaussian noises. These measurements are used to propagate the IMU state from timestep k to $k+1$ based on the following generic nonlinear kinematic model [2]:

$$ {}^V\hat{\mathbf{x}}_{I_{k+1|k}} = f({}^V\hat{\mathbf{x}}_{I_{k|k}}, \mathbf{a}_{m_k}, \boldsymbol{\omega}_{m_k}) \tag{6} $$

where $\hat{\mathbf{x}}_{I_{a|b}}$ denotes the estimate at timestep a processing the measurements up to timestep b.

Specifically, given the continuous-time measurements $\boldsymbol{\omega}_m(t)$ and $\mathbf{a}_m(t)$ in the time interval $t \in [t_k, t_{k+1}]$, and their estimates, i.e. after taking the expectation, $\hat{\boldsymbol{\omega}}(t) = \boldsymbol{\omega}_m(t) - \hat{\mathbf{b}}_\omega(t)$ and $\hat{\boldsymbol{a}}(t) = \boldsymbol{a}_m(t) -{}^I_V \hat{\mathbf{R}}(t)\mathbf{g} - \hat{\mathbf{b}}_a(t)$, we can drive the propagation of IMU state from differential kinematic equations. The solution to the quaternion evolution has the following general form:

$$ {}^{I_t}_V\bar{q} = \boldsymbol{\Theta}(t, t_k){}^{I_k}_V\bar{q} \tag{7} $$

Differentiating and reordering the terms yields the governing equation for $\boldsymbol{\Theta}(t, t_k)$ as

$$ \boldsymbol{\Theta}(t, t_k) = {}^{I_t}_V\bar{q}\ {}^{I_k}_V\bar{q}^{-1} \tag{8} $$

$$ \Rightarrow \dot{\boldsymbol{\Theta}}(t, t_k) = {}^{I_t}_V\dot{\bar{q}}\ {}^{I_k}_V\bar{q}^{-1} \tag{9} $$

$$ = \frac{1}{2}\boldsymbol{\Omega}(\boldsymbol{\omega}(t))\ {}^{I_t}_V\bar{q}\ {}^{I_k}_V\bar{q}^{-1} \tag{10} $$

$$ = \frac{1}{2}\boldsymbol{\Omega}(\boldsymbol{\omega}(t))\boldsymbol{\Theta}(t, t_k) \tag{11} $$

where $\boldsymbol{\Omega}(\boldsymbol{\omega}) = \begin{bmatrix} -\lfloor \boldsymbol{\omega}\times \rfloor & \boldsymbol{\omega} \\ -\boldsymbol{\omega}^\top & 0 \end{bmatrix}$, $\lfloor \boldsymbol{\omega}\times \rfloor$ is the skew symmetric matrix, and $\boldsymbol{\Theta}(t_k, t_k) = \mathbf{I}_4$. If we take $\boldsymbol{\omega}(t) = \boldsymbol{\omega}$ to be constant over the the period $\Delta t = t_{k+1} - t_k$, then $\boldsymbol{\Theta}$ can be expressed as [4]:

$$ \boldsymbol{\Theta}(t_{k+1}, t_k) = \exp\left(\frac{1}{2}\boldsymbol{\Omega}(\boldsymbol{\omega})\Delta t\right) \tag{12} $$

$$ = \cos\left(\frac{|\boldsymbol{\omega}|}{2}\Delta t\right) \cdot \mathbf{I}_4 + \frac{1}{|\boldsymbol{\omega}|}\sin\left(\frac{|\boldsymbol{\omega}|}{2}\Delta t\right) \cdot \boldsymbol{\Omega}(\boldsymbol{\omega}) \tag{13} $$

$$ \simeq \mathbf{I}_4 + \frac{\Delta t}{2}\boldsymbol{\Omega}(\boldsymbol{\omega}) \tag{14} $$

where in the case of a small $|\boldsymbol{\omega}|$ the last equation can be used to approximate. We can formulate the quaternion propagation from t_k to t_{k+1} using the estimated rotational velocity $\hat{\boldsymbol{\omega}}(t) = \hat{\boldsymbol{\omega}}$ as:

$$ {}^{I_{k+1}}_V\hat{\bar{q}} = \exp\left(\frac{1}{2}\boldsymbol{\Omega}(\hat{\boldsymbol{\omega}})\Delta t\right){}^{I_k}_V\hat{\bar{q}} \tag{15} $$

Having defined the integration of the orientation, we can integrate the position and velocity over the measurement interval:

$$ {}^V\hat{\mathbf{p}}_{I_{k+1}} = {}^V\hat{\mathbf{p}}_{I_k} + \int_{t_k}^{t_{k+1}} {}^V\hat{\mathbf{v}}_I(\tau)d\tau \tag{16} $$

$$ = {}^V\hat{\mathbf{p}}_{I_k} + {}^V\hat{\mathbf{v}}_{I_k}\Delta t - \frac{1}{2}\mathbf{g}\Delta t^2 + \int_{t_k}^{t_{k+1}}\int_{t_k}^{s} {}^V_{I_\tau}\hat{\mathbf{R}}\mathbf{a}_m(\tau)d\tau ds \tag{17} $$

$$ = {}^V\hat{\mathbf{p}}_{I_k} + {}^V\hat{\mathbf{v}}_{I_k}\Delta t - \frac{1}{2}\mathbf{g}\Delta t^2 + {}^{I_k}_V\hat{\mathbf{R}}^\top \int_{t_k}^{t_{k+1}}\int_{t_k}^{s} {}^{I_k}_{I_\tau}\hat{\mathbf{R}}\mathbf{a}_m(\tau)d\tau ds \tag{18} $$

$$= {}^V\hat{\mathbf{p}}_{I_k} + {}^V\hat{\mathbf{v}}_{I_k}\Delta t - \frac{1}{2}\mathbf{g}\Delta t^2 + {}^{I_k}_V\hat{\mathbf{R}}^\top \boldsymbol{\alpha} \tag{19}$$

$$\begin{aligned} {}^V\hat{\mathbf{v}}_{I_{k+1}} &= {}^V\hat{\mathbf{v}}_{I_k} + \int_{t_k}^{t_{k+1}} {}^V\hat{\mathbf{a}}(\tau)d\tau & (20)\\ &= {}^V\hat{\mathbf{v}}_{I_k} - \mathbf{g}\Delta t + \int_{t_k}^{t_{k+1}} {}^V_{I_\tau}\hat{\mathbf{R}}\mathbf{a}_m(\tau)d\tau & (21)\\ &= {}^V\hat{\mathbf{v}}_{I_k} - \mathbf{g}\Delta t + {}^{I_k}_V\hat{\mathbf{R}}^\top \int_{I_k}^{t_{k+1}} {}^{t_k}_{I_\tau}\hat{\mathbf{R}}\mathbf{a}_m(\tau)d\tau & (22)\\ &= {}^V\hat{\mathbf{v}}_{I_k} - \mathbf{g}\Delta t + {}^{I_k}_V\hat{\mathbf{R}}^\top \boldsymbol{\beta} & (23) \end{aligned}$$

$$\hat{\mathbf{b}}_{\omega_{k+1}} = \hat{\mathbf{b}}_{\omega_k} \tag{24}$$

$$\hat{\mathbf{b}}_{a_{k+1}} = \hat{\mathbf{b}}_{a_k} \tag{25}$$

All of the above integrals could be solved or numerically solved for if one wishes to use the continuous-time measurement evolution model.

To perform covariance propagation forward in time, we linearize Eq. (6) at the current estimate. For the orientation error propagation, we use the following approximations:

$$\begin{aligned} {}^I_V\mathbf{R} &\approx (\mathbf{I}_3 - \lfloor {}^I_V\tilde{\boldsymbol{\theta}}\times\rfloor){}^I_V\hat{\mathbf{R}} & (26)\\ {}^{I_{k+1}}_{I_k}\mathbf{R} &= \exp({}^{I_{k+1}}_{I_k}\boldsymbol{\theta}) & (27)\\ &= \exp({}^{I_{k+1}}_{I_k}\hat{\boldsymbol{\theta}} + {}^{I_{k+1}}_{I_k}\tilde{\boldsymbol{\theta}}) & (28)\\ &\approx \exp({}^{I_{k+1}}_{I_k}\hat{\boldsymbol{\theta}})\exp(\mathbf{J_r}({}^{I_{k+1}}_{I_k}\hat{\boldsymbol{\theta}}){}^{I_{k+1}}_{I_k}\tilde{\boldsymbol{\theta}}) & (29)\\ &\approx {}^{I_{k+1}}_{I_k}\hat{\mathbf{R}}(\mathbf{I}_3 - \lfloor \mathbf{J_r}({}^{I_{k+1}}_{I_k}\hat{\boldsymbol{\theta}}){}^{I_{k+1}}_{I_k}\tilde{\boldsymbol{\theta}}\times\rfloor) & (30) \end{aligned}$$

where $\mathbf{J_r}$ is right Jacobian matrix of SO(3), and $\exp(\cdot)$ is the SO(3) matrix exponential [5]. The the orientation error propagation become:

$$\begin{aligned} {}^{I_{k+1}}_V\mathbf{R} &= {}^{I_{k+1}}_{I_k}\mathbf{R}{}^{I_k}_V\mathbf{R} & (31)\\ \Rightarrow (\mathbf{I}_3 - \lfloor {}^{I_{k+1}}_V\tilde{\boldsymbol{\theta}}\times\rfloor){}^{I_{k+1}}_V\hat{\mathbf{R}} &\approx {}^{I_{k+1}}_{I_k}\hat{\mathbf{R}}(\mathbf{I}_3 - \lfloor \mathbf{J_r}({}^{I_{k+1}}_{I_k}\hat{\boldsymbol{\theta}}){}^{I_{k+1}}_{I_k}\tilde{\boldsymbol{\theta}}\times\rfloor)(\mathbf{I}_3 - \lfloor {}^{I_k}_V\tilde{\boldsymbol{\theta}}\times\rfloor){}^{I_k}_V\hat{\mathbf{R}} & (32)\\ \Rightarrow -\lfloor {}^{I_{k+1}}_V\tilde{\boldsymbol{\theta}}\times\rfloor{}^{I_{k+1}}_V\hat{\mathbf{R}} &\approx -{}^{I_{k+1}}_{I_k}\hat{\mathbf{R}}\lfloor \mathbf{J_r}({}^{I_{k+1}}_{I_k}\hat{\boldsymbol{\theta}}){}^{I_{k+1}}_{I_k}\tilde{\boldsymbol{\theta}}\times\rfloor{}^{I_k}_V\hat{\mathbf{R}} - {}^{I_{k+1}}_{I_k}\hat{\mathbf{R}}\lfloor {}^{I_k}_V\tilde{\boldsymbol{\theta}}\times\rfloor{}^{I_k}_V\hat{\mathbf{R}} & (33)\\ &\quad + {}^{I_{k+1}}_{I_k}\hat{\mathbf{R}}\lfloor \mathbf{J_r}({}^{I_{k+1}}_{I_k}\hat{\boldsymbol{\theta}}){}^{I_{k+1}}_{I_k}\tilde{\boldsymbol{\theta}}\times\rfloor\lfloor {}^{I_k}_V\tilde{\boldsymbol{\theta}}\times\rfloor{}^{I_k}_V\hat{\mathbf{R}} & (34)\\ \Rightarrow -\lfloor {}^{I_{k+1}}_V\tilde{\boldsymbol{\theta}}\times\rfloor{}^{I_{k+1}}_V\hat{\mathbf{R}} &\approx -{}^{I_{k+1}}_{I_k}\hat{\mathbf{R}}\lfloor \mathbf{J_r}({}^{I_{k+1}}_{I_k}\hat{\boldsymbol{\theta}}){}^{I_{k+1}}_{I_k}\tilde{\boldsymbol{\theta}}\times\rfloor{}^{I_k}_V\hat{\mathbf{R}} - {}^{I_{k+1}}_{I_k}\hat{\mathbf{R}}\lfloor {}^{I_k}_V\tilde{\boldsymbol{\theta}}\times\rfloor{}^{I_k}_V\hat{\mathbf{R}} & (35)\\ \Rightarrow {}^{I_{k+1}}_V\tilde{\boldsymbol{\theta}} &\approx {}^{I_{k+1}}_{I_k}\hat{\mathbf{R}}\mathbf{J_r}({}^{I_{k+1}}_{I_k}\hat{\boldsymbol{\theta}}){}^{I_{k+1}}_{I_k}\tilde{\boldsymbol{\theta}} + {}^{I_{k+1}}_{I_k}\hat{\mathbf{R}}{}^{I_k}_V\tilde{\boldsymbol{\theta}} & (36) \end{aligned}$$

Thus the state transition of orientation error becomes ${}^{I_{k+1}}_{I_k}\hat{\mathbf{R}}$. Since the orientation error is not related to ${}^V\mathbf{p}_I, {}^V\mathbf{v}_I$ the error transitions for these parameters are zero.

For the error propagation of regular vectors, such as ${}^V\mathbf{p}_{I_k}$, ${}^V\mathbf{v}_{I_k}$, and $\mathbf{b}$, we define error state as $\mathbf{x} = \hat{\mathbf{x}} + \tilde{\mathbf{x}}$ where $\mathbf{x}$ is the true state, $\hat{\mathbf{x}}$ is its estimate, $\tilde{\mathbf{x}}$ is the error state. Then the position error propagation:

$${}^V\mathbf{p}_{I_{k+1}} = {}^V\mathbf{p}_{I_k} + {}^V\mathbf{v}_{I_k}\Delta t - \frac{1}{2}\mathbf{g}\Delta t^2 + {}^{I_k}_V\mathbf{R}^\top \boldsymbol{\alpha} \tag{37}$$

$$\Rightarrow {}^V\hat{\mathbf{p}}_{I_{k+1}} + {}^V\tilde{\mathbf{p}}_{I_{k+1}} = {}^V\hat{\mathbf{p}}_{I_k} + {}^V\tilde{\mathbf{p}}_{I_k} + {}^V\hat{\mathbf{v}}_{I_k}\Delta t + {}^V\tilde{\mathbf{v}}_{I_k}\Delta t - \frac{1}{2}\mathbf{g}\Delta t^2 + {}^{I_k}_V\hat{\mathbf{R}}^\top(\mathbf{I}_3 + \lfloor {}^{I_k}_V\tilde{\boldsymbol{\theta}}\times\rfloor)\boldsymbol{\alpha} \tag{38}$$

$$\Rightarrow {}^V\tilde{\mathbf{p}}_{I_{k+1}} = {}^V\tilde{\mathbf{p}}_{I_k} + {}^V\tilde{\mathbf{v}}_{I_k}\Delta t + {}^{I_k}_V\hat{\mathbf{R}}^\top\lfloor {}^{I_k}_V\tilde{\boldsymbol{\theta}}\times\rfloor\boldsymbol{\alpha} \tag{39}$$

$$= {}^V\tilde{\mathbf{p}}_{I_k} + {}^V\tilde{\mathbf{v}}_{I_k}\Delta t - {}^{I_k}_V\hat{\mathbf{R}}^\top\lfloor \boldsymbol{\alpha}\times\rfloor {}^{I_k}_V\tilde{\boldsymbol{\theta}} \tag{40}$$

$$= {}^V\tilde{\mathbf{p}}_{I_k} + {}^V\tilde{\mathbf{v}}_{I_k}\Delta t - \lfloor {}^{I_k}_V\hat{\mathbf{R}}^\top\boldsymbol{\alpha}\times\rfloor {}^{I_k}_V\hat{\mathbf{R}}^\top {}^{I_k}_V\tilde{\boldsymbol{\theta}} \tag{41}$$

$$= {}^V\tilde{\mathbf{p}}_{I_k} + {}^V\tilde{\mathbf{v}}_{I_k}\Delta t - \lfloor {}^V\hat{\mathbf{p}}_{I_{k+1}} - {}^V\hat{\mathbf{p}}_{I_k} - {}^V\hat{\mathbf{v}}_{I_k}\Delta t + \frac{1}{2}\mathbf{g}\Delta t^2\times\rfloor {}^{I_k}_V\hat{\mathbf{R}}^\top {}^{I_k}_V\tilde{\boldsymbol{\theta}} \tag{42}$$

where $\Delta t = t_{k+1} - t_k$. Therefore the transition of position error from ${}^V\mathbf{p}_{I_k}$ is $\mathbf{I}_3$, from orientation $\lfloor {}^V\hat{\mathbf{p}}_{I_{k+1}} - {}^V\hat{\mathbf{p}}_{I_k} - {}^V\hat{\mathbf{v}}_{I_k}\Delta t + \frac{1}{2}\mathbf{g}\Delta t^2\times\rfloor {}^{I_k}_V\hat{\mathbf{R}}^\top$, from velocity $\Delta t\mathbf{I}_3$.

The velocity error propagation is as follows:

$${}^V\mathbf{v}_{I_{k+1}} = {}^V\mathbf{v}_{I_k} - \mathbf{g}\Delta t + {}^{I_k}_V\mathbf{R}^\top\boldsymbol{\beta} \tag{43}$$

$$\Rightarrow {}^V\hat{\mathbf{v}}_{I_{k+1}} + {}^V\tilde{\mathbf{v}}_{I_{k+1}} = {}^V\hat{\mathbf{v}}_{I_k} + {}^V\tilde{\mathbf{v}}_{I_k} - \mathbf{g}\Delta t + {}^{I_k}_V\hat{\mathbf{R}}^\top(\mathbf{I}_3 + \lfloor {}^{I_k}_V\tilde{\boldsymbol{\theta}}\times\rfloor)\boldsymbol{\beta} \tag{44}$$

$$\Rightarrow {}^V\tilde{\mathbf{v}}_{I_{k+1}} = {}^V\tilde{\mathbf{v}}_{I_k} + {}^{I_k}_V\hat{\mathbf{R}}^\top\lfloor {}^{I_k}_V\tilde{\boldsymbol{\theta}}\times\rfloor\boldsymbol{\beta} \tag{45}$$

$$= {}^V\tilde{\mathbf{v}}_{I_k} - {}^{I_k}_V\hat{\mathbf{R}}^\top\lfloor\boldsymbol{\beta}\times\rfloor {}^{I_k}_V\tilde{\boldsymbol{\theta}} \tag{46}$$

$$= {}^V\tilde{\mathbf{v}}_{I_k} - \lfloor {}^{I_k}_V\hat{\mathbf{R}}^\top\boldsymbol{\beta}\times\rfloor {}^{I_k}_V\hat{\mathbf{R}}^\top {}^{I_k}_V\tilde{\boldsymbol{\theta}} \tag{47}$$

$$= {}^V\tilde{\mathbf{v}}_{I_k} - \lfloor {}^V\hat{\mathbf{v}}_{I_{k+1}} - {}^V\hat{\mathbf{v}}_{I_k} + \mathbf{g}\Delta t\times\rfloor {}^{I_k}_V\hat{\mathbf{R}}^\top {}^{I_k}_V\tilde{\boldsymbol{\theta}} \tag{48}$$

Therefore the error transition from velocity becomes $\mathbf{I}_3$ and the transition of velocity from orientation $\lfloor {}^V\hat{\mathbf{v}}_{I_{k+1}} - {}^V\hat{\mathbf{v}}_{I_k} + \mathbf{g}\Delta t\times\rfloor {}^{I_k}_V\hat{\mathbf{R}}^\top$.

For the biases:

$$\hat{\mathbf{b}}_{\omega_{k+1}} + \tilde{\mathbf{b}}_{\omega_{k+1}} = \hat{\mathbf{b}}_{\omega_k} + \tilde{\mathbf{b}}_{\omega_k} \tag{49}$$

$$\Rightarrow \tilde{\mathbf{b}}_{\omega_{k+1}} = \tilde{\mathbf{b}}_{\omega_k} \tag{50}$$

$$\hat{\mathbf{b}}_{a_{k+1}} + \tilde{\mathbf{b}}_{a_{k+1}} = \hat{\mathbf{b}}_{a_k} + \tilde{\mathbf{b}}_{a_k} \tag{51}$$

$$\Rightarrow \tilde{\mathbf{b}}_{a_{k+1}} = \tilde{\mathbf{b}}_{a_k} \tag{52}$$

To summarize, the error state transition matrix of IMU become:

$$\boldsymbol{\Phi}(t_{k+1}, t_k) = \begin{bmatrix} {}^{I_{k+1}}_V\hat{\mathbf{R}} {}^{I_k}_V\hat{\mathbf{R}}^\top & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ -\lfloor {}^V\hat{\mathbf{p}}_{I_{k+1}} - {}^V\hat{\mathbf{p}}_{I_k} - {}^V\hat{\mathbf{v}}_{I_k}\Delta t + \frac{1}{2}\mathbf{g}\Delta t^2\times\rfloor {}^{I_k}_V\hat{\mathbf{R}}^\top & \mathbf{I}_3 & \Delta t\mathbf{I}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ -\lfloor {}^V\hat{\mathbf{v}}_{I_{k+1}} - {}^V\hat{\mathbf{v}}_{I_k} + \mathbf{g}\Delta t\times\rfloor {}^{I_k}_V\hat{\mathbf{R}}^\top & \mathbf{0}_3 & \mathbf{I}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{I}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{I}_3 \end{bmatrix} \tag{53}$$

We use the transition matrix to propagate the covariance matrix of IMU state as:

$$\mathbf{P}_{k+1|k} = \boldsymbol{\Phi}(t_{k+1}, t_k)\mathbf{P}_{k|k}\boldsymbol{\Phi}(t_{k+1}, t_k)^\top + \mathbf{Q}_k \tag{54}$$

where $\mathbf{P}$ and $\mathbf{Q}$ are discrete state and noise covariance, respectively [1].

1.2.2 Visual Measurement Update

We maintain a number of stochastic clones in ${}^V\mathbf{x}_{C_k}$, and perform visual feature tracking to obtain series of visual bearing measurements to 3D environmental features. A measurement $\mathbf{z}_{c_i}$ at timestep

i is expressed as a function of a cloned pose and feature position ${}^V\mathbf{p}_f$:

$$\mathbf{z}_{c_i} = \mathbf{\Pi}({}^{C_i}\mathbf{p}_f) + \mathbf{n}_i \tag{55}$$

$$\mathbf{\Pi}\left([x\ y\ z]^\top\right) = \begin{bmatrix}\frac{x}{z} & \frac{y}{z}\end{bmatrix}^\top \tag{56}$$

$${}^{C_i}\mathbf{p}_f = {}^C_I\mathbf{R}{}^{I_i}_V\mathbf{R}\left({}^V\mathbf{p}_f - {}^V\mathbf{p}_{I_i}\right) + {}^C\mathbf{p}_I \tag{57}$$

where ${}^C_I\mathbf{R}$ and ${}^C\mathbf{p}_I$ represent the camera to IMU extrinsics. To get an estimate of ${}^V\hat{\mathbf{p}}_f$, triangulation is performed using the current state estimates. Then we compute the Jacobian matrix by linearizing Eq. (55) at current estimate and feature position ${}^V\hat{\mathbf{p}}_f = \begin{bmatrix}{}^V\hat{x}_f & {}^V\hat{y}_f & {}^V\hat{z}_f\end{bmatrix}^\top$. The Jacobian matrix respect to ${}^{I_i}_V\bar{q}$, ${}^V\mathbf{p}_{I_i}$, and ${}^V\mathbf{p}_f$ for the update are:

$$\frac{\partial\tilde{\mathbf{z}}_{c_i}}{\partial{}^{I_i}_V\tilde{\boldsymbol{\theta}}} = \frac{\partial\tilde{\mathbf{z}}_{c_i}}{\partial{}^{C_i}\tilde{\mathbf{p}}_f}\frac{\partial{}^{C_i}\tilde{\mathbf{p}}_f}{\partial{}^{I_i}_V\tilde{\boldsymbol{\theta}}} \qquad \frac{\partial\tilde{\mathbf{z}}_{c_i}}{\partial{}^V\tilde{\mathbf{p}}_{I_i}} = \frac{\partial\tilde{\mathbf{z}}_{c_i}}{\partial{}^{C_i}\tilde{\mathbf{p}}_f}\frac{\partial{}^{C_i}\tilde{\mathbf{p}}_f}{\partial{}^V\tilde{\mathbf{p}}_{I_i}} \tag{58}$$

$$\frac{\partial\tilde{\mathbf{z}}_{c_i}}{\partial{}^V\tilde{\mathbf{p}}_f} = \frac{\partial\tilde{\mathbf{z}}_{c_i}}{\partial{}^{C_i}\tilde{\mathbf{p}}_f}\frac{\partial{}^{C_i}\tilde{\mathbf{p}}_f}{\partial{}^V\tilde{\mathbf{p}}_f} \qquad \frac{\partial\tilde{\mathbf{z}}_{c_i}}{\partial{}^{C_i}\tilde{\mathbf{p}}_f} = \begin{bmatrix}\frac{1}{{}^{C_i}\hat{z}_f} & 0 & -\frac{{}^{C_i}\hat{x}_f}{{}^{C_i}\hat{z}_f^2}\\ 0 & \frac{1}{{}^{C_i}\hat{z}_f} & -\frac{{}^{C_i}\hat{y}_f}{{}^{C_i}\hat{z}_f^2}\end{bmatrix} \tag{59}$$

$$\frac{\partial{}^{C_i}\tilde{\mathbf{p}}_f}{\partial{}^{I_i}_V\tilde{\boldsymbol{\theta}}} = {}^C_I\hat{\mathbf{R}}\lfloor{}^{I_i}_V\hat{\mathbf{R}}({}^V\hat{\mathbf{p}}_f - {}^V\hat{\mathbf{p}}_{I_i})\times\rfloor \qquad \frac{\partial{}^{C_i}\tilde{\mathbf{p}}_f}{\partial{}^V\tilde{\mathbf{p}}_{I_i}} = -{}^C_I\hat{\mathbf{R}}{}^{I_i}_V\hat{\mathbf{R}} \tag{60}$$

$$\frac{\partial{}^{C_i}\tilde{\mathbf{p}}_f}{\partial{}^V\tilde{\mathbf{p}}_f} = {}^C_I\hat{\mathbf{R}}{}^{I_i}_V\hat{\mathbf{R}} \tag{61}$$

Stacking the Jacobians and residuals for all visual measurements yields the following general form:

$$\tilde{\mathbf{z}}_c = \mathbf{H}_x{}^V\tilde{\mathbf{x}}_k + \mathbf{H}_f{}^V\tilde{\mathbf{p}}_f + \mathbf{n}_f \tag{62}$$

where $\tilde{\mathbf{z}}_c$ is formed by stacking the individual measurement residuals for a given feature, $\mathbf{H}_x$ and $\mathbf{H}_f$ are the state and feature Jacobians, respectively. Either the feature can now be updated using the standard EKF update or treated as a MSCKF feature [6]. The key idea of the MSCKF is to find the matrix $\mathcal{N}(\mathbf{H}_f^\top)$ whose columns span the left null space of $\mathbf{H}_f$. Multiplying the above linear system on the left by $\mathcal{N}(\mathbf{H}_f^\top)^\top$, we obtain a new measurement function that depends only on the state:

$$\tilde{\mathbf{z}}'_c = \mathbf{H}'_x{}^V\tilde{\mathbf{x}}_k + \mathbf{n}'_f \tag{63}$$

We can directly use this measurement in an EKF update without storing features in the state. This leads to substantial computational savings as the problem size remains bounded over the entire trajectory.

2 GPS Measurement Update

In addition to the visual measurement update, whenever a new GPS measurement in the ENU frame $\{E\}$ is available, we will update the state with it. When the state is in $\{V\}$, the GPS measurement ${}^E\mathbf{p}_{G_k}$ at timestep k can be modeled as:

$$\mathbf{z}_{g_k} := {}^E\mathbf{p}_{G_k} = {}^E\mathbf{p}_V + {}^E_V\mathbf{R}{}^V\mathbf{p}_{G_k} + \mathbf{n}_{g_k} =: \mathbf{h}({}^V\mathbf{x}_k) + \mathbf{n}_{g_k} \tag{64}$$

$${}^V\mathbf{p}_{G_k} = {}^V\mathbf{p}_{I_k} + {}^{I_k}_V\mathbf{R}^{\top I}\mathbf{p}_G \tag{65}$$

where ${}^{I}\mathbf{p}_G$ is the GPS to IMU extrinsic calibration and $\mathbf{n}_{g_k}$ is a white Gaussian noise. Due to the delayed asynchronous nature of the GPS sensor, the state has likely advanced beyond the collection time and thus we express the measurement as a function of the available stochastic clones. Using linear interpolation [6], the IMU pose in Eq. (65) can be expressed as:

$$ {}^{I_k}_{V}\mathbf{R} = \mathrm{Exp}\left(\lambda \ \mathrm{Log}\left({}^{I_b}_{V}\mathbf{R}{}^{I_a}_{V}\mathbf{R}^\top\right)\right){}^{I_a}_{V}\mathbf{R} \tag{66} $$

$$ {}^{V}\mathbf{p}_{I_k} = (1-\lambda){}^{V}\mathbf{p}_{I_a} + \lambda{}^{V}\mathbf{p}_{I_b} \tag{67} $$

$$ \lambda = (t_k + {}^{I}t_G - t_a)/(t_b - t_a) \tag{68} $$

where ${}^{I}t_G$ is the time offset between the GPS and IMU clocks, the bounding poses have timestamps $t_a \le (t_k + {}^{I}t_G) \le t_b$, and $\mathrm{Exp}(\cdot)$, $\mathrm{Log}(\cdot)$ are the SO(3) matrix exponential and logarithmic functions [5].

As evident from Eqs. (64)-(68), the GPS measurement model depends on both the IMU states and the GPS-IMU extrinsic and time offset, thus enabling online spatiotemporal GPS-IMU calibration. To update with this measurement in the MSCKF, we include the extrinsic and timeoffset parameters in the state as:

$$ {}^{V}\mathbf{x}_k = \begin{bmatrix} {}^{V}\mathbf{x}_{I_k}^\top & {}^{V}\mathbf{x}_{C_k}^\top & {}^{I}\mathbf{p}_G^\top & {}^{I}t_G \end{bmatrix}^\top \tag{69} $$

and compute the Jacobian matrix respect to ${}^{I_a}_{V}\bar{q}$, ${}^{V}\mathbf{p}_{I_a}$, ${}^{I_b}_{V}\bar{q}$, ${}^{V}\mathbf{p}_{I_b}$, ${}^{E}_{V}\bar{q}$, ${}^{E}\mathbf{p}_V$, ${}^{I}\mathbf{p}_G$ and ${}^{I}t_G$ as:

$$ \frac{\partial\tilde{\mathbf{z}}}{\partial{}^{I_a}_{V}\tilde{\boldsymbol{\theta}}} = \frac{\partial\tilde{\mathbf{z}}}{\partial{}^{I_k}_{V}\tilde{\boldsymbol{\theta}}}\frac{\partial{}^{I_k}_{V}\tilde{\boldsymbol{\theta}}}{\partial{}^{I_a}_{V}\tilde{\boldsymbol{\theta}}} \qquad \frac{\partial\tilde{\mathbf{z}}}{\partial{}^{V}\tilde{\mathbf{p}}_{I_a}} = \frac{\partial\tilde{\mathbf{z}}}{\partial{}^{V}\tilde{\mathbf{p}}_{I_k}}\frac{\partial{}^{V}\tilde{\mathbf{p}}_{I_k}}{\partial{}^{V}\tilde{\mathbf{p}}_{I_a}} \tag{70} $$

$$ \frac{\partial\tilde{\mathbf{z}}}{\partial{}^{I_b}_{V}\tilde{\boldsymbol{\theta}}} = \frac{\partial\tilde{\mathbf{z}}}{\partial{}^{I_k}_{V}\tilde{\boldsymbol{\theta}}}\frac{{}^{I_k}_{V}\tilde{\boldsymbol{\theta}}}{\partial{}^{I_b}_{V}\tilde{\boldsymbol{\theta}}} \qquad \frac{\partial\tilde{\mathbf{z}}}{\partial{}^{V}\tilde{\mathbf{p}}_{I_b}} = \frac{\partial\tilde{\mathbf{z}}}{\partial{}^{V}\tilde{\mathbf{p}}_{I_k}}\frac{\partial{}^{V}\tilde{\mathbf{p}}_{I_k}}{\partial{}^{V}\tilde{\mathbf{p}}_{I_b}} \tag{71} $$

$$ \frac{\partial\tilde{\mathbf{z}}}{\partial{}^{E}_{V}\tilde{\boldsymbol{\theta}}} = (\lfloor{}^{E}_{V}\hat{\mathbf{R}}({}^{V}\hat{\mathbf{p}}_{I_k} + {}^{I_i}_{V}\hat{\mathbf{R}}^\top{}^{I}\hat{\mathbf{p}}_G)\times\rfloor)_z \qquad \frac{\partial\tilde{\mathbf{z}}}{\partial{}^{E}\tilde{\mathbf{p}}_V} = \mathbf{I}_3 \tag{72} $$

$$ \frac{\partial\tilde{\mathbf{z}}}{\partial{}^{I}\tilde{\mathbf{p}}_G} = {}^{E}_{V}\hat{\mathbf{R}}{}^{I_k}_{V}\hat{\mathbf{R}}^\top \qquad \frac{\partial\tilde{\mathbf{z}}}{\partial{}^{I}\tilde{t}_G} = \frac{\partial\tilde{\mathbf{z}}}{\partial{}^{I_k}_{V}\tilde{\boldsymbol{\theta}}}\frac{\partial{}^{I_k}_{V}\tilde{\boldsymbol{\theta}}}{\partial\lambda}\frac{\partial\lambda}{\partial{}^{I}\tilde{t}_G} + \frac{\partial\tilde{\mathbf{z}}}{\partial{}^{V}\tilde{\mathbf{p}}_{I_k}}\frac{\partial{}^{V}\tilde{\mathbf{p}}_{I_k}}{\partial\lambda}\frac{\partial\lambda}{\partial{}^{I}\tilde{t}_G} \tag{73} $$

$$ \frac{\partial\tilde{\mathbf{z}}}{\partial{}^{I_k}_{V}\tilde{\boldsymbol{\theta}}} = -{}^{E}_{V}\hat{\mathbf{R}}{}^{I_k}_{V}\hat{\mathbf{R}}^\top\lfloor{}^{I}\hat{\mathbf{p}}_G\times\rfloor \qquad \frac{\partial{}^{I_k}_{V}\tilde{\boldsymbol{\theta}}}{\partial{}^{I_a}_{V}\tilde{\boldsymbol{\theta}}} = \mathbf{J_l}(\hat{\lambda}{}^{b}_{a}\hat{\boldsymbol{\theta}})(\mathbf{J_r}(\hat{\lambda}{}^{b}_{a}\hat{\boldsymbol{\theta}})^{-1} - \hat{\lambda}\mathbf{J_r}({}^{b}_{a}\hat{\boldsymbol{\theta}})^{-1}) \tag{74} $$

$$ \frac{\partial\tilde{\mathbf{z}}}{\partial{}^{V}\tilde{\mathbf{p}}_{I_k}} = {}^{E}_{V}\hat{\mathbf{R}} \qquad \frac{\partial{}^{V}\tilde{\mathbf{p}}_{I_k}}{\partial{}^{V}\tilde{\mathbf{p}}_{I_a}} = (1-\hat{\lambda})\mathbf{I}_3 \tag{75} $$

$$ \frac{\partial{}^{I_k}_{V}\tilde{\boldsymbol{\theta}}}{\partial{}^{I_b}_{V}\tilde{\boldsymbol{\theta}}} = -\hat{\lambda}\mathbf{J_l}(\hat{\lambda}{}^{b}_{a}\hat{\boldsymbol{\theta}})\mathbf{J_l}({}^{b}_{a}\hat{\boldsymbol{\theta}})^{-1} \qquad \frac{\partial{}^{V}\tilde{\mathbf{p}}_{I_k}}{\partial{}^{V}\tilde{\mathbf{p}}_{I_b}} = \hat{\lambda}\mathbf{I}_3 \tag{76} $$

$$ \frac{\partial{}^{I_k}_{V}\tilde{\boldsymbol{\theta}}}{\partial\lambda} = \mathbf{J_l}(\hat{\lambda}{}^{b}_{a}\hat{\boldsymbol{\theta}}){}^{b}_{a}\hat{\boldsymbol{\theta}} \qquad \frac{\partial{}^{V}\tilde{\mathbf{p}}_{I_k}}{\partial\lambda} = {}^{V}\hat{\mathbf{p}}_{I_b} - {}^{V}\hat{\mathbf{p}}_{I_a} \tag{77} $$

$$ \frac{\partial\lambda}{\partial{}^{I}\tilde{t}_G} = \frac{1}{t_b - t_a} \qquad {}^{b}_{a}\hat{\boldsymbol{\theta}} = \mathrm{Log}({}^{I_b}_{V}\hat{\mathbf{R}}{}^{I_a}_{V}\hat{\mathbf{R}}^\top) \tag{78} $$

where $(\cdot)_z$ is the third column of the matrix. This is because while here ${}^{E}_{V}\mathbf{R}$ is written as a full rotation matrix, we represent it as one that only rotates about the global gravity aligned z-axis.

3 Observability: VIO Reference Frame

3.1 Observability matrix

As system observability plays an important role for state estimation [7, 8], in this section we perform an observability analysis for the proposed GPS-aided VIO system to gain insights about state/parameter identifiability. For concise presentation, we consider a simplified case where the state does not contain biases or stochastic clones and assumes a single feature with perfectly synchronized and calibrated sensors, while the results can be extended to general cases:

$$ {}^V\mathbf{x}_k = \begin{bmatrix} {}^{I_k}_V\bar{q}^\top & {}^V\mathbf{p}_{I_k}^\top & {}^V\mathbf{v}_{I_k}^\top & {}^V\mathbf{p}_f^\top & {}^E_V\bar{q}^\top & {}^E\mathbf{p}_V^\top \end{bmatrix}^\top \tag{79} $$

Also the measurement functions of (55) and (64) in this case become:

$$ \mathbf{z}_{c_k} = \mathbf{\Pi}({}^{C_k}\mathbf{p}_f) + \mathbf{n}_k \tag{80} $$

$$ {}^{C_k}\mathbf{p}_f = {}^{I_k}_V\mathbf{R}\left({}^V\mathbf{p}_f - {}^V\mathbf{p}_{I_i}\right) \tag{81} $$

$$ \mathbf{z}_{g_k} = {}^E\mathbf{p}_V + {}^E_V\mathbf{R}{}^V\mathbf{p}_{I_k} + \mathbf{n}_{g_k} \tag{82} $$

The linearized error state evolution and residuals of both the GPS and visual measurement are generically given by:

$$ {}^V\tilde{\mathbf{x}}_k = {}^V\mathbf{\Phi}(t_k, t_0){}^V\tilde{\mathbf{x}}_0 + \mathbf{w}_k \tag{83} $$

$$ \tilde{\mathbf{z}}_k = {}^V\mathbf{H}_k{}^V\tilde{\mathbf{x}}_k + \mathbf{n}_k \tag{84} $$

We first compute the state transition matrix which can be easily taken from Eq. (31)-(53):

$$ {}^V\mathbf{\Phi}(t_k, t_0) = \begin{bmatrix} {}^{I_k}_V\hat{\mathbf{R}}{}^{I_0}_V\hat{\mathbf{R}}^\top & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_{3\times 7} \\ -\lfloor {}^V\hat{\mathbf{p}}_{I_k} - {}^V\hat{\mathbf{p}}_{I_0} - {}^V\hat{\mathbf{v}}_{I_0}\Delta t + \frac{1}{2}\mathbf{g}\Delta t^2 \times\rfloor {}^{I_0}_V\hat{\mathbf{R}}^\top & \mathbf{I}_3 & \Delta t\mathbf{I}_3 & \mathbf{0}_{3\times 7} \\ -\lfloor {}^V\hat{\mathbf{v}}_{I_k} - {}^V\hat{\mathbf{v}}_{I_0} + \mathbf{g}\Delta t\times\rfloor {}^{I_0}_V\hat{\mathbf{R}}^\top & \mathbf{0}_3 & \mathbf{I}_3 & \mathbf{0}_{3\times 7} \\ \mathbf{0}_{7\times 3} & \mathbf{0}_{7\times 3} & \mathbf{0}_{7\times 3} & \mathbf{I}_7 \end{bmatrix} \tag{85} $$

where $\Delta t = t_k - t_0$.

Linearization of Eq. (80) respect to the state (79) yields the following measurement Jacobians:

$$ \frac{\partial \tilde{\mathbf{z}}_{c_k}}{\partial {}^V\tilde{\mathbf{x}}_k} = \begin{bmatrix} \frac{\partial \tilde{\mathbf{z}}_{c_k}}{\partial {}^{I_k}_V\tilde{\boldsymbol{\theta}}} & \frac{\partial \tilde{\mathbf{z}}_{c_k}}{\partial {}^V\tilde{\mathbf{p}}_{I_k}} & \frac{\partial \tilde{\mathbf{z}}_{c_k}}{\partial {}^V\tilde{\mathbf{v}}_{I_k}} & \frac{\partial \tilde{\mathbf{z}}_{c_k}}{\partial {}^V\tilde{\mathbf{p}}_f} & \frac{\partial \tilde{\mathbf{z}}_{c_k}}{\partial {}^E_V\tilde{\boldsymbol{\theta}}} & \frac{\partial \tilde{\mathbf{z}}_{c_k}}{\partial {}^E\tilde{\mathbf{p}}_V} \end{bmatrix} \tag{86} $$

$$ \frac{\partial \tilde{\mathbf{z}}_{c_k}}{\partial {}^{I_k}_V\tilde{\boldsymbol{\theta}}} = \frac{\partial \tilde{\mathbf{z}}_{c_k}}{\partial {}^{C_k}\tilde{\mathbf{p}}_f}\frac{\partial {}^{C_k}\tilde{\mathbf{p}}_f}{\partial {}^{I_k}_V\tilde{\boldsymbol{\theta}}} \qquad \frac{\partial \tilde{\mathbf{z}}_{c_k}}{\partial {}^V\tilde{\mathbf{p}}_{I_k}} = \frac{\partial \tilde{\mathbf{z}}_{c_k}}{\partial {}^{C_k}\tilde{\mathbf{p}}_f}\frac{\partial {}^{C_k}\tilde{\mathbf{p}}_f}{\partial {}^V\tilde{\mathbf{p}}_{I_k}} \tag{87} $$

$$ \frac{\partial \tilde{\mathbf{z}}_{c_k}}{\partial {}^V\tilde{\mathbf{v}}_{I_k}} = \mathbf{0}_3 \qquad \frac{\partial \tilde{\mathbf{z}}_{c_k}}{\partial {}^V\tilde{\mathbf{p}}_f} = \frac{\partial \tilde{\mathbf{z}}_{c_k}}{\partial {}^{C_k}\tilde{\mathbf{p}}_f}\frac{\partial {}^{C_k}\tilde{\mathbf{p}}_f}{\partial {}^V\tilde{\mathbf{p}}_f} \tag{88} $$

$$ \frac{\partial \tilde{\mathbf{z}}_{c_k}}{\partial {}^E_V\tilde{\boldsymbol{\theta}}} = \mathbf{0}_{3\times 1} \qquad \frac{\partial \tilde{\mathbf{z}}_{c_k}}{\partial {}^E\tilde{\mathbf{p}}_V} = \mathbf{0}_3 \tag{89} $$

$$ \frac{\partial \tilde{\mathbf{z}}_{c_k}}{\partial {}^{C_k}\tilde{\mathbf{p}}_f} = \begin{bmatrix} \frac{1}{{}^{C_k}\hat{z}_f} & 0 & -\frac{{}^{C_k}\hat{x}_f}{{}^{C_k}\hat{z}_f^2} \\ 0 & \frac{1}{{}^{C_k}\hat{z}_f} & -\frac{{}^{C_k}\hat{y}_f}{{}^{C_k}\hat{z}_f^2} \end{bmatrix} =: \mathbf{H}_\Pi \qquad \frac{\partial {}^{C_k}\tilde{\mathbf{p}}_f}{\partial {}^{I_k}_V\tilde{\boldsymbol{\theta}}} = \lfloor {}^{I_k}_V\hat{\mathbf{R}}({}^V\hat{\mathbf{p}}_f - {}^V\hat{\mathbf{p}}_{I_k})\times\rfloor \tag{90} $$

$$ \frac{\partial {}^{C_k}\tilde{\mathbf{p}}_f}{\partial {}^V\tilde{\mathbf{p}}_{I_k}} = -{}^{I_k}_V\hat{\mathbf{R}} \qquad \frac{\partial {}^{C_k}\tilde{\mathbf{p}}_f}{\partial {}^V\tilde{\mathbf{p}}_f} = {}^{I_k}_V\hat{\mathbf{R}} \tag{91} $$

and linearization of Eq. (82) respect to the state (79) yields the following measurement Jacobians:

$$\frac{\partial \tilde{\mathbf{z}}_{g_k}}{\partial {}^V\tilde{\mathbf{x}}_k} = \begin{bmatrix} \frac{\partial \tilde{\mathbf{z}}_{g_k}}{\partial {}^{I_k}_V\tilde{\boldsymbol{\theta}}} & \frac{\partial \tilde{\mathbf{z}}_{g_k}}{\partial {}^V\tilde{\mathbf{p}}_{I_k}} & \frac{\partial \tilde{\mathbf{z}}_{g_k}}{\partial {}^V\tilde{\mathbf{v}}_{I_k}} & \frac{\partial \tilde{\mathbf{z}}_{g_k}}{\partial {}^V\tilde{\mathbf{p}}_f} & \frac{\partial \tilde{\mathbf{z}}_{g_k}}{\partial {}^E_V\tilde{\boldsymbol{\theta}}} & \frac{\partial \tilde{\mathbf{z}}_{g_k}}{\partial {}^E\tilde{\mathbf{p}}_V} \end{bmatrix} \tag{92}$$

$$\frac{\partial \tilde{\mathbf{z}}_{g_k}}{\partial {}^{I_k}_V\tilde{\boldsymbol{\theta}}} = \mathbf{0}_3 \qquad \frac{\partial \tilde{\mathbf{z}}_{g_k}}{\partial {}^V\tilde{\mathbf{p}}_{I_k}} = {}^E_V\mathbf{R} \tag{93}$$

$$\frac{\partial \tilde{\mathbf{z}}_{g_k}}{\partial {}^V\tilde{\mathbf{v}}_{I_k}} = \mathbf{0}_3 \qquad \frac{\partial \tilde{\mathbf{z}}_{g_k}}{\partial {}^V\tilde{\mathbf{p}}_f} = \mathbf{0}_3 \tag{94}$$

$$\frac{\partial \tilde{\mathbf{z}}_{g_k}}{\partial {}^E_V\tilde{\boldsymbol{\theta}}} = (\lfloor {}^E_V\hat{\mathbf{R}}{}^V\hat{\mathbf{p}}_{I_k}\times\rfloor)_z \qquad \frac{\partial \tilde{\mathbf{z}}_{g_k}}{\partial {}^E\tilde{\mathbf{p}}_V} = \mathbf{I}_3 \tag{95}$$

Stacking Eq. (86) and Eq. (92) we get the general Jacobian of the state as:

$${}^V\mathbf{H}_k = \begin{bmatrix} \mathbf{H}_\Pi\lfloor {}^{I_k}_V\hat{\mathbf{R}}({}^V\hat{\mathbf{p}}_f - {}^V\hat{\mathbf{p}}_{I_k})\times\rfloor & -\mathbf{H}_\Pi {}^{I_k}_V\hat{\mathbf{R}} & \mathbf{0}_{2\times 3} & \mathbf{H}_\Pi {}^{I_k}_V\hat{\mathbf{R}} & \mathbf{0}_{2\times 1} & \mathbf{0}_{2\times 3} \\ \mathbf{0}_3 & {}^E_V\hat{\mathbf{R}} & \mathbf{0}_3 & \mathbf{0}_3 & (\lfloor {}^E_V\hat{\mathbf{R}}{}^V\hat{\mathbf{p}}_{I_k}\times\rfloor)_z & \mathbf{I}_3 \end{bmatrix} \tag{96}$$

Now we can construct the observability matrix ${}^V\mathbf{M}$ (see [9]).

$${}^V\mathbf{M} = \begin{bmatrix} \vdots \\ {}^V\mathbf{H}_k {}^V\boldsymbol{\Phi}(t_k, t_0) \\ \vdots \end{bmatrix}, \quad \mathbf{H}_k {}^V\boldsymbol{\Phi}(t_k, t_0) = \begin{bmatrix} \boldsymbol{\Gamma}_1 & \boldsymbol{\Gamma}_2 & \boldsymbol{\Gamma}_3 & \boldsymbol{\Gamma}_4 & \mathbf{0}_{2\times 1} & \mathbf{0}_{2\times 3} \\ \boldsymbol{\Gamma}_5 & \boldsymbol{\Gamma}_6 & \boldsymbol{\Gamma}_7 & \mathbf{0}_3 & \boldsymbol{\Gamma}_8 & \boldsymbol{\Gamma}_9 \end{bmatrix} \tag{97}$$

$$\boldsymbol{\Gamma}_1 = \mathbf{H}_\Pi {}^{I_k}_V\hat{\mathbf{R}}\lfloor {}^V\hat{\mathbf{p}}_f - {}^V\hat{\mathbf{p}}_{I_0} - {}^V\hat{\mathbf{v}}_{I_0}\Delta t + \frac{1}{2}\mathbf{g}\Delta t^2\times\rfloor {}^{I_0}_V\hat{\mathbf{R}}^\top \tag{98}$$

$$\boldsymbol{\Gamma}_2 = -\mathbf{H}_\Pi {}^{I_k}_V\hat{\mathbf{R}}, \qquad \boldsymbol{\Gamma}_3 = -\Delta t\mathbf{H}_\Pi {}^{I_k}_V\hat{\mathbf{R}} \tag{99}$$

$$\boldsymbol{\Gamma}_4 = \mathbf{H}_\Pi {}^{I_k}_V\hat{\mathbf{R}}, \qquad \boldsymbol{\Gamma}_5 = -{}^E_V\hat{\mathbf{R}}\lfloor {}^V\hat{\mathbf{p}}_{I_k} - {}^V\hat{\mathbf{p}}_{I_0} - {}^V\hat{\mathbf{v}}_{I_0}\Delta t + \frac{1}{2}\mathbf{g}\Delta t^2\times\rfloor {}^{I_0}_V\hat{\mathbf{R}}^\top \tag{100}$$

$$\boldsymbol{\Gamma}_6 = {}^E_V\hat{\mathbf{R}}, \qquad \boldsymbol{\Gamma}_7 = \Delta t {}^E_V\hat{\mathbf{R}} \tag{101}$$

$$\boldsymbol{\Gamma}_8 = (\lfloor {}^E_V\hat{\mathbf{R}}{}^V\hat{\mathbf{p}}_{I_k}\times\rfloor)_z, \qquad \boldsymbol{\Gamma}_9 = \mathbf{I}_3 \tag{102}$$

3.2 Rotation of $\{V\}$ respect to $\{E\}$ along the gravity axis

For the observability analysis, we change of the initial conditions (at t_0) by a perturbation. The first perturbation is a small rotation of $\{V\}$ respect to $\{E\}$ along the gravity axis. This can be mathematically represented as:

$$\begin{aligned} {}^E_{V'}\mathbf{R} &= {}^E_V\mathbf{R}{}^{V'}_V\mathbf{R}^\top \qquad (103) \\ &= {}^E_V\mathbf{R}(\mathbf{I} + \lfloor \alpha\mathbf{g}\times\rfloor) \qquad (104) \end{aligned}$$

where α is small scalar. Now we analyze how this change is equivalent to the state's perturbation. For the orientation:

$$\begin{aligned} (\mathbf{I} - \lfloor {}^{I_0}_V\tilde{\boldsymbol{\theta}}\times\rfloor){}^{I_0}_V\mathbf{R} &= {}^{I_0}_{V'}\mathbf{R} \qquad (105) \\ &= {}^{I_0}_V\mathbf{R}(\mathbf{I} + \lfloor \alpha\mathbf{g}\times\rfloor) \qquad (106) \\ \Rightarrow {}^{I_0}_V\tilde{\boldsymbol{\theta}} &= -\alpha {}^{I_0}_V\mathbf{R}\mathbf{g} \qquad (107) \end{aligned}$$

Apply the perturbation Eq. (104) to initial position of IMU in VIO

$$
\begin{aligned}
{}^{V'}\mathbf{p}_{I_0} &= {}^{V'}\mathbf{p}_V + {}^{V'}_{V}\mathbf{R}\,{}^{V}\mathbf{p}_{I_0} && (108)\\
&= \mathbf{0}_{3\times 1} + (\mathbf{I} - \lfloor \alpha\mathbf{g}\times\rfloor)\,{}^{V}\mathbf{p}_{I_0} && (109)\\
&= (\mathbf{I} - \lfloor \alpha\mathbf{g}\times\rfloor)\,{}^{V}\mathbf{p}_{I_0} && (110)
\end{aligned}
$$

Then the perturbation of the position of IMU in (79) become:

$$
\begin{aligned}
{}^{V}\mathbf{p}_{I_0} + {}^{V}\tilde{\mathbf{p}}_{I_0} &= {}^{V'}\mathbf{p}_{I_0} && (111)\\
&= (\mathbf{I} - \lfloor \alpha\mathbf{g}\times\rfloor)\,{}^{V}\mathbf{p}_{I_0} && (112)\\
\Rightarrow {}^{V}\tilde{\mathbf{p}}_{I_0} &= -\lfloor \alpha\mathbf{g}\times\rfloor\,{}^{V}\mathbf{p}_{I_0} && (113)\\
&= \alpha\lfloor {}^{V}\mathbf{p}_{I_0}\times\rfloor\mathbf{g} && (114)
\end{aligned}
$$

Apply the perturbation Eq. (104) to initial velocity of IMU in VIO

$$
\begin{aligned}
{}^{V'}\mathbf{v}_{I_0} &= {}^{V'}\mathbf{v}_V + {}^{V'}_{V}\mathbf{R}\,{}^{V}\mathbf{v}_{I_0} && (115)\\
&= \mathbf{0}_{3\times 1} + (\mathbf{I} - \lfloor \alpha\mathbf{g}\times\rfloor)\,{}^{V}\mathbf{v}_{I_0} && (116)\\
&= (\mathbf{I} - \lfloor \alpha\mathbf{g}\times\rfloor)\,{}^{V}\mathbf{v}_{I_0} && (117)
\end{aligned}
$$

Then the perturbation of the velocity of IMU in (79) become:

$$
\begin{aligned}
{}^{V}\mathbf{v}_{I_0} + {}^{V}\tilde{\mathbf{v}}_{I_0} &= {}^{V'}\mathbf{v}_{I_0} && (118)\\
&= (\mathbf{I} - \lfloor \alpha\mathbf{g}\times\rfloor)\,{}^{V}\mathbf{v}_{I_0} && (119)\\
\Rightarrow {}^{V}\tilde{\mathbf{v}}_{I_0} &= -\lfloor \alpha\mathbf{g}\times\rfloor\,{}^{V}\mathbf{v}_{I_0} && (120)\\
&= \alpha\lfloor {}^{V}\mathbf{v}_{I_0}\times\rfloor\mathbf{g} && (121)
\end{aligned}
$$

Apply the perturbation Eq. (104) to initial position of feature in VIO

$$
\begin{aligned}
{}^{V'}\mathbf{p}_{f} &= {}^{V'}\mathbf{p}_V + {}^{V'}_{V}\mathbf{R}\,{}^{V}\mathbf{p}_{f} && (122)\\
&= \mathbf{0}_{3\times 1} + (\mathbf{I} - \lfloor \alpha\mathbf{g}\times\rfloor)\,{}^{V}\mathbf{p}_{f} && (123)\\
&= (\mathbf{I} - \lfloor \alpha\mathbf{g}\times\rfloor)\,{}^{V}\mathbf{p}_{f} && (124)
\end{aligned}
$$

Then the perturbation of the position of feature in (79) become:

$$
\begin{aligned}
{}^{V}\mathbf{p}_{f} + {}^{V}\tilde{\mathbf{p}}_{f} &= {}^{V'}\mathbf{p}_{f} && (125)\\
&= (\mathbf{I} - \lfloor \alpha\mathbf{g}\times\rfloor)\,{}^{V}\mathbf{p}_{f} && (126)\\
\Rightarrow {}^{V}\tilde{\mathbf{p}}_{f} &= -\lfloor \alpha\mathbf{g}\times\rfloor\,{}^{V}\mathbf{p}_{f} && (127)\\
&= \alpha\lfloor {}^{V}\mathbf{p}_{f}\times\rfloor\mathbf{g} && (128)
\end{aligned}
$$

Lastly, the perturbation of the rotation from $\{V\}$ to $\{E\}$ become:

$$
\begin{aligned}
(\mathbf{I} - \lfloor {}^{E}_{V}\tilde{\boldsymbol{\theta}}\times\rfloor)\,{}^{E}_{V}\mathbf{R} &= {}^{E}_{V'}\mathbf{R} && (129)\\
&= {}^{E}_{V}\mathbf{R}(\mathbf{I} + \lfloor \alpha\mathbf{g}\times\rfloor) && (130)\\
\Rightarrow {}^{E}_{V}\tilde{\boldsymbol{\theta}} &= -\alpha\,{}^{E}_{V}\mathbf{R}\mathbf{g} && (131)\\
\Rightarrow ({}^{E}_{V}\tilde{\boldsymbol{\theta}})_3 &= (-\alpha\,{}^{E}_{V}\mathbf{R}\mathbf{g})_3 && (132)
\end{aligned}
$$

where $(\cdot)_3$ is the third element of the vector. Therefore, the perturbation of initial condition can be represented as state perturbation as:

$$
{}^{V}\tilde{\mathbf{x}}_k = \begin{bmatrix} -\alpha^{I_0}_{V}\mathbf{R}\mathbf{g} \\ \alpha\lfloor {}^{V}\mathbf{p}_{I_0}\times\rfloor\mathbf{g} \\ \alpha\lfloor {}^{V}\mathbf{v}_{I_0}\times\rfloor\mathbf{g} \\ \alpha\lfloor {}^{V}\mathbf{p}_{f}\times\rfloor\mathbf{g} \\ (-\alpha^{E}_{V}\mathbf{R}\mathbf{g})_3 \\ \mathbf{0}_{3\times 1} \end{bmatrix} \tag{133}
$$

Clearly, ${}^{V}\mathbf{H}_k{}^{V}\mathbf{\Phi}(t_k,t_0){}^{V}\tilde{\mathbf{x}}_k$ yields zero matrix. Therefore, the ${}^{V}\tilde{\mathbf{x}}_k$ is in null space of ${}^{V}\mathbf{H}_k{}^{V}\mathbf{\Phi}(t_k,t_0)$, which means the rotation of $\{V\}$ respect to $\{E\}$ along the gravity axis is unobservable.

3.3 Translation of $\{V\}$ respect to $\{E\}$

Now we try perturbing the ${}^{E}\mathbf{p}_V$ by $\mathbf{T} = \alpha\mathbf{e}_1 + \beta\mathbf{e}_2 + \gamma\mathbf{e}_3$ where $\mathbf{e}_i$ for $i = 1, 2, 3$ are standard basis:

$$
{}^{E}\mathbf{p}_{V'} = {}^{E}\mathbf{p}_V + \mathbf{T} \tag{134}
$$

Apply the perturbation Eq. (134) to the initial position of the IMU:

$$
\begin{aligned}
{}^{V'}\mathbf{p}_{I_0} &= {}^{V'}\mathbf{p}_V + {}^{V'}_{V}\mathbf{R}{}^{V}\mathbf{p}_{I_0} && (135)\\
&= {}^{V'}\mathbf{p}_E + {}^{V'}_{E}\mathbf{R}{}^{E}\mathbf{p}_V + {}^{V}\mathbf{p}_{I_0} && (136)\\
&= -{}^{E}_{V}\mathbf{R}^{\top}{}^{E}\mathbf{p}_{V'} + {}^{E}_{V}\mathbf{R}^{\top}{}^{E}\mathbf{p}_V + {}^{V}\mathbf{p}_{I_0} && (137)\\
&= -{}^{E}_{V}\mathbf{R}^{\top}\mathbf{T} + {}^{V}\mathbf{p}_{I_0} && (138)
\end{aligned}
$$

Then the perturbation of the position of IMU in (79) become:

$$
\begin{aligned}
{}^{V}\mathbf{p}_{I_0} + {}^{V}\tilde{\mathbf{p}}_{I_0} &= {}^{V'}\mathbf{p}_{I_0} && (139)\\
&= -{}^{E}_{V}\mathbf{R}^{\top}\mathbf{T} + {}^{V}\mathbf{p}_{I_0} && (140)\\
\Rightarrow {}^{V}\tilde{\mathbf{p}}_{I_0} &= -{}^{E}_{V}\mathbf{R}^{\top}\mathbf{T} && (141)
\end{aligned}
$$

Apply the perturbation Eq. (134) to the position of the feature:

$$
\begin{aligned}
{}^{V'}\mathbf{p}_f &= {}^{V'}\mathbf{p}_V + {}^{V'}_{V}\mathbf{R}{}^{V}\mathbf{p}_f && (142)\\
&= {}^{V'}\mathbf{p}_E + {}^{V'}_{E}\mathbf{R}{}^{E}\mathbf{p}_V + {}^{V}\mathbf{p}_f && (143)\\
&= -{}^{E}_{V}\mathbf{R}^{\top}{}^{E}\mathbf{p}_{V'} + {}^{E}_{V}\mathbf{R}^{\top}{}^{E}\mathbf{p}_V + {}^{V}\mathbf{p}_f && (144)\\
&= -{}^{E}_{V}\mathbf{R}^{\top}\mathbf{T} + {}^{V}\mathbf{p}_f && (145)
\end{aligned}
$$

Then the perturbation of the position of IMU in (79) become:

$$
\begin{aligned}
{}^{V}\mathbf{p}_f + {}^{V}\tilde{\mathbf{p}}_f &= {}^{V'}\mathbf{p}_f && (146)\\
&= -{}^{E}_{V}\mathbf{R}^{\top}\mathbf{T} + {}^{V}\mathbf{p}_f && (147)\\
\Rightarrow {}^{V}\tilde{\mathbf{p}}_f &= -{}^{E}_{V}\mathbf{R}^{\top}\mathbf{T} && (148)
\end{aligned}
$$

Therefore, the perturbation of $\{V\}$ respect to $\{E\}$ can be represented as state perturbation as:

$$
{}^{V}\tilde{\mathbf{x}}_k = \begin{bmatrix} \mathbf{0}_{3\times1} \\ -{}^{E}_{V}\mathbf{R}^{\top}\mathbf{T} \\ \mathbf{0}_{3\times1} \\ -{}^{E}_{V}\mathbf{R}^{\top}\mathbf{T} \\ \mathbf{0}_{3\times1} \\ \mathbf{T} \end{bmatrix} \tag{149}
$$

Clearly, ${}^{V}\mathbf{H}_k{}^{V}\mathbf{\Phi}(t_k, t_0){}^{V}\tilde{\mathbf{x}}_k$ yields zero matrix. Therefore, the ${}^{V}\tilde{\mathbf{x}}_k$ is in null space of ${}^{V}\mathbf{H}_k{}^{V}\mathbf{\Phi}(t_k, t_0)$, which means the translation of $\{V\}$ respect to $\{E\}$ is unobservable.

3.4 Null space of observability matrix of system in VIO

$$
\mathbf{null}({}^{V}\mathbf{M}) = \begin{bmatrix} \mathbf{0}_3 & -{}^{I_0}_{V}\hat{\mathbf{R}}\mathbf{g} \\ \mathbf{I}_3 & \lfloor {}^{V}\hat{\mathbf{p}}_{I_0} \times \rfloor \mathbf{g} \\ \mathbf{0}_3 & \lfloor {}^{V}\hat{\mathbf{v}}_{I_0} \times \rfloor \mathbf{g} \\ \mathbf{I}_3 & \lfloor {}^{V}\hat{\mathbf{p}}_{f} \times \rfloor \mathbf{g} \\ \mathbf{0}_{1\times3} & (-{}^{E}_{V}\hat{\mathbf{R}}\mathbf{g})_3 \\ -{}^{E}_{V}\hat{\mathbf{R}} & \mathbf{0}_{3\times1} \end{bmatrix}_{16\times4} \tag{150}
$$

The simplified, gathered nullspace from Eq. (133) and Eq. (149) is shown above. The span of the columns of this matrix encodes the unobservable subspace. By inspection, the first block column corresponds to the translation of $\{V\}$ relative to $\{E\}$ and the second block column to the rotation of $\{V\}$ with respect to $\{E\}$ along the axis of gravity. It thus becomes clear that the GPS-VIO system in the VIO frame has these four unobservable directions which are essentially inherited from the standard VIO [8, 10].

4 Observability: ENU Reference Frame

When the system is in $\{E\}$, the parameters in the state is represented in $\{E\}$ as:

$$
{}^{E}\mathbf{x}_k = \begin{bmatrix} {}^{I_k}_{E}\bar{q}^\top & {}^{E}\mathbf{p}_{I_k}^\top & {}^{E}\mathbf{v}_{I_k}^\top & {}^{E}\mathbf{p}_{f}^\top & {}^{E}_{V}\bar{q}^\top & {}^{E}\mathbf{p}_{V}^\top \end{bmatrix}^\top \tag{151}
$$

We can easily compute the observability matrix ${}^{E}\mathbf{M}$ from Eq. (97)-(102) by changing the parameters in $\{V\}$ to $\{E\}$ as:

$$
{}^{E}\mathbf{M} = \begin{bmatrix} \vdots \\ {}^{E}\mathbf{H}_k {}^{E}\mathbf{\Phi}(t_k, t_0) \\ \vdots \end{bmatrix}, \quad {}^{E}\mathbf{H}_k {}^{E}\mathbf{\Phi}(t_k, t_0) = \begin{bmatrix} \mathbf{\Gamma}_1 & \mathbf{\Gamma}_2 & \mathbf{\Gamma}_3 & \mathbf{\Gamma}_4 & \mathbf{0}_{2\times 1} & \mathbf{0}_{2\times 3} \\ \mathbf{\Gamma}_5 & \mathbf{\Gamma}_6 & \mathbf{\Gamma}_7 & \mathbf{0}_3 & \mathbf{\Gamma}_8 & \mathbf{\Gamma}_9 \end{bmatrix} \tag{152}
$$

$$
\mathbf{\Gamma}_1 = \mathbf{H}_{\Pi} {}^{I_k}_{E}\hat{\mathbf{R}} \lfloor {}^{E}\hat{\mathbf{p}}_f - {}^{E}\hat{\mathbf{p}}_{I_0} - {}^{E}\hat{\mathbf{v}}_{I_0}\Delta t + \frac{1}{2}\mathbf{g}\Delta t^2 \times \rfloor {}^{I_0}_{E}\hat{\mathbf{R}}^\top \tag{153}
$$

$$
\mathbf{\Gamma}_2 = -\mathbf{H}_{\Pi} {}^{I_k}_{E}\hat{\mathbf{R}}, \qquad \mathbf{\Gamma}_3 = -\Delta t \mathbf{H}_{\Pi} {}^{I_k}_{E}\hat{\mathbf{R}} \tag{154}
$$

$$
\mathbf{\Gamma}_4 = \mathbf{H}_{\Pi} {}^{I_k}_{E}\hat{\mathbf{R}}, \qquad \mathbf{\Gamma}_5 = -\lfloor {}^{E}\hat{\mathbf{p}}_{I_k} - {}^{E}\hat{\mathbf{p}}_{I_0} - {}^{E}\hat{\mathbf{v}}_{I_0}\Delta t + \frac{1}{2}\mathbf{g}\Delta t^2 \times \rfloor {}^{I_0}_{E}\hat{\mathbf{R}}^\top \tag{155}
$$

$$
\mathbf{\Gamma}_6 = \mathbf{I}_3, \qquad \mathbf{\Gamma}_7 = \Delta t \mathbf{I}_3 \tag{156}
$$

$$
\mathbf{\Gamma}_8 = (\lfloor {}^{E}\hat{\mathbf{p}}_{I_k} \times \rfloor)_z, \qquad \mathbf{\Gamma}_9 = \mathbf{I}_3 \tag{157}
$$

Clearly, the multiplication of ${}^{E}\mathbf{H}_k {}^{E}\mathbf{\Phi}(t_k, t_0)$ with $\mathbf{null}({}^{V}\mathbf{M})$ does not yield a zero matrix which means the 4 unobservable directions of Eq. (79) are now observable. Since VIO is known to have 4 unobservable directions [8, 10], we can conclude that the state in the ENU, see Eq. (151), is fully observable.

While the above results seem to be counter-intuitive given the availability of global GPS measurements, the root cause of this unobservability is the gauge freedom of the 4 d.o.f GPS-VIO frame transformation. Thus even though we utilize global measurements, the system maintains a non-trivial null space. Unobservable directions are known to cause inconsistency issues for linearized estimators as these null spaces falsely disappear due to numerical errors. Therefore the estimator gains information in spurious directions, hurting overall consistency and accuracy, unless special techniques are utilized [8, 11].

5 State transformation from $\{V\}$ to $\{E\}$

Based on the analysis of the previous chapter, we transform the state transform Eq. (79) to Eq. (151) including the propagation of the covariance to achieve full observability. First, we transform the state as:

$$^E\mathbf{x}_k = g(^V\mathbf{x}_k, {}^E_V\mathbf{R}, {}^E\mathbf{p}_V) \tag{158}$$

which in specific:

$$^{I_k}_E\mathbf{R} = {}^{I_k}_V\mathbf{R}{}^E_V\mathbf{R}^\top \tag{159}$$

$$^E\mathbf{p}_{I_k} = {}^E_V\mathbf{R}{}^V\mathbf{p}_{I_k} + {}^E\mathbf{p}_V \tag{160}$$

$$^E\mathbf{v}_{I_k} = {}^E_V\mathbf{R}{}^V\mathbf{v}_{I_k} \tag{161}$$

The Eq. (159)-(161) can be used to warp the all stochastic clone poses and features in general. Biases and calibration parameters have identity transform between the two frames.

Now we linearize function (158) at current estimate to achieve the Jacobian matrix $\mathbf{\Psi}$ and propagate the covariance matrix with it as:

$$^E\tilde{\mathbf{x}}_k = \mathbf{\Psi}{}^V\tilde{\mathbf{x}}_k \ , \quad \mathbf{P}^\oplus = \mathbf{\Psi}\mathbf{P}^\ominus\mathbf{\Psi}^\top \tag{162}$$

where

$$\mathbf{\Psi} = \begin{bmatrix} \mathbf{I}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & ({}^{I_k}_V\hat{\mathbf{R}}{}^E_V\hat{\mathbf{R}}^\top)_z & \mathbf{0}_3 \\ \mathbf{0}_3 & {}^E_V\hat{\mathbf{R}} & \mathbf{0}_3 & \mathbf{0}_3 & (\lfloor {}^E_V\hat{\mathbf{R}}{}^V\hat{\mathbf{p}}_{I_k}\times\rfloor)_z & \mathbf{I}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 & {}^E_V\hat{\mathbf{R}} & \mathbf{0}_3 & (\lfloor {}^E_V\hat{\mathbf{R}}{}^V\hat{\mathbf{v}}_{I_k}\times\rfloor)_z & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & {}^E_V\hat{\mathbf{R}} & (\lfloor {}^E_V\hat{\mathbf{R}}{}^V\hat{\mathbf{p}}_f\times\rfloor)_z & \mathbf{I}_3 \\ \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & 1 & \mathbf{0}_{1\times3} \\ \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_{3\times1} & \mathbf{I}_3 \end{bmatrix} \tag{163}$$

After the transformation, we perform Schur complement [12] to marginalize ${}^E_V\bar{q}$ and ${}^E\mathbf{p}_V$ from the state because they are no longer needed. The final state is:

$$\mathbf{x}_k = \begin{bmatrix} {}^{I_k}_E\bar{q} & {}^E\mathbf{p}_{I_k} & {}^E\mathbf{v}_{I_k} & {}^E\mathbf{p}_f \end{bmatrix} \tag{164}$$

or in general:

$$^E\mathbf{x}_k = \begin{bmatrix} {}^E\mathbf{x}_{I_k}^\top & {}^E\mathbf{x}_{C_k}^\top & {}^I\mathbf{p}_G^\top & {}^It_G \end{bmatrix}^\top \tag{165}$$

$$^E\mathbf{x}_{I_k} = \begin{bmatrix} {}^{I_k}_E\bar{q}^\top & {}^E\mathbf{p}_{I_k}^\top & {}^E\mathbf{v}_{I_k}^\top & \mathbf{b}_{\omega_k}^\top & \mathbf{b}_{a_k}^\top \end{bmatrix}^\top \tag{166}$$

$$^E\mathbf{x}_{C_k} = \begin{bmatrix} {}^{I_{k-1}}_E\bar{q}^\top & {}^E\mathbf{p}_{I_{k-1}}^\top & \cdots & {}^{I_{k-n}}_E\bar{q}^\top & {}^E\mathbf{p}_{I_{k-n}}^\top \end{bmatrix}^\top \tag{167}$$

Note the state is fully observable.

References

[1] Anastasios I Mourikis and Stergios I Roumeliotis. "A multi-state constraint Kalman filter for vision-aided inertial navigation". In: *Proceedings 2007 IEEE International Conference on Robotics and Automation*. IEEE. 2007, pp. 3565–3572.

[2] Nikolas Trawny and Stergios I. Roumeliotis. *Indirect Kalman Filter for 3D Attitude Estimation*. Tech. rep. University of Minnesota, Dept. of Comp. Sci. & Eng., Mar. 2005.

[3] Christoph Hertzberg et al. "Integrating generic sensor fusion algorithms with sound state representations through encapsulation of manifolds". In: *Information Fusion* 14.1 (2013), pp. 57–77.

[4] Peter S Maybeck. *Stochastic models, estimation, and control*. Vol. 3. Academic press, 1982.

[5] Gregory Chirikjian. *Stochastic Models, Information Theory, and Lie Groups, Volume 2: Analytic Methods and Modern Applications*. Vol. 2. Springer Science & Business Media, 2011.

[6] Mingyang Li. "Visual-Inertial Odometry on Resource-Constrained Systems". PhD thesis. UC Riverside, 2014.

[7] Guoquan Huang. "Improving the Consistency of Nonlinear Estimators: Analysis, Algorithms, and Applications". PhD thesis. Department of Computer Science and Engineering, University of Minnesota, 2012. URL: https://conservancy.umn.edu/handle/11299/146717.

[8] Joel A Hesch et al. "Observability-constrained vision-aided inertial navigation". In: (2012).

[9] Z. Chen, K. Jiang, and J.C. Hung. "Local observability matrix and its application to observability analyses". In: *Proc. of the 16th Annual Conference of IEEE*. Pacific Grove, CA, 1990, pp. 100–103.

[10] Jonathan Kelly and Gaurav S Sukhatme. "Visual-inertial sensor fusion: Localization, mapping and sensor-to-sensor self-calibration". In: *The International Journal of Robotics Research* 30.1 (2011), pp. 56–79.

[11] Guoquan Huang, Anastasios I. Mourikis, and Stergios I. Roumeliotis. "A First-Estimates Jacobian EKF for Improving SLAM Consistency". In: *Proc. of the 11th International Symposium on Experimental Robotics*. Athens, Greece, 2008.

[12] Fuzhen Zhang. *The Schur complement and its applications*. Vol. 4. Springer Science & Business Media, 2006.