Online IMU Intrinsic Calibration:
Is It Necessary?

Yulin Yang[*], Patrick Geneva[*], Xingxing Zuo[†] and Guoquan Huang[*]
[*]Robot Perception and Navigation Group, University of Delaware, DE, USA.
Email: {yuyang,pgeneva,ghuang}@udel.edu
[†]Institute of Cyber-System and Control, Zhejiang University, Hangzhou, China.
Email: xingxingzuo@zju.edu.cn

Abstract—This paper addresses the problem of visual-inertial self-calibration while focusing on the necessity of online IMU intrinsic calibration. To this end, we perform observability analysis for visual-inertial navigation systems (VINS) with four different inertial model variants containing intrinsic parameters that encompass one commonly used IMU model for low-cost inertial sensors. The analysis theoretically confirms what is intuitively believed in the literature, that is, the IMU intrinsics are observable given fully-excited 6-axis motion. Moreover, we, for the first time, identify 6 primitive degenerate motions for IMU intrinsic calibration. Each degenerate motion profile will cause a set of intrinsic parameters to be unobservable and any combination of these degenerate motions are still degenerate. This result holds for all four inertial model variants and has *significant* implications on the necessity to perform online IMU intrinsic calibration in many robotic applications. Extensive simulations and real-world experiments are performed to validate both our observability analysis and degenerate motion analysis.

I. INTRODUCTION

Cameras and IMUs are becoming ubiquitous, visual-inertial navigation system (VINS) [11] has gained great popularity in 6DoF pose tracking for autonomous robots – such as micro aerial vehicles (MAV) [25], autonomous-driving cars [7], and unmanned ground vehicles (UGV) [32] – during the past decades. Many efficient and robust VINS algorithms have been developed in recent years (e.g., [20, 14, 4, 21, 6, 3, 9]). There are many factors attribute to VINS performance, among which the sound and accurate sensor calibration – including the rigid transformation between sensors, time offset between IMU-camera, and camera and IMU intrinsics – is crucial.

Extensive works have studied sensor calibration in VINS, both offline and online. For instance, Mirzaei et al. [19] performed the observability analysis for rigid transformation between IMU and camera. They showed that under certain motions, the rigid transformation are not fully observable. Furgale et al. [5] designed Kalibr, a continuous-time interpolation based batch estimator, for IMU-camera extrinsics, time offset and camera intrinsics calibration. Rehder et al. [22] extended Kalibr to incorporate IMU intrinsics (including scaling parameters, axis misalignment, and g-sensitivity). However, most of these approaches need calibration targets (e.g., marker board) and are only suitable for offline non-real-time calibration. Recently, Li et al. [17] incorporated IMU-camera extrinsics, time offset, rolling-shutter read-out time, camera and IMU intrinsics into the multi-state constraint Kalman filter (MSCKF) framework [20]. They showed that these parameters can converge in simulation and demonstrated the system on a real-world experiment. This work is closest to ours but did not provide any observability or degenerate motion analysis, which is the focus and contribution of our work along with extensive validations.

Online sensor calibration in VINS has also attracted significant attentions in recent years, due to its promise to handle poor prior calibration and environmental effects such as varying temperature, humidity, vibrations, non-rigid mounting, and other phenomenons which can degrade the state estimates in the case that the calibration is treated to be true. For example, Li. et al. [16, 15] showed that online extrinsic and time offset calibration within the MSCKF improves localization accuracy. Eckenhoff et al. [2] incorporated the IMU-camera extrinsics, time offset and the camera intrinsic calibration into a multi-camera VINS and showed that online calibration improves estimation accuracy and robustness.

However, blindly performing online calibration is dangerous, as in most cases domain knowledge on specific motions and prior distribution choices are needed to make sure consistent calibration. Schneider et al. [23] introduced an information theoretic metric for selecting the most informative parts of datasets for the task of calibration to prevent uninformative motion segments from degrading the calibration consistency, which is motivated by the fact that not all segments of the trajectory are suitable for the calibration task. Yang et al. [29] identified several degenerate motions for online IMU-camera extrinsic and time offset calibration and these degenerate motions should be avoided when operating VINS for online calibration. In this work, however, we investigate the degenerate motions when jointly estimating the IMU intrinsic parameters, in particular, when deploying VINS on mobile robots which typically have constrained motions. For example, aerial and ground ve-
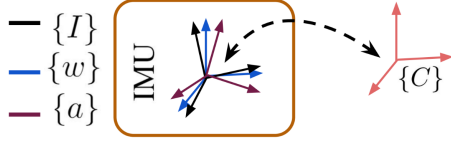
Fig. 1: An IMU sensor composed of accelerometer and gyroscope. The base "inertial" frame can be determined to coincide with either accelerometer frame $\{a\}$ or gyroscope frame $\{w\}$. $\{C\}$ represents the camera frame which has an additional translation.

hicles can only perform a few motion profiles due to their under-actuation, and can easily "fall" into degenerate conditions for calibration. By noting that most literature on IMU intrinsic calibration is limited to either handheld or trajectory segments involving rich motion information [17, 23], this promotes an interesting question: *What are the degenerate motions for IMU intrinsic calibration, and does it prohibit the online IMU intrinsic calibration, especially for underactuated autonomous vehicles?*

In this paper, we aim to respond to these questions by investigating in-depth the observability analysis for visual-inertial self-calibration and performing degenerate motion analysis for IMU intrinsic calibration. In particular, the main contributions of this work include:

- We propose four intrinsic model variants for IMU calibration based on [23] and analytically derive IMU integration and Jacobians .
- We provide observability analysis for these four inertial intrinsic model variants and, for the first time, offer an analysis of the degenerate motions that cause IMU intrinsic parameters unobservable.
- We conduct extensive simulations to verify the convergence of these intrinsic models in the case of full 6-axis motion and validate the identified degenerate motions. We also evaluate the proposed method on different real-world datasets, each of which is unique in its motion profile, allowing for a further discussion on the necessity of online IMU intrinsic calibration.

The paper is organized as follows: in Section II we go over the four inertial model variants; in Section III we derive the full system state transition, and then in Sections IV and V we present the observability analysis and the identified degenerate motions. Finally in Sections VI and VII, we validate our findings and provide some final remarks in Sections VIII and IX.

II. SYSTEM MODELS

A. IMU Intrinsic Model

We define an IMU as containing two separate frames of reference (see Fig. 1): gyroscope frame $\{w\}$, accelerometer frame $\{a\}$. The base "inertial" frame $\{I\}$ should be determined to coincide with either $\{w\}$ or $\{a\}$. Slightly different from the model defined in [23], we define the raw angular velocity reading $^w\boldsymbol{\omega}_m$ from the gyroscope and linear acceleration readings $^a\mathbf{a}_m$ from

accelerometer as:

$$^w\boldsymbol{\omega}_m = \mathbf{T}_w {}^w_I\mathbf{R}^I\boldsymbol{\omega} + \mathbf{b}_g + \mathbf{n}_g \tag{1}$$

$$^a\mathbf{a}_m = \mathbf{T}_a {}^a_I\mathbf{R}\left({}^I\mathbf{a} + {}^I_G\mathbf{R}^G\mathbf{g}\right) + \mathbf{b}_a + \mathbf{n}_a \tag{2}$$

where $\mathbf{T}_w$ and $\mathbf{T}_a$ are invertible 3×3 matrices which represent the scale imperfection and axis misalignment for $\{w\}$ and $\{a\}$, respectively. $^G\mathbf{g} = [0, 0, 9.81]^\top$. $^w_I\mathbf{R}$ and $^a_I\mathbf{R}$ denote the rotation from the gyroscope frame and acceleration frame to base "inertial" frame $\{I\}$, respectively. Note that, if we choose $\{I\}$ coincides with $\{w\}$, then $^w_I\mathbf{R} = \mathbf{I}_3$. Otherwise, $^a_I\mathbf{R} = \mathbf{I}_3$. $\mathbf{b}_g$ and $\mathbf{b}_a$ are the gyroscope and accelerometer biases, which are both modeled as random walks, and $\mathbf{n}_g$ and $\mathbf{n}_a$ are the zero-mean Gaussian noises contaminating the measurements. Inspired by [17, 23], we do not take into account gravity sensitivity [17] and the translation between the gyroscope and accelerometer frames. We can write the true (or corrected) angular velocity $^I\boldsymbol{\omega}$ and linear acceleration $^I\mathbf{a}$ as:

$$^I\boldsymbol{\omega} = {}^I_w\mathbf{R}\mathbf{D}_w\left(^w\boldsymbol{\omega}_m - \mathbf{b}_g - \mathbf{n}_g\right) \tag{3}$$

$$^I\mathbf{a} = {}^I_a\mathbf{R}\mathbf{D}_a\left(^a\mathbf{a}_m - \mathbf{b}_a - \mathbf{n}_a\right) - {}^I_G\mathbf{R}^G\mathbf{g} \tag{4}$$

where $\mathbf{D}_w = \mathbf{T}_w^{-1}$ and $\mathbf{D}_a = \mathbf{T}_a^{-1}$. In practice we calibrate $\mathbf{D}_a$, $\mathbf{D}_w$ and $^I_a\mathbf{R}$ (or $^I_w\mathbf{R}$) to prevent the need to have an unneeded matrix inversion in the measurement equation. While the IMU reading is defined as Eq. (3) and (4), we can only calibrate either $^I_w\mathbf{R}$ or $^I_a\mathbf{R}$, since the base "inertial" frame must coincide with either $\{w\}$ or $\{a\}$. Note that if both $^I_w\mathbf{R}$ and $^I_a\mathbf{R}$ are calibrated, it will make the rotation between the IMU and camera [17] unobservable. This is experimentally validated in simulation in Section VI-D.

B. Four Variants of IMU Intrinsic Model

In the following, we present four different IMU intrinsic model variants among which each has 15 degrees of freedom. We summarize them and evaluate their performances in online filter-based VIO state estimation in Section VI and VII. The four variants can be listed as:

- *imu1*: Variant one includes 6 parameters for $\mathbf{D}_{w6}$, the rotation $^I_w\mathbf{R}$, and 6 parameters for $\mathbf{D}_{a6}$. $\mathbf{D}_{a6}$ and $\mathbf{D}_{w6}$ are uptriangular matrices with the structure as $\mathbf{D}_{*6} = \begin{bmatrix} d_{*11} & d_{*12} & d_{*13} \\ 0 & d_{*22} & d_{*23} \\ 0 & 0 & d_{*33} \end{bmatrix}$.
- *imu2*: Variant two contains $\mathbf{D}_{a6}$, the rotation $^I_a\mathbf{R}$, and $\mathbf{D}_{w6}$. We note that this coincides with the model used in [23].
- *imu3*: Variant three combines Variant one's $\mathbf{D}_{w6}$ and $^I_w\mathbf{R}$ into a general 3×3 matrix containing 9 parameters in total. Thus, in this variant we estimate the upper-triangle $\mathbf{D}_{a6}$ and a full matrix $\mathbf{D}_{w9}$ as

$$\mathbf{D}_{*9} = \begin{bmatrix} d_{*11} & d_{*12} & d_{*13} \\ d_{*21} & d_{*22} & d_{*23} \\ d_{*31} & d_{*31} & d_{*33} \end{bmatrix} \tag{5}$$

- *imu4*: Variant four is an extension of variant two with a combination of the $\mathbf{D}_{a6}$ and $_a^I\mathbf{R}$. Thus, in this variant we estimate the upper-triangle $\mathbf{D}_{w6}$ and a full matrix $\mathbf{D}_{a9}$, as defined in Eq. (5).

We note that *imu2* is used in the rest of paper for the system derivations and observability analysis.

C. System Dynamic Model

The state vector $\mathbf{x}$, containing the IMU state $\mathbf{x}_I$, IMU intrinsics $\mathbf{x}_{in}$ and feature state $\mathbf{x}_f$, is described as:

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_I^\top & \mathbf{x}_{in}^\top & \mathbf{x}_f^\top \end{bmatrix}^\top \tag{6}$$

$$\mathbf{x}_I = \begin{bmatrix} _G^I\bar{q}^\top & ^G\mathbf{p}_I^\top & ^G\mathbf{v}_I^\top & \mathbf{b}_g^\top & \mathbf{b}_a^\top \end{bmatrix}^\top \tag{7}$$

$$\mathbf{x}_{in} = \begin{bmatrix} \mathbf{x}_{Dw}^\top & \mathbf{x}_{Da}^\top & _a^I\bar{q}^\top \end{bmatrix}^\top \tag{8}$$

where $_G^I\bar{q}$ denotes quaternion with JPL convention [26] and corresponds to the rotation matrix $_G^I\mathbf{R}$, which represents the rotation from $\{G\}$ to $\{I\}$. $^G\mathbf{p}_I$ and $^G\mathbf{v}_I$ denote the IMU position and velocity in $\{G\}$. $\mathbf{x}_{Dw}$ and $\mathbf{x}_{Da}$ represent the column-wise non-zero IMU intrinsic parameters, see II-B. Specifically, they are both defined as $\mathbf{x}_{D*} = \begin{bmatrix} d_{*11} & d_{*12} & d_{*22} & d_{*13} & d_{*23} & d_{*33} \end{bmatrix}^\top$. The corresponding error states can be defines as $\tilde{\mathbf{x}} = \mathbf{x} - \hat{\mathbf{x}}$, where $\hat{\mathbf{x}}$ denotes the current best estimate. Note that for quaternion, we use the quaternion left multiplicative error $\bar{q} \approx \begin{bmatrix} \frac{1}{2}\delta\boldsymbol{\theta}^\top & 1 \end{bmatrix}^\top \otimes \hat{\bar{q}}$, where $\otimes$ denotes quaternion multiplication [26].

The evolution of this system can be written as [26]:

$$_G^I\dot{\bar{q}} = \frac{1}{2}\boldsymbol{\Omega}(^I\boldsymbol{\omega})_G^I\bar{q}, \ ^G\dot{\mathbf{p}}_I = {}^G\mathbf{v}_I, \ ^G\dot{\mathbf{v}}_I = {}_I^G\mathbf{R}^I\mathbf{a} \tag{9}$$

$$\dot{\mathbf{b}}_g = \mathbf{n}_{wg}, \ \dot{\mathbf{b}}_a = \mathbf{n}_{wa}, \ \dot{\mathbf{x}}_{in} = \mathbf{0}_{15\times 1}, \ \dot{\mathbf{x}}_f = \mathbf{0}_{3\times 1} \tag{10}$$

where $\mathbf{n}_{wg}$ and $\mathbf{n}_{wa}$ are zero-mean Gaussian noises driving $\mathbf{b}_g$ and $\mathbf{b}_a$, respectively.

D. Camera Measurement Model

Leveraging a simplified model from OpenVINS project [6] with the global 3d point representation (i.e. $\mathbf{x}_f = {}^G\mathbf{p}_f$), the feature measurements can be written:

$$\mathbf{z}_c = \begin{bmatrix} ^C\mathbf{p}_{fx}/^C\mathbf{p}_{fz} \\ ^C\mathbf{p}_{fy}/^C\mathbf{p}_{fz} \end{bmatrix} + \mathbf{n}_c \tag{11}$$

$$^C\mathbf{p}_f = {}_I^C\mathbf{R}_G^I\mathbf{R}\left(^G\mathbf{p}_f - {}^G\mathbf{p}_I\right) + {}^C\mathbf{p}_I \tag{12}$$

where $\mathbf{n}_c$ is the Gaussian image noise and the measurement Jacobian is: $\mathbf{H} = \mathbf{H}_\pi\mathbf{H}_c$:

$$\mathbf{H}_\pi = \frac{1}{^C\hat{\mathbf{p}}_{fz}^2}\begin{bmatrix} ^C\hat{\mathbf{p}}_{fz} & 0 & -^C\hat{\mathbf{p}}_{fx} \\ 0 & ^C\hat{\mathbf{p}}_{fz} & -^C\hat{\mathbf{p}}_{fy} \end{bmatrix} \tag{13}$$

$$\mathbf{H}_c = {}_I^C\hat{\mathbf{R}}_G^I\hat{\mathbf{R}}\begin{bmatrix} \lfloor ^G\hat{\mathbf{p}}_f - {}^G\hat{\mathbf{p}}_I\rfloor_I^G\hat{\mathbf{R}} & -\mathbf{I}_3 & \mathbf{0}_{3\times 24} & \mathbf{I}_3 \end{bmatrix} \tag{14}$$

III. ANALYTIC IMU INTEGRATION WITH STATE TRANSITION MATRIX

In the following section, the analytic IMU integration and corresponding error state transition matrix are presented. To simplify, we define $^w\hat{\boldsymbol{\omega}}$ and $^a\hat{\mathbf{a}}$ as:

$$^w\hat{\boldsymbol{\omega}} = {}^w\boldsymbol{\omega}_m - \hat{\mathbf{b}}_g = \begin{bmatrix} ^w\hat{w}_1 & ^w\hat{w}_2 & ^w\hat{w}_3 \end{bmatrix}^\top \tag{15}$$

$$^a\hat{\mathbf{a}} = {}^a\mathbf{a}_m - \hat{\mathbf{b}}_a = \begin{bmatrix} ^a\hat{a}_1 & ^a\hat{a}_2 & ^a\hat{a}_3 \end{bmatrix}^\top \tag{16}$$

Note that $^w\hat{\boldsymbol{\omega}}$ and $^a\hat{\mathbf{a}}$ are assumed to be constant between time interval t_k to t_{k+1}. Hence, we can have:

$$^I\hat{\boldsymbol{\omega}} = \hat{\mathbf{D}}_w^w\hat{\boldsymbol{\omega}}, \ ^I\hat{\mathbf{a}} = {}_a^I\hat{\mathbf{R}}\hat{\mathbf{D}}_a^a\hat{\mathbf{a}} \tag{17}$$

A. Analytic State Integration

Based on the above assumption, we can compute the integration of IMU state from t_k to t_{k+1} as:

$$_G^{I_{k+1}}\hat{\mathbf{R}} = {}_{I_k}^{I_{k+1}}\hat{\mathbf{R}}_G^{I_k}\hat{\mathbf{R}} \simeq \exp\left(-^{I_k}\hat{\boldsymbol{\omega}}\Delta t_k\right)_G^{I_k}\hat{\mathbf{R}} \tag{18}$$

$$^G\hat{\mathbf{p}}_{I_{k+1}} \simeq {}^G\hat{\mathbf{p}}_{I_k} + {}^G\hat{\mathbf{v}}_{I_k}\Delta t_k + {}_{I_k}^G\hat{\mathbf{R}}\boldsymbol{\Xi}_2{}^{I_k}\hat{\mathbf{a}} - \frac{1}{2}{}^G\mathbf{g}\Delta t_k^2 \tag{19}$$

$$^G\hat{\mathbf{v}}_{I_{k+1}} \simeq {}^G\hat{\mathbf{v}}_{I_k} + {}_{I_k}^G\hat{\mathbf{R}}\boldsymbol{\Xi}_1{}^{I_k}\hat{\mathbf{a}} - {}^G\mathbf{g}\Delta t_k \tag{20}$$

where $\Delta t_k = t_{k+1} - t_k$. We also have the following preintegration terms $\boldsymbol{\Xi}_1$ and $\boldsymbol{\Xi}_2$ defined as:

$$\boldsymbol{\Xi}_1 = \int_{t_k}^{t_{k+1}} {}_{I_\tau}^{I_k}\mathbf{R}d\tau, \ \boldsymbol{\Xi}_2 = \int_{t_k}^{t_{k+1}}\int_{t_k}^s {}_{I_\tau}^{I_k}\mathbf{R}d\tau ds \tag{21}$$

B. State Transition Matrix

We now define the linearized error state system as:

$$\delta\boldsymbol{\theta}_{k+1} \simeq {}_{I_k}^{I_{k+1}}\hat{\mathbf{R}}\delta\boldsymbol{\theta}_k + \mathbf{J}_r\left(t_k\right)\Delta t_k{}^{I_k}\tilde{\boldsymbol{\omega}} \tag{22}$$

$$^G\tilde{\mathbf{v}}_{I_{k+1}} \simeq {}^G\tilde{\mathbf{v}}_{I_k} - {}_{I_k}^G\hat{\mathbf{R}}\lfloor\boldsymbol{\Xi}_1\rfloor\delta\boldsymbol{\theta}_k + {}_{I_k}^G\hat{\mathbf{R}}\boldsymbol{\Xi}_1{}^{I_k}\tilde{\mathbf{a}} \tag{23}$$
$$- {}_{I_k}^G\hat{\mathbf{R}}\boldsymbol{\Xi}_3{}^{I_k}\tilde{\boldsymbol{\omega}}$$

$$^G\tilde{\mathbf{p}}_{I_{k+1}} \simeq {}^G\tilde{\mathbf{p}}_{I_k} + {}^G\tilde{\mathbf{v}}_k\Delta t_k - {}_{I_k}^G\hat{\mathbf{R}}\lfloor\boldsymbol{\Xi}_2\rfloor\delta\boldsymbol{\theta}_k \tag{24}$$
$$+ {}_{I_k}^G\hat{\mathbf{R}}\boldsymbol{\Xi}_2{}^{I_k}\tilde{\mathbf{a}} - {}_{I_k}^G\hat{\mathbf{R}}\boldsymbol{\Xi}_4{}^{I_k}\tilde{\boldsymbol{\omega}}$$

where $\mathbf{J}_r\left(t_k\right) \triangleq \mathbf{J}_r\left(^{I_k}\hat{\boldsymbol{\omega}}_k\Delta t_k\right)$ [4] denotes the $SO(3)$ right Jacobian. The $^{I_k}\tilde{\boldsymbol{\omega}}$, $^{I_k}\tilde{\mathbf{a}}$, $\mathbf{H}_{Dw}$ and $\mathbf{H}_{Da}$ terms are defined as:

$$^{I_k}\tilde{\boldsymbol{\omega}} = -\hat{\mathbf{D}}_w\tilde{\mathbf{b}}_g + \mathbf{H}_{Dw}\tilde{\mathbf{x}}_{Dw} - \hat{\mathbf{D}}_w\mathbf{n}_g \tag{25}$$

$$^{I_k}\tilde{\mathbf{a}} = \lfloor^{I_k}\hat{\mathbf{a}}\rfloor\delta\boldsymbol{\theta}_a + {}_a^I\hat{\mathbf{R}}\mathbf{H}_{Da}\tilde{\mathbf{x}}_{Da} - {}_a^I\hat{\mathbf{R}}\hat{\mathbf{D}}_a(\tilde{\mathbf{b}}_a + \mathbf{n}_a) \tag{26}$$

$$\mathbf{H}_{Dw} = \begin{bmatrix} ^w\hat{w}_1\mathbf{e}_1 & ^w\hat{w}_2\mathbf{e}_1 & ^w\hat{w}_2\mathbf{e}_2 & ^w\hat{w}_3\mathbf{I}_3 \end{bmatrix} \tag{27}$$

$$\mathbf{H}_{Da} = \begin{bmatrix} ^a\hat{a}_1\mathbf{e}_1 & ^a\hat{a}_2\mathbf{e}_1 & ^a\hat{a}_2\mathbf{e}_2 & ^a\hat{a}_3\mathbf{I}_3 \end{bmatrix} \tag{28}$$

The integrated components $\boldsymbol{\Xi}_3$ and $\boldsymbol{\Xi}_4$ are as follows:

$$\boldsymbol{\Xi}_3 = \int_{t_k}^{t_{k+1}} {}_{I_\tau}^{I_k}\mathbf{R}\lfloor^{I_\tau}\mathbf{a}\rfloor\mathbf{J}_r\left(\tau\right)\Delta\tau d\tau \tag{29}$$

$$\boldsymbol{\Xi}_4 = \int_{t_k}^{t_{k+1}}\int_{t_k}^s {}_{I_\tau}^{I_k}\mathbf{R}\lfloor^{I_\tau}\mathbf{a}\rfloor\mathbf{J}_r\left(\tau\right)\Delta\tau d\tau ds \tag{30}$$

where $\Delta\tau = \tau - t_k$ and the integration terms $\boldsymbol{\Xi}_i, i = 1\ldots 4$ can be computed analytically assuming constant measurements, see [28]. Once computed, they can be used repeatedly for state transitions matrices and Jacobians. Finally, we have the overall linearized error state system:

$$\tilde{\mathbf{x}}_{k+1} \simeq \boldsymbol{\Phi}_{(k+1,k)}\tilde{\mathbf{x}}_k + \mathbf{G}_k\mathbf{n}_{dk} \tag{31}$$

$$\boldsymbol{\Phi}_{(k+1,k)} = \begin{bmatrix} \boldsymbol{\Phi}_{I(k+1,k)} & \boldsymbol{\Phi}_{I,in} & \mathbf{0}_{15\times 3} \\ \mathbf{0}_{15\times 15} & \boldsymbol{\Phi}_{in} & \mathbf{0}_{15\times 3} \\ \mathbf{0}_{3\times 15} & \mathbf{0}_{3\times 15} & \boldsymbol{\Phi}_f \end{bmatrix} \tag{32}$$

where $\boldsymbol{\Phi}_{in} = \mathbf{I}_{15}$, $\boldsymbol{\Phi}_f = \mathbf{I}_3$, $\boldsymbol{\Phi}_{I(k+1,k)}$, $\boldsymbol{\Phi}_{I,in}$ and $\mathbf{G}_k$ can be found in supplementary materials [30]. , and $\mathbf{n}_{dk}$ is the discrete-time IMU noises.

IV. Observability Analysis

Observability analysis plays an important role in state estimation for VINS [10, 18]. This analysis allows for determining the minimum measurements needed to uniquely determine the state and identify degenerate motions which can possibly hurt system performance. To the best of our knowledge this is the first time that the observability properties and degenerate motions of IMU intrinsic calibration within the VINS framework has been investigated. Following the logic of [8], we can construct the observability matrix as:

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_1 \\ \mathbf{M}_2 \\ \vdots \\ \mathbf{M}_k \end{bmatrix} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \boldsymbol{\Phi}_{(2,1)} \\ \vdots \\ \mathbf{H}_k \boldsymbol{\Phi}_{(k,1)} \end{bmatrix} \tag{33}$$

Without loss of generality, the k-th row of $\mathbf{M}_k$ is:

$$\mathbf{M}_k = \mathbf{H}_{\pi}{}_I^C\hat{\mathbf{R}}{}_G^{I_k}\hat{\mathbf{R}}\times \tag{34}$$
$$\begin{bmatrix} \boldsymbol{\Gamma}_1 & \boldsymbol{\Gamma}_2 & \boldsymbol{\Gamma}_3 & \boldsymbol{\Gamma}_4 & \boldsymbol{\Gamma}_5 & \boldsymbol{\Gamma}_6 & \boldsymbol{\Gamma}_7 & \boldsymbol{\Gamma}_8 & \boldsymbol{\Gamma}_9 \end{bmatrix}$$

where:

$$\boldsymbol{\Gamma}_1 = \lfloor {}^G\hat{\mathbf{p}}_f - {}^G\hat{\mathbf{p}}_{I_1} - {}^G\hat{\mathbf{v}}_{I_1}\Delta t_k + \frac{1}{2}{}^G\mathbf{g}\Delta t_k^2 \rfloor {}_{I_1}^G\hat{\mathbf{R}}$$

$$\boldsymbol{\Gamma}_2 = -\mathbf{I}_3, \ \boldsymbol{\Gamma}_3 = -\mathbf{I}_3\Delta t_k, \ \boldsymbol{\Gamma}_5 = \mathbf{H}_{aa}^{pI}\hat{\mathbf{R}}\hat{\mathbf{D}}_a$$

$$\boldsymbol{\Gamma}_4 = -(\lfloor {}^G\hat{\mathbf{p}}_f - {}^G\hat{\mathbf{p}}_{I_k} \rfloor {}_{I_k}^G\hat{\mathbf{R}}\mathbf{J}_r(t_k)\Delta t_k + \mathbf{H}_w^p)\hat{\mathbf{D}}_w$$

$$\boldsymbol{\Gamma}_6 = (\lfloor {}^G\hat{\mathbf{p}}_f - {}^G\hat{\mathbf{p}}_{I_k} \rfloor {}_{I_k}^G\hat{\mathbf{R}}\mathbf{J}_r(t_k)\Delta t_k + \mathbf{H}_w^p)\mathbf{H}_{Dw}$$

$$\boldsymbol{\Gamma}_7 = -\mathbf{H}_{aa}^{pI}\hat{\mathbf{R}}\mathbf{H}_{Da}, \ \boldsymbol{\Gamma}_8 = -\mathbf{H}_a^p\lfloor {}^{I_k}\hat{\mathbf{a}} \rfloor, \ \boldsymbol{\Gamma}_9 = \mathbf{I}_3$$

where $\mathbf{H}_a^p = {}_{I_k}^G\hat{\mathbf{R}}\boldsymbol{\Xi}_2$ and $\mathbf{H}_w^p = {}_{I_k}^G\hat{\mathbf{R}}\boldsymbol{\Xi}_4$.

Given general random motion, one can easily find four unobservable directions, which are typical for VINS system [8]. However, $\boldsymbol{\Gamma}_6$, $\boldsymbol{\Gamma}_7$ and $\boldsymbol{\Gamma}_8$ are heavily motion affected and time-varying, complicating the analysis. Combined with our numerical simulations of a monocular camera and IMU, Section VI, intrinsic parameters are observable give fully-excited motions. While we omit the derivations and simulation results here due to space, the other three intrinsic model variants are also fully observable in the case of random motion.

V. Degenerate Motion Analysis

From the above analysis, it is clear that since $\boldsymbol{\Gamma}_6$, $\boldsymbol{\Gamma}_7$ and $\boldsymbol{\Gamma}_8$ are heavily motion affected, and thus, they are extremely susceptible to being unobservable under certain motions. Since bias terms and IMU intrinsics are tightly-coupled, through observation of the observability matrix $\mathbf{M}$, we provide a selection of basic motion types which can cause the IMU intrinsics to become unobservable with the given *imu2*. While we only provide the results for *imu2*, the results are applicable to the other three.

1) ${}^w w_1$ constant: If ${}^w w_1$ is constant, d_{a11} will be unobservable with unobservable directions as:

$$\mathbf{N}_{w1} = \begin{bmatrix} \mathbf{0}_{1\times 9} & (\hat{\mathbf{D}}_w^{-1}\mathbf{e}_1)^{\top}{}^w w_1 & \mathbf{0}_{1\times 3} & 1 & \mathbf{0}_{1\times 17} \end{bmatrix}^{\top} \tag{35}$$

TABLE I: Summary of basic degenerate motions for online IMU intrinsics calibration (*imu2*).

Motion Types	Nullspace Dim.	Unobservable Parameters
constant ${}^w\omega_1$	1	d_{w11}
constant ${}^w\omega_2$	2	d_{w12}, d_{w22}
constant ${}^w\omega_3$	3	$d_{w13}, d_{w23}, d_{w33}$
constant ${}^a a_1$	3	d_{a11}, pitch and yaw of ${}_a^I\mathbf{R}$
constant ${}^a a_2$	3	d_{a12}, d_{a22}, roll of ${}_a^I\mathbf{R}$
constant ${}^a a_3$	3	$d_{a13}, d_{a23}, d_{a33}$

2) ${}^w w_2$ constant: If ${}^w w_2$ is constant, d_{a12} and d_{a22} will be unobservable with unobservable directions as:

$$\mathbf{N}_{w2} = \begin{bmatrix} \mathbf{0}_{1\times 9} & (\hat{\mathbf{D}}_w^{-1}\mathbf{e}_1)^{\top}{}^w w_2 & \mathbf{0}_{1\times 4} & 1 & \mathbf{0}_{1\times 16} \\ \mathbf{0}_{1\times 9} & (\hat{\mathbf{D}}_w^{-1}\mathbf{e}_2)^{\top}{}^w w_2 & \mathbf{0}_{1\times 5} & 1 & \mathbf{0}_{1\times 15} \end{bmatrix}^{\top} \tag{36}$$

3) ${}^w w_3$ constant: If ${}^w w_3$ is constant, d_{a13}, d_{a23} and d_{a33} are unobservable with unobservable directions as:

$$\mathbf{N}_{w3} = \begin{bmatrix} \mathbf{0}_{1\times 9} & (\hat{\mathbf{D}}_w^{-1}\mathbf{e}_1)^{\top}{}^w w_3 & \mathbf{0}_{1\times 6} & 1 & \mathbf{0}_{1\times 14} \\ \mathbf{0}_{1\times 9} & (\hat{\mathbf{D}}_w^{-1}\mathbf{e}_2)^{\top}{}^w w_3 & \mathbf{0}_{1\times 7} & 1 & \mathbf{0}_{1\times 13} \\ \mathbf{0}_{1\times 9} & (\hat{\mathbf{D}}_w^{-1}\mathbf{e}_3)^{\top}{}^w w_3 & \mathbf{0}_{1\times 8} & 1 & \mathbf{0}_{1\times 12} \end{bmatrix}^{\top} \tag{37}$$

4) ${}^a a_1$ constant: If ${}^a a_1$ is constant, d_{a11}, pitch and yaw of ${}_a^I\mathbf{R}$ are unobservable with unobservable directions as:

$$\mathbf{N}_{a1} = \begin{bmatrix} \mathbf{0}_{12\times 1} & \mathbf{0}_{12\times 1} & \mathbf{0}_{12\times 1} \\ \hat{\mathbf{D}}_a^{-1}\mathbf{e}_1{}^a a_1 & \hat{\mathbf{D}}_a^{-1}\mathbf{e}_2\hat{d}_{a11}{}^a a_1 & \hat{\mathbf{D}}_a^{-1}\mathbf{e}_3\hat{d}_{a11}\hat{d}_{a22}{}^a a_1 \\ \mathbf{0}_{6\times 1} & \mathbf{0}_{6\times 1} & \mathbf{0}_{6\times 1} \\ 1 & 0 & 0 \\ 0 & \hat{d}_{a22} & 0 \\ 0 & -\hat{d}_{a12} & 0 \\ 0 & \hat{d}_{a23} & -\hat{d}_{a33}\hat{d}_{a22} \\ 0 & -\hat{d}_{a13} & \hat{d}_{a12}\hat{d}_{a33} \\ 0 & 0 & \hat{d}_{a13}\hat{d}_{a22} - \hat{d}_{a12}\hat{d}_{a23} \\ \mathbf{0}_{3\times 1} & {}_a^I\hat{\mathbf{R}}\mathbf{e}_3 & {}_a^I\hat{\mathbf{R}}(\mathbf{e}_1\hat{d}_{a12} + \mathbf{e}_2\hat{d}_{a22}) \\ \mathbf{0}_{3\times 1} & \mathbf{0}_{3\times 1} & \mathbf{0}_{3\times 1} \end{bmatrix}$$

5) ${}^a a_2$ constant: If ${}^a a_2$ is constant, d_{a12}, d_{a22} and roll of ${}_a^I\mathbf{R}$ are unobservable with unobservable directions as:

$$\mathbf{N}_{a2} = \begin{bmatrix} \mathbf{0}_{12\times 1} & \mathbf{0}_{12\times 1} & \mathbf{0}_{12\times 1} \\ \hat{\mathbf{D}}_a^{-1}\mathbf{e}_1{}^a a_2 & \hat{\mathbf{D}}_a^{-1}\mathbf{e}_2{}^a a_2 & \hat{\mathbf{D}}_a^{-1}\mathbf{e}_3\hat{d}_{a22}{}^a a_2 \\ \mathbf{0}_{6\times 1} & \mathbf{0}_{6\times 1} & \mathbf{0}_{6\times 1} \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \hat{d}_{a33} \\ 0 & 0 & -\hat{d}_{a23} \\ \mathbf{0}_{3\times 1} & \mathbf{0}_{3\times 1} & {}_a^I\hat{\mathbf{R}}\mathbf{e}_1 \\ \mathbf{0}_{3\times 1} & \mathbf{0}_{3\times 1} & \mathbf{0}_{3\times 1} \end{bmatrix} \tag{38}$$

6) ${}^a a_3$ constant: If ${}^a a_3$ is constant, d_{a13}, d_{a23} and d_{a33} are unobservable with unobservable directions as:

$$\mathbf{N}_{a3} = \begin{bmatrix} \mathbf{0}_{1\times 12} & (\hat{\mathbf{D}}_a^{-1}\mathbf{e}_1)^{\top}{}^a a_3 & \mathbf{0}_{1\times 9} & 1 & \mathbf{0}_{1\times 8} \\ \mathbf{0}_{1\times 12} & (\hat{\mathbf{D}}_a^{-1}\mathbf{e}_2)^{\top}{}^a a_3 & \mathbf{0}_{1\times 10} & 1 & \mathbf{0}_{1\times 7} \\ \mathbf{0}_{1\times 12} & (\hat{\mathbf{D}}_a^{-1}\mathbf{e}_3)^{\top}{}^a a_3 & \mathbf{0}_{1\times 11} & 1 & \mathbf{0}_{1\times 6} \end{bmatrix}^{\top} \tag{39}$$

From the analysis, the IMU intrinsic calibration is sensitive to sensor motion and thus all 6 axes need to be excited to make sure all IMU intrinsics can be calibrated. The findings are summarized in Table I. Note that any combination of these basic motions is still degenerate

TABLE II: Simulation parameters and prior standard deviations that perturbations of measurements and initial states were drawn from.

Parameter	Value	Parameter	Value
IMU Scale	0.005	IMU Skew	0.002
Rot. atoI (rad)	0.001	Rot. wtoI (rad)	0.001
Gyro. White Noise	1.6968e-04	Gyro. Rand. Walk	1.9393e-05
Accel. White Noise	2.0000e-3	Accel. Rand. Walk	3.0000e-3
Focal Len. (px/m)	0.35	Cam. Center (px)	0.45
d1 and d2	0.01	d3 and d4	0.001
Rot. CtoI (rad)	0.001	Pos. IinC (m)	0.01
Pixel Proj. (px)	1	Timeoff (s)	5e-4
Cam Freq. (hz)	20	IMU Freq. (hz)	200
Avg. Feats	50	Num. SLAM	25
Num. Clones	11	Feat. Rep.	GLOBAL
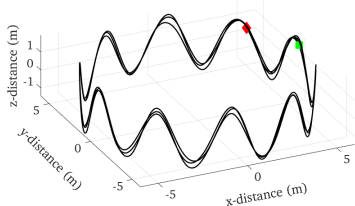
Fig. 2: The 216 meter long sinusoidal simulation trajectory. Starting location denoted with the green square and ending location denoted with the red diamond.

and causes all related parameters to be unobservable. The degenerate motion analysis can also be extended to other 3 model variants, which we omit here for brevity.

VI. SIMULATION VALIDATIONS

A. Simulation with Fully-Excited Motion

We leverage the OpenVINS [6] framework to verify our observability analysis under different conditions. The basic configurations for our simulator are listed in Table II. In the following general trajectory simulation, we perform full calibration (including IMU-camera extrinsics, time offset, both camera and *imu2* IMU intrinsics) for VINS with a single monocular camera. We refer the reader to the OpenVINS [6] for details on how this changes the state.

The trajectory, shown in Fig. 2, is a general 3D sinusoid with full excitation of all 6 axes. From the results shown in Fig. 3, the estimation errors and 3σ bounds for all the intrinsic parameters of *imu2* can converge quite nicely, verifying that the analysis for general motion holds true. We plot six different realizations of the initial calibration guesses based on the specified priors, and it is clear that all are able to converge from different initial guesses to a small error.

B. Comparison of Different Inertial Model Variants

We next compare these four proposed model variants. The sinusoidal trajectory, see Fig. 2, is used but online camera intrinsics, extrinsics, and IMU-CAM time offset calibration is disabled. Shown in Table III, it is clear that the choice between these four variants has little impact on estimation accuracy which indicates they provide almost the same amount of correction to the inertial

TABLE III: Average ATE and NEES over 10 runs with true or bad calibration, with and without online IMU intrinsic calibration. This used the sinusoidal trajectory with 6-axis excitation and only calibrated the IMU intrinsics. The notation "true" means we use the groundtruth calibration, while "bad" means we use the perturbed calibration states. Note that we only report a single configuration when we use the true parameters and don't calibrate.

IMU Model	ATE (deg)	ATE (m)	Ori. NEES	Pos. NEES
true w/ calib imu1	0.182	0.060	1.969	0.953
true w/ calib imu2	0.184	0.063	1.978	1.046
true w/ calib imu3	0.183	0.060	1.967	0.948
true w/ calib imu4	0.184	0.063	1.977	1.045
bad w/ calib imu1	0.182	0.061	1.977	0.984
bad w/ calib imu2	0.179	0.063	1.978	1.091
bad w/ calib imu3	0.181	0.060	1.974	0.976
bad w/ calib imu4	0.178	0.063	1.979	1.093
true w/o calib vio	0.173	0.059	1.968	0.981
bad w/o calib imu1	0.259	0.213	9.655	15.457
bad w/o calib imu2	0.260	0.190	9.181	12.310
bad w/o calib imu3	0.275	0.214	9.446	15.601
bad w/o calib imu4	0.269	0.200	9.337	14.248

readings. Note that we perturb each model based on which parameters it estimates and use the priors specified in Table II.

Also seen is the accuracy of the standard VIO system which does not calibrate any parameters online and uses the groundtruth calibration values. As expected this has the best accuracy due to the use of the true parameters. The VIO system which does *not* estimate the IMU intrinsics but instead treats the perturbed states as being true is inconsistent, as seen by the large Normalized Estimation Error Squared (NEES), and has the largest estimation error. As compared to online IMU intrinsic calibration, there is only a small loss in estimation accuracy which is far smaller in magnitude than if we were to treat these bad calibration values as being true.

C. Simulation with Degenerate Motion

We now verify the identified degenerate motions and present results for two special motions: 1-axis rotation (see Fig. 4) and constant local a_x (see Fig. 5). In both simulations, we only calibrate the IMU intrinsics and omit the results of full system calibration as they give the same conclusion. We modify the same sinusoidal trajectory, see Fig. 2, by removing roll and pitch orientation changes for the first degenerate motion (with only yaw rotation) and removed the pitch and made the current yaw angle tangent to the trajectory in the x-y plane for the second degenerate motion (with constant local acceleration along x-axis).

Shown in Fig. 4, the first 3 parameters d_{w11}, d_{w12} and d_{w22} for $\mathbf{D}_w$ did not converge at all (the 3σ bounds are almost straight lines), which matches our analysis, see Table I, which says these should be unobservable in the case that $^w w_x$ (roll) and $^w w_y$ (pitch) are constant as in this 1-axis case. Fig. 5 shows where we have enforced that the local acceleration along the x-axis, a_x, is constant. The d_{a11} and the pitch and yaw of $^I_a\mathbf{R}$ terms do not converge, thus validating our analysis.
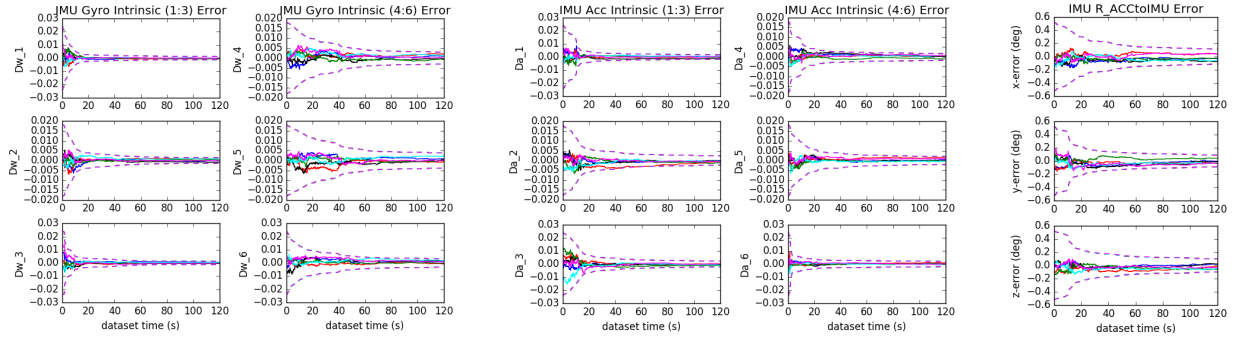
Fig. 3: 3 sigma bounds and estimation errors for six different runs with different realization of the measurement noise for fully excited motion (using *imu2*). All the IMU intrinsic parameters converge nicely. Note that all parameters, including camera intrinsics, extrinsics, time offset, and IMU intrinsic parameters are calibrated.
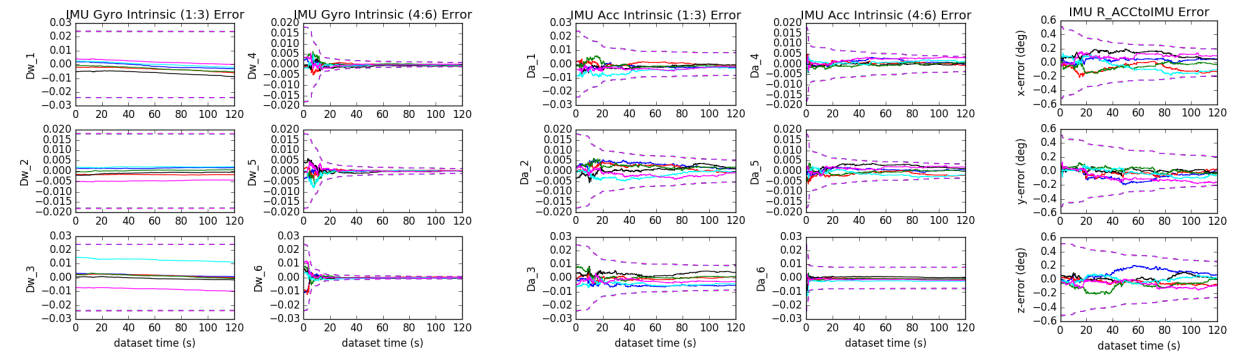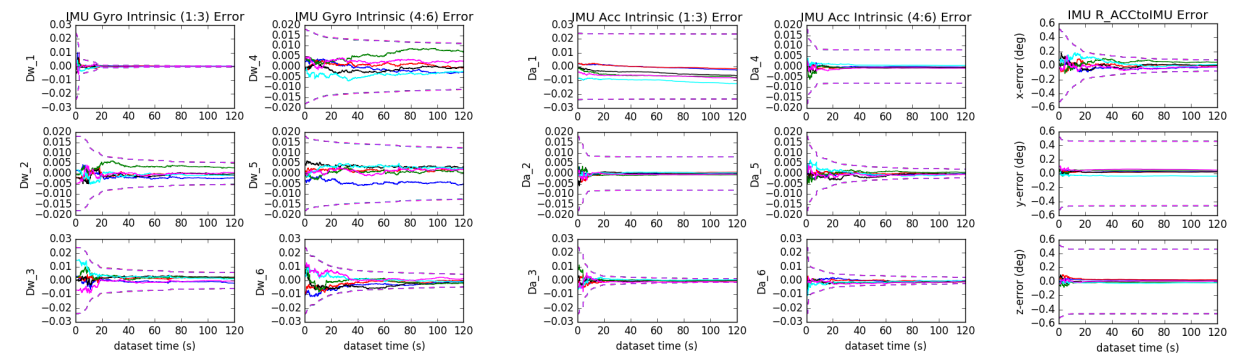

Fig. 4: 3 sigma bounds and estimation errors for six different runs with different realization of the measurement noise and initial calibration states for 1-axis rotation trajectory (using *imu2*). d_{w11}, d_{w12}, and d_{w22} (the first 3 parameters from $\mathbf{D}_w$) are not observable and they did not converge. Note that only the IMU intrinsic parameters are calibrated.

Fig. 5: 3 sigma bounds and estimation errors for six different runs with different realization of the measurement noise and initial calibration states for degenerate motion with constant acceleration on x-axis (using *imu2*). d_{a11} and the pitch and yaw of $_a^I\mathbf{R}$ are not observable and they did not converge. Note that only the IMU intrinsic parameters are calibrated.

D. Simulation with Over Parameterization

As we mentioned in Section II-A, if we calibrate both 9 parameters for gyroscope and accelerometer, the rotation between IMU and the camera will be affected. If we calibrate all three relative rotations, the intermediate inertial frame $\{I\}$ is not constrained. If we change the relative rotation from ItoC, then this perturbed rotation can be absorbed into the atoI and wtoI terms, and thus means we have an extra 3DoF rotation not constrained by our measurement. As shown in Fig. 6, when using the

full 6-axis excitation sinusoidal trajectory, see Section VI-A, the convergence of $_I^C\mathbf{R}$ becomes much worse if we calibrate IMU-CAM extrinsics and all 18 parameters for the IMU intrinsics even when the *same* priors and measurements are used.

VII. REAL-WORLD EXPERIMENTS

We further evaluate IMU intrinsic calibration on 3 real-world datasets with different motion/actuation profiles: TUM VI datasets with general handheld motion,
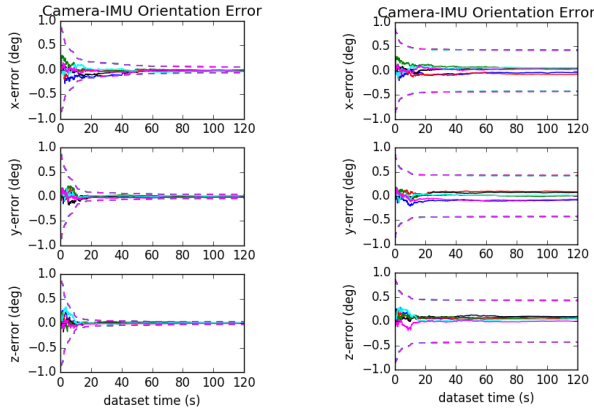
Fig. 6: Camera to IMU orientation errors when using IMU *imu2* (left) and the over paramterized *imu5* (right). Note that *imu5* uses $\mathbf{D}_{w9}$ and $\mathbf{D}_{a9}$. Only the IMU intrinsics and relative pose between IMU and camera were online calibrated.

EuRoc MAV datasets with underactuated 3D motion and KAIST dataset with planar motion.

A. TUM-VI Dataset: General Handheld Motion

We first evaluate on the handheld real-world TUM VI benchmark [24] which provides grayscale stereo images at 20 Hz, a time-synchronized 200Hz IMU, and accurate pose groundtruth from an external motion capture system. We initialized $\mathbf{D}_a$, $\mathbf{D}_w$, $_a^I\mathbf{R}$, $_w^I\mathbf{R}$ as identity matrices since these are unknown values, with inflated priors to account that these are unknown on startup. Only the IMU intrinsics are calibrated while the rest of the parameters are set to their values provided by the dataset. While the filter employs First-Estimates Jacobains (FEJ) [12], in practice we do not FEJ the intrinsic calibrations to prevent large linearization errors. We track a maximum of 200 features with a max of 30 SLAM features being kept in the state vectors and only use the left camera image as our input monocular camera feed.

As shown in Table IV, we evaluate the Absolute Trajectory Error (ATE) [33] for the proposed IMU intrinsic model variants and standard VIO without online IMU intrinsic calibration. We can see an improvement due to the estimation of the IMU intrinsics, while it is important to note that, as shown in the simulation, there is little difference between the four variants. This shows that if the trajectory has full excitation, as in the case for handheld trajectories, the estimator accuracy improves due to online IMU intrinsic calibration.

B. EuRoC MAV Dataset: Under-actuated Motion

The EuRoC MAV dataset [1] contains a series of trajectories from a MAV with a visual-inertial sensor attached and provides 20 Hz grayscale stereo images and 200 Hz inertial readings in addition to an external groundtruth motion capture system. We use the same estimator as in the TUM VI datasets (see Sec. VII-A)

and report the ATE results in Table V, which clearly differ from the results of the handheld datasets. This is primarily due to the fact that an underactuated MAV cannot fully excite its 6DoF motion for a given small time interval and thus undergoes (nearly) degenerate motions locally (for *imu2*), hurting the sliding-window filter (i.e., MSCKF). To see this, we have computed the sample standard deviations of IMU readings for both MAV and TUM datasets and found that the MAV datasets demonstrate significantly smaller $\boldsymbol{\omega}$ and $\mathbf{a}$ changes in local time windows than the TUM datasets (see supplementary materials [30] for details).

C. KAIST Complex Urban Dataset: Planar Motion

We now evaluate the performance of online intrinsic calibration on the *very* common motion profile of planar motion. As in the Table I, planar motion is the combination of three degenerate motions and in the case of *imu2* 6 parameters are unobservable. We leverage the KAIST Complex Urban dataset [13], which provides stereo grayscale images at 10 Hz and a 200 Hz IMU. They additionally provide a batch-optimized groundtruth which can be used as a reference to compare against. As compared to the previous experiments, we use the stereo pair as our camera along with the IMU to remove the scale ambiguity for monocular VINS caused by constant acceleration [27, 31] and also enabled all online calibration to handle inaccuracies in the dataset's provided transforms and measurement times.

Shown in Fig. 7, the system with IMU intrinsic calibration has larger drift on the Urban 39 dataset. When looking at the Root Mean Squared Error (RMSE) in respect to the dataset's groundtruth, for the standard VIO is 1.58 degrees with 13.03 meters (0.12%), while with *imu2* online IMU intrinsic calibration it is 1.41 degrees with 23.13 meters (0.22%). We propose that this larger error is due to the introduced unobservable directions in online IMU intrinsic calibration when experiencing planar motion.

VIII. Discussion: Necessity of Online IMU Calibration?

Recall the question raised in the introduction was *whether or not we should calibrate IMU intrinsics online, and if this is not the case, under what conditions does online calibration fail.* As shown in our degenerate motion analysis, there are a large number of motion types that prohibit accurate calibration of the IMU intrinsics. In the general case with random motion such as in the TUM VI datasets, we are able to improve accuracy by online estimating these parameters since they are all observable.

More importantly, in the motion cases most commonly seen in aerial and ground vehicles, there is typically at least one unobservable direction due to these robots traveling with either underactuated 3D motion or planar motion. The impact on performance was shown in the

TABLE IV: Absolute Trajectory Error (ATE) on TUM VI room equences (with units degrees/meters). *imu1* denotes $\mathbf{D}_{a6}, \mathbf{D}_{\omega6}, {}^I_w\mathbf{R}$; *imu2* denotes $\mathbf{D}_{a6}, \mathbf{D}_{\omega6}, {}^I_a\mathbf{R}$; *imu3* denotes $\mathbf{D}_{a6}, \mathbf{D}_{\omega9}$; *imu4* denotes $\mathbf{D}_{a9}, \mathbf{D}_{\omega6}$.

IMU Model	dataset-room1	dataset-room2	dataset-room3	dataset-room4	dataset-room5	dataset-room6	Average
VIO	1.430 / 0.089	1.173 / 0.064	1.934 / 0.088	1.333 / 0.054	1.140 / 0.092	0.888 / 0.056	1.317 / 0.074
imu1	0.954 / 0.066	1.153 / 0.059	1.809 / 0.074	1.175 / 0.038	1.028 / 0.073	1.017 / 0.033	1.189 / 0.057
imu2	0.877 / 0.077	1.170 / 0.051	1.974 / 0.076	1.148 / 0.039	0.950 / 0.081	0.825 / 0.038	1.157 / 0.060
imu3	0.957 / 0.065	1.142 / 0.058	1.836 / 0.075	1.211 / 0.039	1.006 / 0.073	1.021 / 0.034	1.196 / 0.057
imu4	0.893 / 0.077	1.173 / 0.052	1.896 / 0.076	1.134 / 0.038	1.259 / 0.097	0.823 / 0.038	1.196 / 0.063

TABLE V: Absolute Trajectory Error (ATE) on EuRoC MAV Vicon room sequences (with units degrees/meters). *imu1* denotes $\mathbf{D}_{a6}, \mathbf{D}_{\omega6}, {}^I_w\mathbf{R}$; *imu2* denotes $\mathbf{D}_{a6}, \mathbf{D}_{\omega6}, {}^I_a\mathbf{R}$; *imu3* denotes $\mathbf{D}_{a6}, \mathbf{D}_{\omega9}$; *imu4* denotes $\mathbf{D}_{a9}, \mathbf{D}_{\omega6}$.

IMU Model	V1_01_easy	V1_02_medium	V1_03_difficult	V2_01_easy	V2_02_medium	V2_03_difficult	Average
VIO	0.657 / 0.043	1.805 / 0.060	2.437 / 0.069	0.869 / 0.109	1.373 / 0.080	1.277 / 0.180	1.403 / 0.090
imu1	0.601 / 0.055	1.924 / 0.065	2.334 / 0.073	1.201 / 0.115	1.342 / 0.086	1.710 / 0.168	1.519 / 0.094
imu2	0.552 / 0.054	1.990 / 0.062	2.197 / 0.083	0.960 / 0.107	1.453 / 0.085	1.666 / 0.216	1.470 / 0.101
imu3	0.606 / 0.055	1.905 / 0.065	2.359 / 0.073	1.180 / 0.114	1.335 / 0.088	1.640 / 0.167	1.504 / 0.094
imu4	0.569 / 0.056	1.969 / 0.069	2.165 / 0.076	0.846 / 0.127	1.636 / 0.094	1.577 / 0.195	1.461 / 0.103

EuRoC MAV and KAIST Urban datasets where the use of online IMU intrinsic calibration hurts estimator accuracy! This leads to our recommendation and answer to the question: due to the high likelihood of (or even almost sure) experiencing degenerate motions for some short of period of time in most robotic applications, we do *not* recommend performing online IMU intrinsic calibration during real-time operations. The exception to this is the handheld sensor or mobile device case, which often exhibit full 6DoF motions and thus is recommended to perform online IMU calibration to improve estimation accuracy. For both of these applications, we still recommend using an offline batch optimization to calibrate the IMU intrinsics as an initial guess for the filter or to treat the intrinsics as true if one knows they are going to experience degenerate motions.

IX. CONCLUSIONS AND FUTURE WORK

In this paper we have addressed the problem of online IMU calibration for visual-inertial navigation, and presented four IMU intrinsic model variants derived from one commonly used inertial model in practice. Based on the observability analysis, we have identified six basic degenerate motion patterns, of which, any combination results in unobservable directions of IMU intrinsic parameters. Extensive validation on simulated and real-world datasets were performed to verify both the observability and degenerate motion analysis. As shown through our experiments, online IMU intrinsic calibration is risky due to its dependence on the motion profile to ensure observability. In the case of autonomous (ground) vehicles, most trajectories would have degenerate motion, thus resulting in us *not* recommending online calibration of IMU intrinsics for these types of robots. In the case of handheld motion, however we found that the estimation of IMU intrinsics improved performance
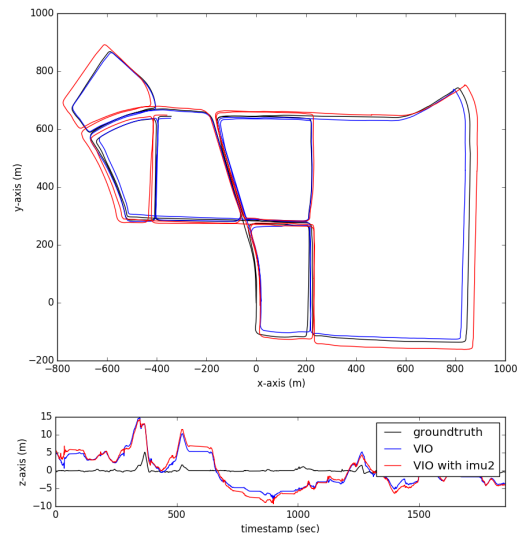

Fig. 7: Trajectory plots for the KAIST Urban 39 dataset 10km in total length. Figure is best seen in color.

as expected. In the future, we will investigate a complete degenerate motion analysis and different ways to selectively perform online calibration to avoid additional unobservable direction caused in degenerate motions.

X. ACKNOWLEDGMENTS

REFERENCES

[1] M. Burri, J. Nikolic, P. Gohl, T. Schneider, J. Rehder, S. Omari, M. W. Achtelik, and R. Siegwart. The euroc micro aerial vehicle datasets. *The International Journal of Robotics Research*, 35(10):1157–1163, 2016.

[2] K. Eckenhoff, P. Geneva, J. Bloecker, and G. Huang. Multi-camera visual-inertial navigation with online intrinsic and extrinsic calibration. In *Proc. International Conference on Robotics and Automation*, Montreal, Canada, May 2019.

[3] K. Eckenhoff, P. Geneva, and G. Huang. Closed-form preintegration methods for graph-based visual-inertial navigation. *International Journal of Robotics Research*, 38(5):563–586, 2019.

[4] C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza. On-manifold preintegration for real-time visual–inertial odometry. *IEEE Transactions on Robotics*, 33(1):1–21, Feb. 2017.

[5] P. Furgale, J. Rehder, and R. Siegwart. Unified temporal and spatial calibration for multi-sensor systems. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1280–1286, Tokyo, Japan, Nov. 3–8, 2013.

[6] P. Geneva, K. Eckenhoff, W. Lee, Y. Yang, and G. Huang. Openvins: A research platform for visual-inertial estimation. In *IROS 2019 Workshop on Visual-Inertial Navigation: Challenges and Applications*, Macau, China, Nov. 2019.

[7] L. Heng, B. Choi, Z. Cui, M. Geppert, S. Hu, B. Kuan, P. Liu, R. Nguyen, Y. C. Yeo, A. Geiger, et al. Project autovision: Localization and 3d scene perception for an autonomous vehicle with a multi-camera system. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 4695–4702. IEEE, 2019.

[8] J. A. Hesch, D. G. Kottas, S. L. Bowman, and S. I. Roumeliotis. Consistency analysis and improvement of vision-aided inertial navigation. *IEEE Transactions on Robotics*, 30(1):158–176, Feb 2014.

[9] Z. Huai and G. Huang. Robocentric visual–inertial odometry. *The International Journal of Robotics Research*, 0(0):0278364919853361, 0.

[10] G. Huang. *Improving the Consistency of Nonlinear Estimators: Analysis, Algorithms, and Applications*. PhD thesis, Department of Computer Science and Engineering, University of Minnesota, 2012.

[11] G. Huang. Visual-inertial navigation: A concise review. In *Proc. International Conference on Robotics and Automation*, Montreal, Canada, May 2019.

[12] G. Huang, A. I. Mourikis, and S. I. Roumeliotis. Observability-based rules for designing consistent EKF SLAM estimators. *International Journal of Robotics Research*, 29(5):502–528, Apr. 2010.

[13] J. Jeong, Y. Cho, Y.-S. Shin, H. Roh, and A. Kim. Complex urban dataset with multi-level sensors from highly diverse urban environments. *The International Journal of Robotics Research*, 38(6):642–657, 2019.

[14] S. Leutenegger, S. Lynen, M. Bosse, R. Siegwart, and P. Furgale. Keyframe-based visual-inertial odometry using nonlinear optimization. *International Journal of Robotics Research*, Dec. 2014.

[15] M. Li and A. I. Mourikis. Online temporal calibration for Camera-IMU systems: Theory and algorithms. *International Journal of Robotics Research*, 33(7):947–964, June 2014.

[16] M. Li and A. I. Mourikis. Vision-aided inertial navigation with rolling-shutter cameras. *International Journal of Robotics Research*, 33(11):1490–1507, Sept. 2014.

[17] M. Li, H. Yu, X. Zheng, and A. I. Mourikis. High-fidelity sensor modeling and self-calibration in vision-aided inertial navigation. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 409–416, May 2014.

[18] A. Martinelli. Observability properties and deterministic algorithms in visual-inertial structure from motion. *Found. Trends Robot*, 3(3):139–209, Dec. 2014.

[19] F. M. Mirzaei and S. I. Roumeliotis. A kalman filter-based algorithm for IMU-camera calibration: Observability analysis and performance evaluation. *IEEE Transactions on Robotics*, 24(5):1143–1156, Oct. 2008.

[20] A. I. Mourikis and S. I. Roumeliotis. A multi-state constraint Kalman filter for vision-aided inertial navigation. In *International Conference on Robotics and Automation*, pages 3565–3572, Rome, Italy, Apr. 2007.

[21] T. Qin, P. Li, and S. Shen. Vins-mono: A robust and versatile monocular visual-inertial state estimator. *IEEE Transactions on Robotics*, 34(4):1004–1020, Aug 2018.

[22] J. Rehder, J. Nikolic, T. Schneider, T. Hinzmann, and R. Siegwart. Extending kalibr: Calibrating the extrinsics of multiple imus and of individual axes. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4304–4311, May 2016.

[23] T. Schneider, M. Li, C. Cadena, J. Nieto, and R. Siegwart. Observability-aware self-calibration of visual and inertial sensors for ego-motion estimation. *IEEE Sensors Journal*, 19(10):3846–3860, May 2019.

[24] D. Schubert, T. Goll, N. Demmel, V. Usenko, J. Stückler, and D. Cremers. The tum vi benchmark for evaluating visual-inertial odometry. In *2018*

IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 1680–1687. IEEE, 2018.

[25] S. Shen, Y. Mulgaonkar, N. Michael, and V. Kumar. Multi-sensor fusion for robust autonomous flight in indoor and outdoor environments with a rotorcraft MAV. In *Proc. of the IEEE International Conference on Robotics and Automation*, pages 4974–4981, Hong Kong, China, May 31 – June 7, 2014.

[26] N. Trawny and S. I. Roumeliotis. Indirect Kalman filter for 3D attitude estimation. Technical report, University of Minnesota, Dept. of Comp. Sci. & Eng., Mar. 2005.

[27] K. J. Wu, C. X. Guo, G. Georgiou, and S. I. Roumeliotis. Vins on wheels. In *International Conference on Robotics and Automation*, pages 5155–5162, Singapore, May. 2017. IEEE.

[28] Y. Yang, B. P. W. Babu, C. Chen, G. Huang, and L. Ren. Analytic combined imu integrator for visual-inertial navigation. In *Proc. of the IEEE International Conference on Robotics and Automation*, Paris, France, 2020.

[29] Y. Yang, P. Geneva, K. Eckenhoff, and G. Huang. Degenerate motion analysis for aided INS with online spatial and temporal calibration. *IEEE Robotics and Automation Letters (RA-L)*, 4(2):2070–2077, 2019.

[30] Y. Yang, P. Geneva, X. Zuo, and G. Huang. Supplementary materials: Online imu intrinsic calibration: Is it necessary? In *Technical Report of Robot Perception and Navigation Group*, July 2020. Available: http://udel.edu/~yuyang/downloads/tr_intrinsic.pdf.

[31] Y. Yang and G. Huang. Observability analysis of aided ins with heterogeneous features of points, lines and planes. *IEEE Transactions on Robotics*, 35(6):399–1418, Dec. 2019.

[32] M. Zhang, X. Zuo, Y. Chen, and M. Li. Localization for ground robots: On manifold representation, integration, re-parameterization, and optimization. *arXiv preprint arXiv:1909.03423*, 2019.

[33] Z. Zhang and D. Scaramuzza. A tutorial on quantitative trajectory evaluation for visual (-inertial) odometry. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 7244–7251. IEEE, 2018.