Unobservable Directions of VINS Under Special Motions

Kejian J. Wu and Stergios I. Roumeliotis
{kejian|stergios}@cs.umn.edu

Department of Computer Science & Engineering
University of Minnesota

MARS LAB

Multiple Autonomous
Robotic Systems Labroratory

Technical Report
Number -2016-002
September 2016

Dept. of Computer Science & Engineering
University of Minnesota
4-192 EE/CS Building
200 Union St. S.E.
Minneapolis, MN 55455
Tel: (612) 625-2217
Fax: (612) 625-0572
URL: mars.cs.umn.edu

Unobservable Directions of VINS Under Special Motions

Kejian J. Wu and Stergios I. Roumeliotis
{kejian|stergios}@cs.umn.edu

Department of Computer Science & Engineering
University of Minnesota

Multiple Autonomous Robotic Systems Laboratory, TR-2016-002
September 2016

Abstract

In this paper, we study the unobservable directions of vision-aided inertial navigation systems (VINS) under special motion profiles. It is well-known that the VINS has four unobservable directions, corresponding to the global translation and the rotation around gravity, under the assumption of *generic* 3D motions. Here we study the case when the motion is constrained to be of either constant local acceleration or no rotation, as commonly found with ground vehicles. We show and prove the additional unobservable directions due to these special motions *analytically*, as well as the corresponding physical meanings of these new directions.

1 Vision-aided Inertial Navigation System (VINS)

First, we provide a brief review of the VINS. The VINS estimates the following state vector:

$$\mathbf{x} = \begin{bmatrix} {}^I\mathbf{q}_G^T & \mathbf{b}_g^T & {}^G\mathbf{v}_I^T & \mathbf{b}_a^T & {}^G\mathbf{p}_I^T \mid {}^G\mathbf{f}_1^T & \cdots & {}^G\mathbf{f}_N^T \end{bmatrix}^T \tag{1}$$

where ${}^I\mathbf{q}_G$ is the unit quaternion that represents the orientation of the global frame $\{G\}$ in the IMU frame $\{I\}$ at time t. ${}^G\mathbf{v}_I$ and ${}^G\mathbf{p}_I$ are the the velocity and position of $\{I\}$ in $\{G\}$, respectively, and the gyroscope and accelerometer biases are denoted by $\mathbf{b}_g$ and $\mathbf{b}_a$, respectively. Finally, the positions of point features in $\{G\}$ are denoted by ${}^G\mathbf{f}_j,\ j = 1, ..., N$.

The IMU provides measurements of the rotational velocity, $\boldsymbol{\omega}_m$, and the linear acceleration, $\mathbf{a}_m$, as:

$$\begin{aligned} \boldsymbol{\omega}_m(t) &= {}^I\boldsymbol{\omega}(t) + \mathbf{b}_g(t) + \mathbf{n}_g(t) \\ \mathbf{a}_m(t) &= \mathbf{C}({}^I\mathbf{q}_G(t))({}^G\mathbf{a}(t) - {}^G\mathbf{g}) + \mathbf{b}_a(t) + \mathbf{n}_a(t) \end{aligned} \tag{2}$$

where the noise terms, $\mathbf{n}_g(t)$ and $\mathbf{n}_a(t)$ are modelled as zero-mean, white Gaussian noise processes, $\mathbf{C}({}^I\mathbf{q}_G(t))$ denotes the rotational matrix associated with the quaternion ${}^I\mathbf{q}_G$ at time t, while the gravitational acceleration, ${}^G\mathbf{g}$, is considered a known constant. The IMU's rotational velocity ${}^I\boldsymbol{\omega}(t)$ and linear acceleration ${}^G\mathbf{a}(t)$, in (2), can be used to derive the continuous-time system equations:

$$\begin{aligned} {}^I\dot{\mathbf{q}}_G(t) &= \frac{1}{2}\boldsymbol{\Omega}(\boldsymbol{\omega}_m(t) - \mathbf{b}_g(t) - \mathbf{n}_g(t))\,{}^I\mathbf{q}_G(t) \\ \dot{\mathbf{b}}_g(t) &= \mathbf{n}_{wg}(t) \\ {}^G\dot{\mathbf{v}}_I(t) &= \mathbf{C}({}^I\mathbf{q}_G(t))^T(\mathbf{a}_m(t) - \mathbf{b}_a(t) - \mathbf{n}_a(t)) + {}^G\mathbf{g} \\ \dot{\mathbf{b}}_a(t) &= \mathbf{n}_{wa}(t) \\ {}^G\dot{\mathbf{p}}_I(t) &= {}^G\mathbf{v}_I(t) \\ {}^G\dot{\mathbf{f}}_j(t) &= \mathbf{0}, \quad j = 1, \ldots, N \end{aligned} \tag{3}$$

where, $\mathbf{\Omega}(\boldsymbol{\omega}) \triangleq \begin{bmatrix} -\lfloor\boldsymbol{\omega}\rfloor & \boldsymbol{\omega} \\ -\boldsymbol{\omega}^T & 0 \end{bmatrix}$ for $\boldsymbol{\omega} \in \mathbb{R}^3$, $\lfloor\cdot\rfloor$ denotes the skew-symmetric matrix, while the IMU biases are modelled as random walks driven by white, zero-mean Gaussian noise processes $\mathbf{n}_{wg}(t)$ and $\mathbf{n}_{wa}(t)$, respectively.

As the camera-IMU pair moves, the camera provides measurements of point features extracted from the images. Each such measurement, $\mathbf{z}_j$, is modeled as the perspective projection of the point feature $\mathbf{f}_j$, expressed in the current IMU frame[1] $\{I\}$, onto the image plane:

$$\mathbf{z}_j = \frac{1}{z}\begin{bmatrix} x \\ y \end{bmatrix} + \mathbf{n}_j, \quad \begin{bmatrix} x \\ y \\ z \end{bmatrix} \triangleq {}^I\mathbf{f}_j = \mathbf{C}({}^I\mathbf{q}_G)({}^G\mathbf{f}_j - {}^G\mathbf{p}_I) \tag{4}$$

where the measurement noise, $\mathbf{n}_j$, is modeled as zero mean, white Gaussian. For modeling the IMU propagation [see (3)] and camera observations [see (4)], including their error equations and analytical Jacobians, we follow [1].

2 VINS: Observability Analysis Under Specific Motion Profiles

Observability is a fundamental property of a dynamic system and provides important insights. Previous works have studied the observability properties of VINS, and employed the results of their analysis to improve the consistency of the estimator [1]. Specifically, in [2, 1], it was shown that, *for generic motions*, a VINS has four unobservable directions (three for global translation and one for global yaw).

In this paper, we are interested in the case when the VINS is deployed on a ground vehicle, whose motion is approximately planar, and, for the most part, along a straight line (e.g., when moving forward) or a circular arc (e.g., when turning). In particular, we are interested in the impact that such motions have on the VINS's observability properties, and hence the accuracy of the corresponding estimator.

2.1 Constant Acceleration

Consider that the platform moves with constant *local* linear acceleration (e.g., on a circle), i.e.,

$${}^I\mathbf{a}(t) \triangleq \mathbf{C}({}^I\mathbf{q}_G(t))^G\mathbf{a}(t) \equiv {}^I\mathbf{a}, \quad \forall\, t \geq t_0 \tag{5}$$

where ${}^I\mathbf{a}$ is a constant vector with respect to time, we prove the following theorem:

Theorem 1: The linearized VINS model of (3)-(4) has the following *additional* unobservable direction, besides the global translation and yaw, *if and only if* the condition in (5) is satisfied:

$$\mathbf{N}_s = \begin{bmatrix} \mathbf{0}_{3\times1} \\ \mathbf{0}_{3\times1} \\ {}^G\mathbf{v}_{I_0} \\ -{}^I\mathbf{a} \\ {}^G\mathbf{p}_{I_0} \\ {}^G\mathbf{f}_1 \\ \vdots \\ {}^G\mathbf{f}_N \end{bmatrix} \tag{6}$$

Proof: See Appendix A.

Remark: The unobservable direction in (6) corresponds to the scale, as shown in Appendix B.

The physical interpretation of Theorem 1 is that, when the local acceleration is non-varying, one *cannot* distinguish the magnitude of the true body acceleration from that of the accelerometer bias, as both of them are, at least temporarily, constant. As a consequence, the magnitude of the true body acceleration can be arbitrary, leading to scale ambiguity.

[1] For clarity of presentation, we assume that the IMU-camera frames coincide. In practice, we estimate the IMU-camera extrinsics online.

At this point, we should note that in most cases in practice, a planar ground vehicle moves with (almost) constant acceleration, such as when following a straight line path with constant speed or acceleration, or when making turns along a circular arc with constant speed, etc. Based on Theorem 1, these motions render the scale estimated by the VINS inaccurate.

2.2 No Rotation

Consider that the platform has no rotational motion, i.e., the orientation remains the same across time:

$$ {}^{I_t}_{G}\mathbf{C} \triangleq \mathbf{C}({}^{I}\mathbf{q}_G(t)) \equiv {}^{I_0}_{G}\mathbf{C}, \quad \forall\, t \geq t_0 \tag{7}$$

where I_t denotes the IMU frame at time t. The VINS observability property, in this case, is stated in the following theorem:

Theorem 2: The linearized VINS model of (3)-(4) has the following *additional* unobservable directions, besides the global translation, *if and only if* the condition in (7) is satisfied:

$$\mathbf{N}_o = \begin{bmatrix} {}^{I_0}_{G}\mathbf{C} \\ \mathbf{0}_{3\times 3} \\ -\lfloor {}^{G}\mathbf{v}_{I_0} \rfloor \\ {}^{I_0}_{G}\mathbf{C}\lfloor {}^{G}\mathbf{g} \rfloor \\ -\lfloor {}^{G}\mathbf{p}_{I_0} \rfloor \\ -\lfloor {}^{G}\mathbf{f}_1 \rfloor \\ \vdots \\ -\lfloor {}^{G}\mathbf{f}_N \rfloor \end{bmatrix} \tag{8}$$

Proof: See Appendix C.

Remark: The unobservable directions in (8) correspond to the global orientation (all three dof instead of only the yaw), as shown in Appendix D. Note also that the global yaw unobservable direction (see (57) in [1]) can be expressed as $\mathbf{N}_o {}^{G}\mathbf{g}$, which reflects the fact that the 1-dof global yaw direction, defined as the rotation about the gravity vector ${}^{G}\mathbf{g}$, is contained in the 3-dof global orientation directions $\mathbf{N}_o$.

The physical interpretation of Theorem 2 is that, when there is no rotational motion, one *cannot* distinguish the direction of the local gravitational acceleration from that of the accelerometer bias, as both of them are, at least temporarily, constant. As a consequence, the roll and pitch angles become ambiguous.

The motion profile considered in Theorem 2 is the case typically followed by a robot moving on a straight line, or (for a holonomic vehicle) sliding sideways. In such cases, due to the lack of observability, the orientation estimates generated by the VINS become inaccurate.

A Proof of Theorem 1

In this section, we prove that the scale in (6) is an unobservable direction of the VINS model, if and only if the platform is moving with constant *local* linear acceleration [see (5)]. We follow the approach presented in [1], that examines the right null space of the observability matrix of the corresponding linearized VINS model. As is the case in [1], and for clarity of presentation, we include only one feature in the state vector (the extension to multiple features is straightforward).

As previously shown (see (51) in [1]), any block row, $\mathbf{M}_k$, of the observability matrix has the following structure:

$$\mathbf{M}_k = \mathbf{H}_k \mathbf{\Phi}_{k,1} = \mathbf{\Gamma}_1 \begin{bmatrix} \mathbf{\Gamma}_2 & \mathbf{\Gamma}_3 & -\delta t_k \mathbf{I}_3 & \mathbf{\Gamma}_4 & -\mathbf{I}_3 & \mathbf{I}_3 \end{bmatrix} \tag{9}$$

for any time $t_k \geq t_0$, with the matrices $\mathbf{\Gamma}_i, i = 1, \ldots, 4$, defined by (52)-(55) in [1]. From the property of the observability matrix, the scale direction, $\mathbf{N}_s$, is unobservable, if and only if, $\mathbf{M}_k \mathbf{N}_s = \mathbf{0}$ [3]. From (9) and (6), together with the definition of the matrices $\mathbf{\Gamma}_i$, we obtain:

$$\mathbf{M}_k\mathbf{N}_s = \mathbf{\Gamma}_1(-{}^G\mathbf{v}_{I_0}\delta t_k - \mathbf{\Gamma}_4{}^I\mathbf{a} - {}^G\mathbf{p}_{I_0} + {}^G\mathbf{f}) \tag{10}$$

$$\text{with } \mathbf{\Gamma}_4{}^I\mathbf{a} = \int_{t_0}^{t_k}\int_{t_0}^{s} {}^G_{I_\tau}\mathbf{C}\; d\tau ds \cdot {}^I\mathbf{a} \tag{11}$$

$$= \int_{t_0}^{t_k}\int_{t_0}^{s} {}^G_{I_\tau}\mathbf{C}{}^I\mathbf{a}\; d\tau ds \tag{12}$$

$$= \int_{t_0}^{t_k}\int_{t_0}^{s} {}^G_{I_\tau}\mathbf{C}{}^{I_\tau}\mathbf{a}(\tau)\; d\tau ds \tag{13}$$

$$= \int_{t_0}^{t_k}\int_{t_0}^{s} {}^G\mathbf{a}(\tau)\; d\tau ds \tag{14}$$

$$= \int_{t_0}^{t_k} ({}^G\mathbf{v}_{I_s} - {}^G\mathbf{v}_{I_0})\; ds \tag{15}$$

$$= {}^G\mathbf{p}_{I_k} - {}^G\mathbf{p}_{I_0} - {}^G\mathbf{v}_{I_0}\delta t_k \tag{16}$$

where the equality from (12) to (13) holds if and only if the constant acceleration assumption in (5) is satisfied. Substituting (16) into (10) yields:

$$\begin{aligned}\mathbf{M}_k\mathbf{N}_s &= \mathbf{\Gamma}_1({}^G\mathbf{f} - {}^G\mathbf{p}_{I_k}) = \mathbf{H}_{c,k}{}^{I_k}_G\mathbf{C}({}^G\mathbf{f} - {}^G\mathbf{p}_{I_k}) \\ &= \mathbf{H}_{c,k}{}^{I_k}\mathbf{f} = \mathbf{0}\end{aligned} \tag{17}$$

where the last equality holds since the camera perspective-projection Jacobian matrix, $\mathbf{H}_{c,k}$, has as its right null space the feature position in the IMU frame (see (30) in [1]).

Lastly, this new unobservable direction is in addition to the four directions corresponding to global translation and yaw, i.e., $\mathbf{N}_s$ and $\mathbf{N}_1$ in (57) of [1] are independent, since the 4th block element of $\mathbf{N}_1$ is zero while that of $\mathbf{N}_s$ is not.

B The Scale Unobservable Direction

In this section, we show that the unobservable direction in (6) corresponds to the scale. Assume that there exists a VINS state vector $\mathbf{x}$ and the corresponding measurements from the IMU [see (2)] and the camera [see (4)]. Consider the case where both the entire trajectory of the platform and the scene are "scaled up" by a factor of α, or equivalently, the global coordinate system $\{G\}$ is "shrunken down" by the factor α. This corresponds to a change of the original state $\mathbf{x}$ into a new state $\mathbf{x}'$. Specifically, as for the IMU's position, $\mathbf{p}_I$, and the feature's position, $\mathbf{f}_j$, with respect to $\{G\}$, the scale change can be written as:[2]

$${}^G\mathbf{p}'_I = \alpha{}^G\mathbf{p}_I \tag{18}$$

$${}^G\mathbf{f}'_j = \alpha{}^G\mathbf{f}_j, \quad j = 1,\ldots,N \tag{19}$$

where ${}^G\mathbf{p}'_I$ and ${}^G\mathbf{f}'_j$ are the new positions after the scaling. By taking the first and second-order time derivative on both sides of (18), we obtain the scaled velocity and body acceleration of the IMU as:

$${}^G\mathbf{v}'_I = \alpha{}^G\mathbf{v}_I \tag{20}$$

$${}^G\mathbf{a}'_I = \alpha{}^G\mathbf{a}_I \tag{21}$$

Note that, on the other hand, the scale change does not affect the IMU's orientation with respect to the global frame (as the "scale" referred here is with respect to translation only), i.e.:

$${}^I_G\mathbf{C}' = {}^I_G\mathbf{C} \tag{22}$$

and hence the rotational velocity remains the same as well:

$${}^I\boldsymbol{\omega}' = {}^I\boldsymbol{\omega} \tag{23}$$

[2]Note that the presented analysis holds true for any time $t \geq t_0$. Hence we omit the time index for the clarity of presentation.

Moreover, to ensure that this scale change is unobservable, the measurements from the IMU and the camera need to be unchanged. As for the camera observations, for each feature j, from (4) (18) (19) (22) we have:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \triangleq {}^I\mathbf{f}'_j = {}^I_G\mathbf{C}'({}^G\mathbf{f}'_j - {}^G\mathbf{p}'_I) = {}^I_G\mathbf{C}(\alpha {}^G\mathbf{f}_j - \alpha {}^G\mathbf{p}_I) = \alpha {}^I_G\mathbf{C}({}^G\mathbf{f}_j - {}^G\mathbf{p}_I) = \alpha {}^I\mathbf{f}_j = \alpha \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{24}$$

$$\Rightarrow \quad \mathbf{z}'_j = \frac{1}{z'}\begin{bmatrix} x' \\ y' \end{bmatrix} = \frac{1}{\alpha z}\begin{bmatrix} \alpha x \\ \alpha y \end{bmatrix} = \frac{1}{z}\begin{bmatrix} x \\ y \end{bmatrix} = \mathbf{z}_j \tag{25}$$

Hence, after scaling, the camera measurements do not change due to the perspective projection model. This result is to be expected, since a camera's observation is scale invariant, therefore it's insensitive to any scale change. As for the IMU measurements, we first examine the rotational velocity measured by the gyroscope. Since the gyroscope measurements [see (2)] need to stay the same before and after the scaling, i.e.:

$$\boldsymbol{\omega}_m = {}^I\boldsymbol{\omega} + \mathbf{b}_g = {}^I\boldsymbol{\omega}' + \mathbf{b}'_g \tag{26}$$

by substituting (23), we obtain:

$$\mathbf{b}'_g = \mathbf{b}_g \tag{27}$$

Similarly, for the linear acceleration measurements from the accelerometer, from (2) and (21) (22), we obtain:

$$\mathbf{a}_m = {}^I_G\mathbf{C}({}^G\mathbf{a}_I - {}^G\mathbf{g}) + \mathbf{b}_a = {}^I_G\mathbf{C}'({}^G\mathbf{a}'_I - {}^G\mathbf{g}) + \mathbf{b}'_a = {}^I_G\mathbf{C}(\alpha {}^G\mathbf{a}_I - {}^G\mathbf{g}) + \mathbf{b}'_a \tag{28}$$

$$\Rightarrow \quad \mathbf{b}'_a = \mathbf{b}_a - (\alpha - 1){}^I_G\mathbf{C}{}^G\mathbf{a}_I = \mathbf{b}_a - (\alpha - 1){}^I\mathbf{a} \tag{29}$$

Collecting the equations (22) (27) (20) (29) (18) (19), we put together the VINS state element changes due to the scaling, by a factor of α, as:

$$\begin{aligned} {}^I_G\mathbf{C}' &= {}^I_G\mathbf{C} \\ \mathbf{b}'_g &= \mathbf{b}_g \\ {}^G\mathbf{v}'_I &= \alpha {}^G\mathbf{v}_I = {}^G\mathbf{v}_I + (\alpha - 1){}^G\mathbf{v}_I \\ \mathbf{b}'_a &= \mathbf{b}_a - (\alpha - 1){}^I\mathbf{a} \\ {}^G\mathbf{p}'_I &= \alpha {}^G\mathbf{p}_I = {}^G\mathbf{p}_I + (\alpha - 1){}^G\mathbf{p}_I \\ {}^G\mathbf{f}'_j &= \alpha {}^G\mathbf{f}_j = {}^G\mathbf{f}_j + (\alpha - 1){}^G\mathbf{f}_j, \quad j = 1, \dots, N \end{aligned} \tag{30}$$

If we define the original and the new error state as $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{x}}'$, corresponding to the original and the new VINS state $\mathbf{x}$ and $\mathbf{x}'$ (see [1] for the definition of the VINS error state), respectively, then (30) can be rewritten in the error-state form as:

$$\begin{bmatrix} {}^I\delta\boldsymbol{\theta}'_G \\ \tilde{\mathbf{b}}'_g \\ {}^G\tilde{\mathbf{v}}'_I \\ \tilde{\mathbf{b}}'_a \\ {}^G\tilde{\mathbf{p}}'_I \\ {}^G\tilde{\mathbf{f}}'_1 \\ \vdots \\ {}^G\tilde{\mathbf{f}}'_N \end{bmatrix} = \begin{bmatrix} {}^I\delta\boldsymbol{\theta}_G \\ \tilde{\mathbf{b}}_g \\ {}^G\tilde{\mathbf{v}}_I + (\alpha - 1){}^G\mathbf{v}_I \\ \tilde{\mathbf{b}}_a - (\alpha - 1){}^I\mathbf{a} \\ {}^G\tilde{\mathbf{p}}_I + (\alpha - 1){}^G\mathbf{p}_I \\ {}^G\tilde{\mathbf{f}}_1 + (\alpha - 1){}^G\mathbf{f}_1 \\ \vdots \\ {}^G\tilde{\mathbf{f}}_N + (\alpha - 1){}^G\mathbf{f}_N \end{bmatrix} = \begin{bmatrix} {}^I\delta\boldsymbol{\theta}_G \\ \tilde{\mathbf{b}}_g \\ {}^G\tilde{\mathbf{v}}_I \\ \tilde{\mathbf{b}}_a \\ {}^G\tilde{\mathbf{p}}_I \\ {}^G\tilde{\mathbf{f}}_1 \\ \vdots \\ {}^G\tilde{\mathbf{f}}_N \end{bmatrix} + (\alpha - 1)\begin{bmatrix} \mathbf{0}_{3\times 1} \\ \mathbf{0}_{3\times 1} \\ {}^G\mathbf{v}_I \\ -{}^I\mathbf{a} \\ {}^G\mathbf{p}_I \\ {}^G\mathbf{f}_1 \\ \vdots \\ {}^G\mathbf{f}_N \end{bmatrix} \tag{31}$$

where we see that the right-most vector is exactly the same as in (6), hence

$$\tilde{\mathbf{x}}' = \tilde{\mathbf{x}} + (\alpha - 1)\mathbf{N}_s \tag{32}$$

To summarize, if the entire trajectory of the platform and the scene are scaled by a factor of α (as the starting point of this analysis), then the VINS error state (and hence the state) will be changed along the direction of $\mathbf{N}_s$ by a factor of $\alpha - 1$ [see (32)], *without* changing the measurements from the camera [see (25)] or the IMU [see (26) and (28)]. Moreover, it is obvious that the reverse statement holds true as well. Therefore, we conclude that the direction $\mathbf{N}_s$ in (6) is unobservable, and it corresponds to the scale change in terms of its physical meaning.

C Proof of Theorem 2

In what follows, we prove that the 3-dof global orientation in (8) is an unobservable direction of the VINS model, if and only if the platform does not rotate [see (7)]. Similarly to the proof presented in Appendix A, in this case, we need to show that $\mathbf{M}_k\mathbf{N}_o = \mathbf{0}$. From (9) and (8), together with the definition of the matrices $\mathbf{\Gamma}_i, i = 1, \dots, 4$, we obtain:

$$\mathbf{M}_k\mathbf{N}_o = \mathbf{\Gamma}_1(\mathbf{\Gamma}_4\, {}^{I_0}_{G}\mathbf{C} - \frac{1}{2}\delta t_k^2\,\mathbf{I}_3)\lfloor {}^{G}\mathbf{g}\rfloor \tag{33}$$

$$= \mathbf{\Gamma}_1(\int_{t_0}^{t_k}\int_{t_0}^{s} {}^{G}_{I_\tau}\mathbf{C}\; d\tau ds \cdot {}^{I_0}_{G}\mathbf{C} - \frac{1}{2}\delta t_k^2\,\mathbf{I}_3)\lfloor {}^{G}\mathbf{g}\rfloor \tag{34}$$

$$= \mathbf{\Gamma}_1(\int_{t_0}^{t_k}\int_{t_0}^{s} {}^{G}_{I_0}\mathbf{C}\; d\tau ds \cdot {}^{I_0}_{G}\mathbf{C} - \frac{1}{2}\delta t_k^2\,\mathbf{I}_3)\lfloor {}^{G}\mathbf{g}\rfloor \tag{35}$$

$$= \mathbf{\Gamma}_1(\int_{t_0}^{t_k}\int_{t_0}^{s} 1\; d\tau ds \cdot {}^{G}_{I_0}\mathbf{C}{}^{I_0}_{G}\mathbf{C} - \frac{1}{2}\delta t_k^2\,\mathbf{I}_3)\lfloor {}^{G}\mathbf{g}\rfloor$$

$$= \mathbf{\Gamma}_1(\frac{1}{2}\delta t_k^2\,\mathbf{I}_3 - \frac{1}{2}\delta t_k^2\,\mathbf{I}_3)\lfloor {}^{G}\mathbf{g}\rfloor = \mathbf{0} \tag{36}$$

where the equality from (34) to (35) holds if and only if the no rotation (i.e., constant orientation) assumption in (7) is satisfied.

Lastly, these new unobservable directions are in addition to the three directions corresponding to global translation, i.e., $\mathbf{N}_o$ and $\mathbf{N}_{t,1}$ in (57) of [1] are independent, since the first block element of $\mathbf{N}_{t,1}$ is zero while that of $\mathbf{N}_o$ is a (full-rank) rotational matrix.

D The Orientation Unobservable Directions

In this section, we show that the unobservable directions in (8) correspond to the 3-dof global orientation. Assume that there exists a VINS state vector $\mathbf{x}$ and the corresponding measurements from the IMU [see (2)] and the camera [see (4)]. Consider the case where the global frame $\{G\}$ is rotated by a *small* angle $\delta\boldsymbol{\phi}$ into a new global frame $\{G'\}$, where $\delta\boldsymbol{\phi}$ is a 3×1 vector whose direction and magnitude represent the axis and angle of the rotation, respectively. Hence,

$$ {}^{G}_{G'}\mathbf{C} = \mathbf{C}(\delta\boldsymbol{\phi}) \simeq \mathbf{I}_3 - \lfloor\delta\boldsymbol{\phi}\rfloor \tag{37}$$

by the small-angle approximation of the rotational matrix based on the assumption that the amount of rotation is small. Due to this change of the global frame (from $\{G\}$ to $\{G'\}$), the original VINS state, $\mathbf{x}$, which is expressed with respect to $\{G\}$, is now changed to a new state, $\mathbf{x}'$, which is expressed with respect to $\{G'\}$. Specifically, as for the orientation of the IMU:[3]

$$ {}^{I}_{G'}\mathbf{C} = {}^{I}_{G}\mathbf{C}\,{}^{G}_{G'}\mathbf{C} = {}^{I}_{G}\mathbf{C}(\mathbf{I}_3 - \lfloor\delta\boldsymbol{\phi}\rfloor) = (\mathbf{I}_3 - \lfloor{}^{I}_{G}\mathbf{C}\,\delta\boldsymbol{\phi}\rfloor){}^{I}_{G}\mathbf{C} \tag{38}$$

Since the transformation involves only rotation, the new position state of the IMU can be obtained as:

$$ {}^{G'}\mathbf{p}_I = {}^{G'}_{G}\mathbf{C}\,{}^{G}\mathbf{p}_I = {}^{G}_{G'}\mathbf{C}^T\,{}^{G}\mathbf{p}_I = (\mathbf{I}_3 - \lfloor\delta\boldsymbol{\phi}\rfloor)^T\,{}^{G}\mathbf{p}_I = (\mathbf{I}_3 + \lfloor\delta\boldsymbol{\phi}\rfloor){}^{G}\mathbf{p}_I = {}^{G}\mathbf{p}_I + \lfloor\delta\boldsymbol{\phi}\rfloor{}^{G}\mathbf{p}_I = {}^{G}\mathbf{p}_I - \lfloor{}^{G}\mathbf{p}_I\rfloor\delta\boldsymbol{\phi} \tag{39}$$

and similarly for the feature's position:

$$ {}^{G'}\mathbf{f}_j = {}^{G'}_{G}\mathbf{C}\,{}^{G}\mathbf{f}_j = {}^{G}\mathbf{f}_j - \lfloor{}^{G}\mathbf{f}_j\rfloor\delta\boldsymbol{\phi}, \quad j = 1, \dots, N \tag{40}$$

By taking the first-order time derivative on both sides of (39), we obtain the new velocity state of the IMU:

$$ {}^{G'}\mathbf{v}_I = {}^{G}\mathbf{v}_I - \lfloor{}^{G}\mathbf{v}_I\rfloor\delta\boldsymbol{\phi} \tag{41}$$

Note that, on the other hand, the transformation of the *global* frame does not affect the trajectory, and hence the motion, of the IMU when expressed in the IMU's *local* frame of reference $\{I\}$. Therefore, the *local*

[3] Note that the presented analysis holds true for any time $t \geq t_0$. Hence we omit the time index for the clarity of presentation.

rotational velocity and linear acceleration of the IMU are the same before and after the transformation of the global frame, i.e.:

$$ {}^{I}\boldsymbol{\omega}' = {}^{I}\boldsymbol{\omega} \tag{42} $$

$$ {}^{I}\mathbf{a}' = {}^{I}\mathbf{a} \tag{43} $$

Moreover, to ensure that the change of the global frame's orientation is unobservable, the measurements from the IMU and the camera need to be unchanged. As for the camera observations, for each feature j, from (4) (38) (39) (40) we have:

$$ \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \triangleq {}^{I}\mathbf{f}'_j = {}^{I}_{G'}\mathbf{C}({}^{G'}\mathbf{f}_j - {}^{G'}\mathbf{p}_I) = {}^{I}_{G}\mathbf{C}\, {}^{G}_{G'}\mathbf{C}({}^{G'}_{G}\mathbf{C}\, {}^{G}\mathbf{f}_j - {}^{G'}_{G}\mathbf{C}\, {}^{G}\mathbf{p}_I) = {}^{I}_{G}\mathbf{C}({}^{G}\mathbf{f}_j - {}^{G}\mathbf{p}_I) = {}^{I}\mathbf{f}_j = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{44} $$

$$ \Rightarrow \quad \mathbf{z}'_j = \frac{1}{z'} \begin{bmatrix} x' \\ y' \end{bmatrix} = \frac{1}{z} \begin{bmatrix} x \\ y \end{bmatrix} = \mathbf{z}_j \tag{45} $$

Hence, after the transformation of the global frame, the camera measurements do not change. This result is to be expected, since a camera's observation depends only on the relative position of the feature with respect to the camera's frame, therefore it's insensitive to any change in the global frame itself. As for the IMU measurements, we first examine the rotational velocity measured by the gyroscope. Since the gyroscope measurements [see (2)] need to stay the same before and after the transformation of the global frame, i.e.:

$$ \boldsymbol{\omega}_m = {}^{I}\boldsymbol{\omega} + \mathbf{b}_g = {}^{I}\boldsymbol{\omega}' + \mathbf{b}'_g \tag{46} $$

by substituting (42), we obtain:

$$ \mathbf{b}'_g = \mathbf{b}_g \tag{47} $$

Similarly, for the linear acceleration measurements from the accelerometer, from (2) and the definition that ${}^{I}\mathbf{a} = {}^{I}_{G}\mathbf{C}\, {}^{G}\mathbf{a}_I$, we obtain:

$$ \mathbf{a}_m = {}^{I}_{G}\mathbf{C}({}^{G}\mathbf{a}_I - {}^{G}\mathbf{g}) + \mathbf{b}_a = {}^{I}\mathbf{a} - {}^{I}_{G}\mathbf{C}\, {}^{G}\mathbf{g} + \mathbf{b}_a = {}^{I}\mathbf{a}' - {}^{I}_{G'}\mathbf{C}\, {}^{G'}\mathbf{g}' + \mathbf{b}'_a \tag{48} $$

Substituting (43) yields:

$$ \mathbf{b}'_a = \mathbf{b}_a + {}^{I}_{G'}\mathbf{C}\, {}^{G'}\mathbf{g}' - {}^{I}_{G}\mathbf{C}\, {}^{G}\mathbf{g} \tag{49} $$

Note that according to the definition, the gravity vector, $\mathbf{g}$, is a known *constant* in the corresponding global frame, i.e., $\mathbf{g}$ is *fixed* with respect to the global frame. Hence, as the global frame rotates from $\{G\}$ to $\{G'\}$, the gravity vector rotates simultaneously from $\mathbf{g}$ to $\mathbf{g}'$, such that:

$$ {}^{G'}\mathbf{g}' = {}^{G}\mathbf{g} \tag{50} $$

Substituting (50) and (38) into (49), we obtain:

$$ \begin{aligned} \mathbf{b}'_a &= \mathbf{b}_a + {}^{I}_{G'}\mathbf{C}\, {}^{G'}\mathbf{g}' - {}^{I}_{G}\mathbf{C}\, {}^{G}\mathbf{g} = \mathbf{b}_a + {}^{I}_{G'}\mathbf{C}\, {}^{G}\mathbf{g} - {}^{I}_{G}\mathbf{C}\, {}^{G}\mathbf{g} = \mathbf{b}_a + ({}^{I}_{G'}\mathbf{C} - {}^{I}_{G}\mathbf{C})^{G}\mathbf{g} \\ &= \mathbf{b}_a + ({}^{I}_{G}\mathbf{C} - {}^{I}_{G}\mathbf{C}\lfloor\delta\boldsymbol{\phi}\rfloor - {}^{I}_{G}\mathbf{C})^{G}\mathbf{g} = \mathbf{b}_a - {}^{I}_{G}\mathbf{C}\lfloor\delta\boldsymbol{\phi}\rfloor^{G}\mathbf{g} = \mathbf{b}_a + {}^{I}_{G}\mathbf{C}\lfloor{}^{G}\mathbf{g}\rfloor\delta\boldsymbol{\phi} \end{aligned} \tag{51} $$

Collecting the equations (38) (47) (41) (51) (39) (40), we put together the VINS state element changes due to the rotation of the global frame, by a small angle $\delta\boldsymbol{\phi}$, as:

$$ \begin{aligned} {}^{I}_{G'}\mathbf{C} &= (\mathbf{I}_3 - \lfloor{}^{I}_{G}\mathbf{C}\,\delta\boldsymbol{\phi}\rfloor){}^{I}_{G}\mathbf{C} \\ \mathbf{b}'_g &= \mathbf{b}_g \\ {}^{G'}\mathbf{v}_I &= {}^{G}\mathbf{v}_I - \lfloor{}^{G}\mathbf{v}_I\rfloor\delta\boldsymbol{\phi} \\ \mathbf{b}'_a &= \mathbf{b}_a + {}^{I}_{G}\mathbf{C}\lfloor{}^{G}\mathbf{g}\rfloor\delta\boldsymbol{\phi} \\ {}^{G'}\mathbf{p}_I &= {}^{G}\mathbf{p}_I - \lfloor{}^{G}\mathbf{p}_I\rfloor\delta\boldsymbol{\phi} \\ {}^{G'}\mathbf{f}_j &= {}^{G}\mathbf{f}_j - \lfloor{}^{G}\mathbf{f}_j\rfloor\delta\boldsymbol{\phi}, \quad j = 1, \ldots, N \end{aligned} \tag{52} $$

If we define the original and the new error state as $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{x}}'$, corresponding to the original and the new VINS state $\mathbf{x}$ and $\mathbf{x}'$ (see [1] for the definition of the VINS error state), respectively, then (52) can be rewritten in the error-state form as:

$$\begin{bmatrix} {}^I\delta\boldsymbol{\theta}_{G'} \\ \tilde{\mathbf{b}}'_g \\ {}^{G'}\tilde{\mathbf{v}}_I \\ \tilde{\mathbf{b}}'_a \\ {}^{G'}\tilde{\mathbf{p}}_I \\ {}^{G'}\tilde{\mathbf{f}}_1 \\ \vdots \\ {}^{G'}\tilde{\mathbf{f}}_N \end{bmatrix} = \begin{bmatrix} {}^I\delta\boldsymbol{\theta}_G + {}^I_G\mathbf{C}\,\delta\boldsymbol{\phi} \\ \tilde{\mathbf{b}}_g \\ {}^G\tilde{\mathbf{v}}_I - \lfloor {}^G\mathbf{v}_I \rfloor \delta\boldsymbol{\phi} \\ \tilde{\mathbf{b}}_a + {}^I_G\mathbf{C}\lfloor {}^G\mathbf{g} \rfloor \delta\boldsymbol{\phi} \\ {}^G\tilde{\mathbf{p}}_I - \lfloor {}^G\mathbf{p}_I \rfloor \delta\boldsymbol{\phi} \\ {}^G\tilde{\mathbf{f}}_1 - \lfloor {}^G\mathbf{f}_1 \rfloor \delta\boldsymbol{\phi} \\ \vdots \\ {}^G\tilde{\mathbf{f}}_N - \lfloor {}^G\mathbf{f}_N \rfloor \delta\boldsymbol{\phi} \end{bmatrix} = \begin{bmatrix} {}^I\delta\boldsymbol{\theta}_G \\ \tilde{\mathbf{b}}_g \\ {}^G\tilde{\mathbf{v}}_I \\ \tilde{\mathbf{b}}_a \\ {}^G\tilde{\mathbf{p}}_I \\ {}^G\tilde{\mathbf{f}}_1 \\ \vdots \\ {}^G\tilde{\mathbf{f}}_N \end{bmatrix} + \begin{bmatrix} {}^I_G\mathbf{C} \\ \mathbf{0}_{3\times 3} \\ -\lfloor {}^G\mathbf{v}_I \rfloor \\ {}^I_G\mathbf{C}\lfloor {}^G\mathbf{g} \rfloor \\ -\lfloor {}^G\mathbf{p}_I \rfloor \\ -\lfloor {}^G\mathbf{f}_1 \rfloor \\ \vdots \\ -\lfloor {}^G\mathbf{f}_N \rfloor \end{bmatrix} \delta\boldsymbol{\phi} \tag{53}$$

where we see that the matrix multiplied with $\delta\boldsymbol{\phi}$ on the right-hand side is exactly the same as in (8), hence

$$\tilde{\mathbf{x}}' = \tilde{\mathbf{x}} + \mathbf{N}_o\delta\boldsymbol{\phi} \tag{54}$$

To summarize, if the global frame is rotated by a small angle $\delta\boldsymbol{\phi}$ (as the starting point of this analysis), or equivalently, the entire trajectory of the platform and the scene are rotated by $-\delta\boldsymbol{\phi}$ with respect to the global frame, then the VINS error state (and hence the state) will be changed along the direction as a linear combination of the columns of $\mathbf{N}_o$ [see (54)], *without* changing the measurements from the camera [see (45)] or the IMU [see (46) and (48)]. Moreover, it is obvious that the reverse statement holds true as well. Therefore, we conclude that the directions $\mathbf{N}_o$ in (8) are unobservable, and they correspond to the change of the 3-dof global orientation in terms of the physical meaning. In particular, if $\delta\boldsymbol{\phi} = ||\delta\boldsymbol{\phi}||\mathbf{e}_1$, it would correspond to a rotation about the global frame's x-axis, i.e., a change in the roll angle. Similarly for the pitch and yaw angles. Hence, the three columns of $\mathbf{N}_o$ correspond to the roll, pitch, and yaw angle change, respectively, of the orientation of the IMU's frame with respect to the global frame.

References

[1] J. A. Hesch, D. G. Kottas, S. L. Bowman, and S. I. Roumeliotis, "Consistency analysis and improvement of vision-aided inertial navigation," *IEEE Trans. on Robotics*, vol. 30, no. 1, pp. 158–176, Feb. 2014.

[2] A. Martinelli, "Vision and imu data fusion: Closed-form solutions for attitude, speed, absolute scale, and bias determination," *IEEE Trans. on Robotics*, vol. 28, no. 1, pp. 44–60, Feb. 2012.

[3] Z. Chen, K. Jiang, and J. C. Hung, "Local observability matrix and its application to observability analyses," in *Proc. of 16th Annual Conference IEEE Industrial Electronics Society*, Pacific Grove, CA, Nov. 27–30 1990, pp. 100–103.