

A Direct Algorithm for Multi-Gyroscope Infield Calibration

Tianheng Wang and Stergios I. Roumeliotis

Abstract—In this paper, we address the problem of estimating the rotational extrinsics, as well as the scale factors of two gyroscopes rigidly mounted on the same device. In particular, we formulate the problem as a least-squares minimization and introduce a *direct* algorithm that computes the estimated quantities without any iterations, hence avoiding local minima and improving efficiency. Furthermore, we show that the rotational extrinsics are observable while the scale factors can be determined up to a global scale for general configurations of the gyroscopes. To this end, we also study special placements of the gyroscopes where a pair, or all, of their axes are parallel and analyze their impact on the scale factors’ observability. Lastly, we evaluate our algorithm in simulation and real-world experiments to assess its performance as a function of key motion and sensor characteristics.

I. INTRODUCTION AND RELATED WORK

Most mobile and portable devices (e.g., robots, cell phones, wearables) are equipped with an inertial measurement unit (IMU) for estimating its motion, often in conjunction with additional sensors, such as cameras, laser scanners, ultrawideband, or ultrasound sensors [1]–[13]. An IMU typically comprises a 3-axial accelerometer and a 3-axial gyroscope each measuring linear accelerations and rotational velocities, respectively, expressed in its own frame of reference. Using more than one IMU has the potential to increase motion estimation accuracy by fusing their measurements after accounting for their *extrinsics* [14]–[16].

Multi-IMU extrinsics can be determined in a number of ways.¹ In certain cases, IMUs are calibrated for both their extrinsics and intrinsics by employing specialized equipment such as a rate table [20], [21], in conjunction with a camera and using a calibration target [22], or an external optical system [23]. Although such approaches can provide excellent calibration accuracy, they have high cost, due to their dependence on hardware infrastructure. Additionally, their results are valid potentially for *limited time*, as both extrinsics and intrinsics may change with time and/or device use due to, e.g., humidity, temperature, and mechanical stress. For these reasons, recent research has focused on *infield* calibration of IMUs. When additional sensors, such as a camera, are available, the IMUs extrinsics can be estimated along with the device’s pose as part of a visual-inertial algorithm [14]. Such approaches, however, have a number of limitations. Firstly, the processing cost can be quite high due to the dependence on the aiding camera(s) and the requirement to extract, match, and process visual features within the visual-inertial estimator. Secondly, these methods can only

be used when additional sensors are present. Moreover, even when the aiding sensor, such as a camera, is available, the operating conditions (e.g., lighting, distance to the scene, texture, motion) may be unfavorable and negatively impact the calibration quality. Lastly, a joint localization, mapping, and calibration problem is typically high dimensional, non-linear and thus there are no guarantees for convergence to the global minimum,

Besides the potential of multiple IMUs to increase the accuracy of visual-inertial state estimation, there exist other use cases, where the key parameters to be estimated and monitored are the *rotational extrinsics between the gyroscopes*. Specifically, gyroscopes can be attached on the back of the binocular displays of a head-mounted device, or on ultrasound sensors installed on an underwater robot. In these applications, tracking the extrinsics of the gyroscopes for changes due to, e.g., temperature or shock, allows us to compensate for binocular disparity in the displays [24], [25], or account for changes in the viewing directions of exteroceptive sensors with non-overlapping fields of view [9]. Achieving this without using any additional sensors besides IMUs is highly desirable especially when considering operating in challenging conditions or under processing constraints.

Alternative approaches that directly calibrate the extrinsics of pairs of IMUs include the works of [26] and [27] where rotation between the gyroscopes and translation of the accelerometers are estimated by assuming *known and constant* intrinsics (scale factors and skewness). This assumption, however, is often violated in practice especially for the *scale factors* of gyroscopes found in MEMS IMUs, which drift due to temperature, humidity, shock, and mechanical strain.

To address the limitations of previous approaches, in this paper we make the following main contributions:

- 1) We formulate the problem of estimating the rotational extrinsics and intrinsics of a pair of gyroscopes as least-squares (LS) and introduce an algorithm for *directly* computing the observable subset of them.
- 2) We show that the extrinsics and scale factors, up to a global scale, are, in general, observable. Additionally, we consider degenerate cases of gyroscope placement and analyze their impact on scale factor observability.
- 3) We evaluate the presented algorithm in simulations and experiments to assess its performance as a function of key motion and sensor parameters. Additionally, we evaluate the theoretically predicted impact of skewness error on the resulting calibration accuracy. Lastly, we consider the effect of device flexibility and propose a method for detecting and rejecting rotation data during the times that the rigidity assumption is violated.

The authors are with Apple Inc., Cupertino, CA 95014, USA. Email: {tianheng.wang2|stergios}@apple.com

¹Our focus is on multi-IMU calibration. For single-IMU calibration, the interested reader is referred to [17]–[19] and references therein.

II. PROBLEM FORMULATION

Our objective is to use angular velocity measurements from two gyroscopes to estimate their 3 degrees of freedom (d.o.f.) rotational extrinsics and scale factors. In what follows, we assume that the gyroscopes' data are properly time stamped and re-sampled² to ensure that their measurements correspond to the same time instants.

A. Problem Formulation

Each of the gyroscopes' measurements, ${}^i\boldsymbol{\omega}_m(k)$, $i = 1, 2$, at time-step k can be written as

$${}^i\boldsymbol{\omega}_m(k) = \mathbf{S}_i {}^i\boldsymbol{\omega}(k) + \mathbf{b}_i(k) + \mathbf{n}_i(k) \quad (1)$$

where ${}^i\boldsymbol{\omega}(k)$ is the true rotational velocity expressed w.r.t. frame $\{i\}$, $\mathbf{b}_i(k)$ and $\mathbf{n}_i(k)$ are the bias and noise for the i th gyroscope, respectively, and \mathbf{S}_i is the *upper-triangular* intrinsic matrix that represents the impact of scale and skewness.³ Since the two gyroscopes are rigidly attached to the same body, the true rotational velocities they measure are related through the rotational matrix $\mathbf{C} = \frac{1}{2}\mathbf{C}$ that transforms vectors expressed w.r.t. frame $\{2\}$ to frame $\{1\}$, i.e.,

$${}^1\boldsymbol{\omega}(k) = \mathbf{C} {}^2\boldsymbol{\omega}(k). \quad (2)$$

Solving (1) for ${}^i\boldsymbol{\omega}(k)$, $i = 1, 2$, and substituting in (2), yields

$${}^1\boldsymbol{\omega}_m(k) = \mathbf{A} {}^2\boldsymbol{\omega}_m(k) + \mathbf{b}(k) + \mathbf{n}(k) \quad (3)$$

where we have defined

$$\mathbf{A} := \mathbf{S}_1 \mathbf{C} \mathbf{S}_2^{-1} \quad (4)$$

$$\mathbf{b}(k) := \mathbf{b}_1(k) - \mathbf{A} \mathbf{b}_2(k) \quad (5)$$

$$\mathbf{n}(k) := \mathbf{n}_1(k) - \mathbf{A} \mathbf{n}_2(k). \quad (6)$$

Here $\mathbf{n}_1(k)$ and $\mathbf{n}_2(k)$ are considered to be independent white Gaussian noise sequences with $\text{Cov}\{\mathbf{n}_i(k)\} = \sigma_i^2 \mathbf{I}_3$. Since \mathbf{A} is close to orthonormal matrix, we have

$$\text{Cov}\{\mathbf{n}(k)\} \simeq \sigma^2 \mathbf{I}_3, \quad \sigma^2 = \sigma_1^2 + \sigma_2^2. \quad (7)$$

Hereafter, we assume that the bias in each gyro is approximately constant for the duration of the experiment (a few minutes) and ignore its dependence on time, which allows us to form the optimization problem:

$$\min_{\mathbf{A}, \mathbf{b}} \mathcal{C}(\mathbf{A}, \mathbf{b}) = \frac{1}{2\sigma^2} \sum_{k=1}^N \|\mathbf{A} {}^2\boldsymbol{\omega}_m(k) - \mathbf{b}\|^2 \quad (8)$$

At this point, we make the following key observations:

- 1) The optimization problem (8) is *quadratic* in the *combined* bias \mathbf{b} and matrix \mathbf{A} . This motivates our approach to first *directly* solve for \mathbf{A} , and then decompose it into the gyroscopes' intrinsics and extrinsics [see (4)].
- 2) The time steps from which we use rotational velocity measurements to form the cost function in (8) need *not* be consecutive. Instead, as shown in Section IV, one should select measurements from time instants when the rotational norm is large (and thus improve the signal to

²Any unknown time offset between the gyroscopes is typically computed by correlating their measurements' *norms*, and then compensated for via time alignment. Resampling can be done via cubic-spline interpolation.

³Note that scale is often referred to as *sensitivity*, while skewness is also known as *cross-axis sensitivity* and typically represents the effect of non-orthogonality of the gyroscopes' axes.

noise ratio), but also avoid periods of very fast motion where the rigidity assumption is violated.

- 3) The 9 elements of matrix \mathbf{A} are functions of the 15 unknown calibration d.o.f. corresponding to the 6 d.o.f. of each of the gyroscopes' intrinsics in \mathbf{S}_i , $i = 1, 2$ and 3 d.o.f. of the rotation matrix \mathbf{C} .

Due to the limited information available to the two gyroscopes when in motion, potentially up to 9 out of the 15 unknown d.o.f in \mathbf{A} can be determined. Besides the extrinsics, the gyroscopes' scale factors are usually susceptible to drift. In contrast, the skewness remains approximately constant over a wide range of conditions, and its impact on accuracy can be theoretically predicted as confirmed in Section IV-B. For this reason, from here on we assume that the gyroscopes' measurements have been compensated for skewness using the values provided by the manufacturer, and model the intrinsics as a diagonal matrix of scale factors:

$$\mathbf{S}_i := \text{diag}\{s_x, s_y, s_z\}. \quad (9)$$

This reduces the dimensionality of the problem from 15 to 9 and makes a potential decomposition of matrix \mathbf{A} into its factors possible. Note, however, that each of the 9 elements of \mathbf{A} is a *4th-order polynomial* in the unknowns. Specifically, \mathbf{A} is a bilinear function of the scale-factor matrices \mathbf{S}_1 and \mathbf{S}_2^{-1} . Additionally, the elements of the orientation matrix \mathbf{C} are 2nd-order polynomials of the orientation expressed using, e.g., Cayley or quaternion parameterization [28]. The optimization problem in (8) could be solved by linearizing \mathbf{A} in (4) w.r.t. the elements of \mathbf{S}_i , $i = 1, 2$ and \mathbf{C} , and iteratively minimizing the cost function using, e.g., Gauss-Newton. Such approach, however, is susceptible to local minima and may require many iterations to converge. For this reason, in the next section, we present the main contribution of this paper: A *direct solver* that is able to determine the global minimum of (8) in a single step.

B. Direct Least Squares Algorithm

The optimization problem (8) can be solved as follows

$$\min_{\mathbf{A}, \mathbf{b}} \mathcal{C}(\mathbf{A}, \mathbf{b}) \Leftrightarrow \min_{\mathbf{A}} \min_{\mathbf{b}|\mathbf{A}} \mathcal{C}(\mathbf{A}, \mathbf{b}) \quad (10)$$

where $\min_{\mathbf{b}|\mathbf{A}}$ means minimizing over \mathbf{b} given \mathbf{A} . The inner problem can be solved by setting the derivative of \mathcal{C} in (8) with respect to \mathbf{b} equal to zero, which yields:

$$\begin{aligned} \nabla_{\mathbf{b}} \mathcal{C} = 0 &\Rightarrow \sum_{k=1}^N (\mathbf{A} {}^2\boldsymbol{\omega}_m(k) - \mathbf{b}) \Big|_{\mathbf{b}=\mathbf{b}^*} = 0 \\ &\Rightarrow \mathbf{b}^* = \mathbf{A} \bar{\boldsymbol{\omega}}_m \end{aligned} \quad (11)$$

where

$${}^i\bar{\boldsymbol{\omega}}_m := \frac{1}{N} \sum_{k=1}^N {}^i\boldsymbol{\omega}_m(k), \quad i = 1, 2. \quad (12)$$

After substituting $\mathbf{b} = \mathbf{b}^*$ from (11) into (10), we have

$$\min_{\mathbf{A}, \mathbf{b}} \mathcal{C}(\mathbf{A}, \mathbf{b}) = \min_{\mathbf{A}} \mathcal{C}(\mathbf{A}, \mathbf{b}^*(\mathbf{A})) = \min_{\mathbf{A}} \mathcal{C}'(\mathbf{A}) \quad (13)$$

where

$$\mathcal{C}'(\mathbf{A}) = \frac{1}{2\sigma^2} \sum_{k=1}^N \left\| {}^1\check{\omega}_m(k) - \mathbf{A}^2\check{\omega}_m(k) \right\|^2 \quad (14)$$

$$= \frac{1}{2\sigma^2} \left\| \Omega_1 - \mathbf{A}\Omega_2 \right\|^2 \quad (15)$$

in which

$${}^i\check{\omega}_m(k) := {}^i\omega_m(k) - {}^i\bar{\omega}_m \quad (16)$$

$$\Omega_i := \left[{}^i\check{\omega}_m(1) \ \dots \ {}^i\check{\omega}_m(N) \right], \quad i = 1, 2. \quad (17)$$

The optimal solution \mathbf{A}^* of \mathbf{A} that minimizes (15) satisfies

$$\mathbf{A}^* \cdot (\Omega_2 \Omega_2^\top) = \Omega_1 \Omega_1^\top \quad (18)$$

i.e., normal equation of least squares.

Lemma 1: To recover the matrix \mathbf{A} in (18), two conditions need to be satisfied:

- 1) At least 4 pairs of measurements that span 3 d.o.f. are available.
- 2) If ${}^2\omega_m(k)$, $k = 1, 2, 3$, span 3 d.o.f., then ${}^2\omega_m(4) = \sum_{k=1}^3 \alpha_k {}^2\omega_m(k)$, and $\sum_{k=1}^3 \alpha_k \neq 1$ is required.

Proof: See Appendix I. ■

Once we have computed the optimal value of \mathbf{A} [from (18)], we employ (4) to recover the scale-factor matrices \mathbf{S}_1 and \mathbf{S}_2 for the two gyros, as well as the rotational matrix \mathbf{C} between them. To do so, we first express (4) w.r.t. \mathbf{C} :

$$\mathbf{C} = \mathbf{S}_1^{-1} \mathbf{A} \mathbf{S}_2 \quad (19)$$

and employ the orthonormality constraint:

$$\mathbf{C} \mathbf{C}^\top = \mathbf{I}_3 \Rightarrow \mathbf{A} \mathbf{S}_2^\top \mathbf{A}^\top = \mathbf{S}_1^\top. \quad (20)$$

Defining the matrix \mathbf{A} element-wise as $\mathbf{A} := [a_{ij}]$, $i, j = 1, 2, 3$, and substituting from (9), (20) can be re-written as:

$$\begin{bmatrix} \mathbf{A}_I & -\mathbf{I}_3 \\ \mathbf{A}_0 & \mathbf{0}_{3 \times 3} \end{bmatrix} \begin{bmatrix} \mathbf{s}_2 \\ \mathbf{s}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{3 \times 1} \end{bmatrix} \quad (21)$$

where the vectors \mathbf{s}_1 and \mathbf{s}_2 comprise the diagonal elements of the matrices \mathbf{S}_1^\top and \mathbf{S}_2^\top [see (9)], respectively, i.e.,

$$\mathbf{s}_i := [{}^i s_x^2 \quad {}^i s_y^2 \quad {}^i s_z^2]^\top, \quad i = 1, 2 \quad (22)$$

and

$$\mathbf{A}_I := \begin{bmatrix} a_{11}^2 & a_{12}^2 & a_{13}^2 \\ a_{21}^2 & a_{22}^2 & a_{23}^2 \\ a_{31}^2 & a_{32}^2 & a_{33}^2 \end{bmatrix} \quad (23)$$

$$\mathbf{A}_0 := \begin{bmatrix} a_{11}a_{21} & a_{12}a_{22} & a_{13}a_{23} \\ a_{21}a_{31} & a_{22}a_{32} & a_{23}a_{33} \\ a_{31}a_{11} & a_{32}a_{12} & a_{33}a_{13} \end{bmatrix}. \quad (24)$$

The process of solving (21) for the scale factors s_1 and s_2 , as well as the rotational matrix \mathbf{C} for three different configurations of the gyroscopes is presented in Section III.

C. Computational Complexity

At this point, we should note that the proposed algorithm has linear complexity $\mathcal{O}(N)$ in the number of gyro samples N ; this is due to (18) where the product of two pairs of $3 \times N$ matrices is computed before solving for the 3×3 matrix \mathbf{A} . The rest of the operations involve inverting or finding the null space of 3×3 matrices, and thus have constant cost.

III. OBSERVABILITY ANALYSIS

Hereafter, we turn our attention to the issue of determining how many d.o.f. are observable when considering two special and the general case of the gyroscopes' relative orientation.

A. Observability of Orientation and Scale Factors

Lemma 2: If the two gyroscopes are placed with their axes parallel to each other, then only 6 d.o.f. are observable: 3 for orientation and 3 for the pairwise ratios of the scale factors.

Proof: When the gyroscopes are placed with their axes parallel, the rotation matrix \mathbf{C} becomes a *signed* permutation matrix and, hence, \mathbf{A} a generalized permutation matrix. In this case, it is easy to see that \mathbf{A}_I in (23), which comprises the squares of the elements of \mathbf{A} , will also be a generalized permutation matrix of rank 3, while \mathbf{A}_0 in (24) will be $\mathbf{0}_{3 \times 3}$, i.e., of rank 0, as each row/column of \mathbf{A} contains only one non-zero element. Thus, (21) becomes:

$$\mathbf{A}_I \cdot \mathbf{s}_2 = \mathbf{s}_1 \quad (25)$$

from which we can only recover the pairwise ratios of the squared values of the scale factors corresponding to each of the gyroscopes parallel axes. Lastly, the 3 d.o.f. of the extrinsics matrix \mathbf{C} are computed by converting \mathbf{A} into a signed permutation matrix. ■

Lemma 3: If the two gyroscopes are placed so that only one pair of their axes are parallel, then 7 d.o.f. are observable: 3 for the gyroscopes' orientation, 1 for the ratio of the scale factors corresponding to the common axis \mathbf{e}_j , and 3 for the remaining 4 scale factors.

Proof: Without loss of generality, we consider the case when the gyroscopes' z axes are parallel and point in the same direction, i.e., their relative orientation corresponds to $\mathbf{C} = \mathbf{C}(\mathbf{e}_3, \theta)$, where $[\mathbf{e}_1 \ \mathbf{e}_2 \ \mathbf{e}_3] = \mathbf{I}_3$ and $\theta \in \mathbb{R}^*$. The analysis of all other cases is analogous after permuting the gyroscopes' axes so that their new z axes are aligned.⁴

Since \mathbf{A} has the same zeros structure as \mathbf{C} [see (4)], the non-diagonal elements corresponding to the row/column of the axis of rotation \mathbf{e}_j will be zero. Analogously, \mathbf{A}_I has zeros at the same locations as \mathbf{C} , while \mathbf{A}_0 will have a right null space of rank one. In particular, if $\mathbf{C} = \mathbf{C}(\mathbf{e}_3, \theta)$, then

$$\mathbf{A}_I = \begin{bmatrix} a_{11}^2 & a_{12}^2 & 0 \\ a_{21}^2 & a_{22}^2 & 0 \\ 0 & 0 & a_{33}^2 \end{bmatrix}, \quad \mathbf{A}_0 = \begin{bmatrix} a_{11}a_{21} & a_{12}a_{22} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Substituting in (21) and re-arranging terms yields a relation for the ratio of the z -axes' scale factors:

$$a_{33} = {}^1 s_z / {}^2 s_z \quad (26)$$

and one for the scale factors of the remaining four axes

$$\begin{bmatrix} a_{11}^2 & a_{12}^2 & -1 & 0 \\ a_{21}^2 & a_{22}^2 & 0 & -1 \\ a_{11}a_{21} & a_{12}a_{22} & 0 & 0 \end{bmatrix} \begin{bmatrix} {}^2 s_x^2 \\ {}^2 s_y^2 \\ {}^1 s_x^2 \\ {}^1 s_y^2 \end{bmatrix} = \mathbf{0}_{4 \times 1}. \quad (27)$$

⁴For example, the case when ${}^1x \parallel {}^2y$, and point in opposite directions, can be mapped to the one detailed here by first pre-multiplying the measurements of gyroscopes 1 and 2 with the signed permutation matrices $\mathbf{P}_1 = [\mathbf{e}_3 \ \mathbf{e}_2 \ \mathbf{e}_1]$ and $\mathbf{P}_2 = [\mathbf{e}_1 \ -\mathbf{e}_3 \ -\mathbf{e}_2]$, respectively.

As evident, the rows of the preceding matrix are linearly independent and hence the above equation has a unique right null space ζ from which the 4 remaining scale factors can be determined up to a global scale $\lambda \in \mathbb{R}^*$

$$\begin{bmatrix} {}^2s_x & {}^2s_y & {}^1s_x & {}^1s_y \end{bmatrix}^\top = \lambda \cdot \check{\zeta} \quad (28)$$

where $\check{\zeta}(i) = \sqrt{\zeta(i)}$, $i = 1, \dots, 4$. Lastly, substituting from (28) and (26) in (19), yields the rotational matrix

$$\mathbf{C}(\mathbf{e}_3, \theta) = \begin{bmatrix} a_{11} \frac{\check{\zeta}(1)}{\check{\zeta}(3)} & a_{12} \frac{\check{\zeta}(2)}{\check{\zeta}(3)} & 0 \\ a_{21} \frac{\check{\zeta}(1)}{\check{\zeta}(4)} & a_{22} \frac{\check{\zeta}(2)}{\check{\zeta}(4)} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (29)$$

with $\theta = \text{atan2}(a_{21}\check{\zeta}(1), a_{22}\check{\zeta}(2))$. ■

Lemma 4: If the two gyroscopes are placed so that none of their axes are parallel, then 8 d.o.f. are observable: 3 for the relative orientation and 5 for the gyroscopes' scale factors.

Proof: In the general case, where none of the elements of \mathbf{C} is zero, matrix \mathbf{A}_0 has rank 2.⁵ Note also, that \mathbf{A}_0 cannot be of rank 3, as this would result in $\mathbf{s}_2 = \mathbf{0}_{3 \times 1}$ and $\mathbf{s}_1 = \mathbf{0}_{3 \times 1}$, which is impossible. From (21), we have $\mathbf{A}_0 \mathbf{s}_2 = \mathbf{0}_{3 \times 1}$ which allows us to compute the optimal value of \mathbf{s}_2 and subsequently of \mathbf{s}_1 as:

$$\mathbf{s}_2^* = \text{null}(\mathbf{A}_0) \quad , \quad \mathbf{s}_1^* = \mathbf{A}_I \mathbf{s}_2^*. \quad (30)$$

Note that since \mathbf{s}_2^* is a *unit vector*, the *absolute* scale for the gyros cannot be recovered. Once \mathbf{s}_1^* and \mathbf{s}_2^* have been determined, we compute the corresponding \mathbf{S}_1^* and \mathbf{S}_2^* from (22) and (9). Subsequently, we recover the unknown rotational matrix \mathbf{C} from (19), as

$$\mathbf{C}^* = \mathbf{S}_1^{*-1} \mathbf{A}^* \mathbf{S}_2^*. \quad (31)$$

Lastly, we ensure that $\mathbf{C}^* \in \text{SO}(3)$ by finding the nearest orthonormal matrix to \mathbf{C}^* as in [29]. ■

Summarizing the previous four Lemmas, we have shown that

Theorem 1: The relative orientation of two gyroscopes is always observable when two conditions are satisfied:

- 1) At least 4 pairs of measurements that span 3 d.o.f. are available.
- 2) If ${}^2\omega_m(k)$, $k = 1, 2, 3$, span 3 d.o.f., then ${}^2\omega_m(4) = \sum_{k=1}^3 \alpha_k {}^2\omega_m(k)$, and $\sum_{k=1}^3 \alpha_k \neq 1$ is required.

Theorem 2: The scale factors of two gyroscopes are observable, up to a global scale, except for special cases:

- If one pair of the gyroscopes' axes are parallel: 4 scale factors are observable, up to a global scale, and only the ratio of the remaining two, corresponding to the parallel axes, can be determined;
- If all of the gyroscopes' axes are parallel, only the pairwise ratios of their scale factors can be determined.

A key practical implication from Theorem 2, is that when installing multiple gyroscopes on a device, care should be taken to place them so that they have no parallel axes.

⁵Due to numerical errors, the matrix \mathbf{A}_0 is almost never rank deficient. Thus, we typically compute \mathbf{s}_2^* as the eigenvector corresponding to the smallest eigenvalue of matrix $\mathbf{A}_0^\top \mathbf{A}_0$.

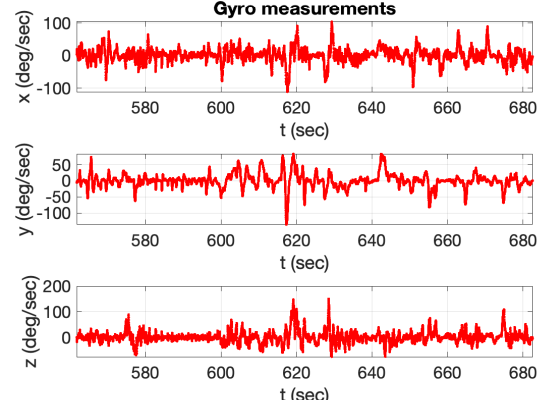


Fig. 1: Example of gyroscope 1 measurements.

B. Recovering Global Scale

As shown in Theorem 2, global scale is not observable without additional information, even when the gyroscopes are not in a special configuration. To recover scale, one needs to consider either *additional sensors* or *prior information*. Specifically, if one or more cameras are also used in conjunction with a pair of IMUs in a visual-inertial estimator, then instead of six, only one d.o.f., that of global scale, would need to be estimated, hence reducing the dimensionality of the problem and improving accuracy. On the other hand, if it is known that particular axes of the gyroscopes, e.g., the x axes, have stable scale factors close to one (i.e., ${}^i s_x \simeq 1$, $i = 1, 2$), then global scale can be recovered by solving the following constrained optimization problem:

$$\min_{{}^1s_x, {}^2s_x} f({}^1s_x, {}^2s_x) = ({}^1s_x - 1)^2 + ({}^2s_x - 1)^2 \quad (32)$$

$$\text{s.t. } {}^1s_x / {}^2s_x = \lambda \quad (33)$$

where $\lambda := {}^1s_x / {}^2s_x$ is the scale factor ratio estimated from the direct method of Section III. The solution is given by

$${}^1s_x^* = \frac{\lambda(\lambda + 1)}{\lambda^2 + 1}, \quad {}^2s_x^* = \frac{\lambda + 1}{\lambda^2 + 1}. \quad (34)$$

The remaining scale factors 1s_y , 1s_z , 2s_y , and 2s_z are recovered by maintaining their ratios w.r.t. 1s_x and 2s_x . In particular ${}^i s_y^* = ({}^i s_y / {}^i s_x) {}^i s_x^*$ and ${}^i s_z^* = ({}^i s_z / {}^i s_x) {}^i s_x^*$, where the parameters without * are those before optimization.

IV. SIMULATION VALIDATION

In our simulations, we employ data from human motion (see Fig. 1) to emulate the ideal measurements of gyroscope 1 sampled at 100 Hz. Then, we use the true rotational extrinsics \mathbf{C} to generate the ideal data of gyroscope 2. Next, we introduce intrinsic errors (in scale factors and skewness for some tests) and finally, we contaminate the gyroscopes' data [see (1)] with additive noise $\mathbf{n}_i(k) \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_3)$, $\sigma = 0.1$ deg/s and time-varying bias modelled as:

$$\mathbf{b}_i(k+1) = \mathbf{b}_i(k) + \boldsymbol{\nu}_i(k+1) \quad (35)$$

where $\boldsymbol{\nu}_i(k) \sim \mathcal{N}(\mathbf{0}, \sigma_\nu^2 \mathbf{I})$, $\sigma_\nu = 57.3 \mu\text{deg/s}$.

Two key metrics include the root mean squared error (RMSE) accuracy and signal-to-noise ratio (SNR):

$$\text{RMSE} := \left(\frac{1}{M} \sum_{i=1}^M e_i^2 \right)^{\frac{1}{2}}, \quad \text{SNR} = \left(\frac{1}{\sigma^2} \sum_{k=1}^N \|\omega_k\|^2 \right)^{\frac{1}{2}} \quad (36)$$

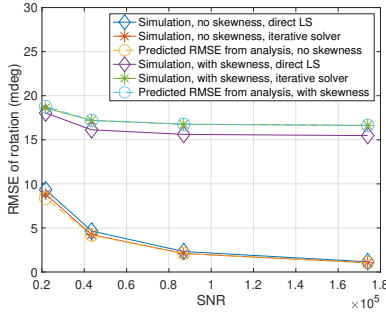


Fig. 2: Rotation RMSE w.r.t. SNR. Circles represent the predicted RMSE due to skewness (40).

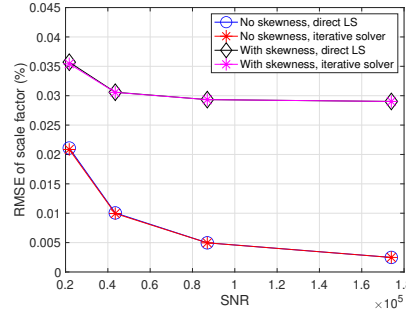


Fig. 3: Scale factor RMSE (%) w.r.t. SNR.

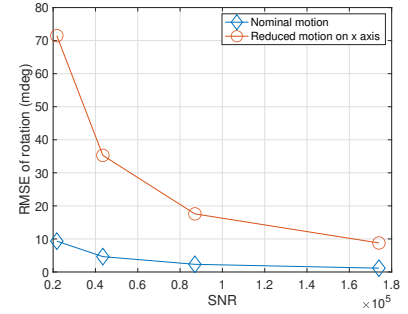


Fig. 4: Rotation RMSE with reduced velocity along x axis, compared with RMSE with nominal motion.

where M is the number of Monte-Carlo runs. The error for the i th Monte-Carlo run, e_i , is defined as follows, for orientation and intrinsics, respectively:

$$e_i = \left\| \text{Log}(\hat{\mathbf{C}}\mathbf{C}^T) \right\| \quad (37)$$

$$e_i = \frac{1}{\sqrt{5}} \left\| \frac{1}{\hat{s}_x} \begin{bmatrix} 1 \\ \hat{s}_y \\ \hat{s}_z \end{bmatrix} - \frac{1}{s_x} \begin{bmatrix} 1 \\ s_y \\ s_z \end{bmatrix} \right\| \quad (38)$$

where $\hat{\mathbf{C}}$ and \mathbf{C} are the estimated and true rotations, and ${}^i\hat{\mathbf{s}} = [{}^i\hat{s}_x \ {}^i\hat{s}_y \ {}^i\hat{s}_z]^T$ and ${}^i\mathbf{s} = [{}^i s_x \ {}^i s_y \ {}^i s_z]^T$ are the estimated and true gyroscopes' scale factors for $i = 1, 2$.

Asymptotically, the RMSE coincides with the square-root of error *covariance* for maximum-likelihood estimators [30]. Furthermore, a lower bound on the error covariance of the estimated orientation is derived in Appendix II

$$\text{Tr}\{\mathbb{E}\{\mathbf{P}\}\} \geq 4.5(\mathbb{E}\{\text{SNR}^2\})^{-1}. \quad (39)$$

According to this expression, a lower bound on the RMSE is expected to be approximately proportional to the inverse of SNR, and hence to the

- Inverse of the gyroscope velocity norm
- Standard deviation of noise σ , and
- $1/\sqrt{N}$ assuming the data follow a stationary process.

A. Performance with Known Skewness

In order to validate the relationship between RMSE and SNR, we first consider the case without skewness error (i.e., the intrinsic matrices \mathbf{S}_1 and \mathbf{S}_2 are diagonal) and “scale” the SNR of the true gyroscope data in Fig. 1, by factors of 0.5, 1, 2, and 4. This is accomplished by appropriately scaling the velocities' norms, noise σ , or data sample size. As evident from the results in Figs. 2 and 3, respectively, for the estimated rotation and scale factors, their RMSEs are indeed proportional to the inverse of the SNR.

Next, we turn our attention to the fact that imbalance in the rotational velocities along each of the three axes affects calibration accuracy. To illustrate this, we purposely reduce the rates about the x axis 10 fold, while amplifying the ones around the y and z axes such that the SNR remains the same. As shown in Fig. 4, the rotational RMSE increases significantly. This is to be expected as the information matrix, and hence the covariance, depends on the individual axes' rotational velocities [see (52)]. For this reason, in practice,

the measurements used for calibration should be selected so as to ensure sufficient motion along all three axes.

B. Performance with Imperfect Skewness Calibration

Next, we investigate the impact of skewness errors. In [31], we derive the following closed-form expression for the rotation error due to skewness:

$$\tilde{\boldsymbol{\theta}} \simeq \frac{1}{2} (\mathbf{w}_1 - \mathbf{C}\mathbf{w}_2 + \mathbf{n}) \quad (40)$$

where \mathbf{n} is the LS error when solving for \mathbf{A} from (18), which goes to zero as the number of measurements increases,

$$\mathbf{w}_i \simeq [-\tilde{\mathbf{S}}_i(2, 3) \ \tilde{\mathbf{S}}_i(1, 3) \ -\tilde{\mathbf{S}}_i(1, 2)]^T, \quad i = 1, 2 \quad (41)$$

and $\tilde{\mathbf{S}}_i(1, 2)$, $\tilde{\mathbf{S}}_i(1, 3)$, and $\tilde{\mathbf{S}}_i(2, 3)$ are the skewness errors in the corresponding intrinsic matrices. According to (40), the rotation error is asymptotically bounded by the skewness errors. To assess this in simulation, the three upper off-diagonal elements in \mathbf{S}_1 and \mathbf{S}_2 , are randomly drawn⁶ from $\mathcal{N}(0, \sigma_s^2)$ with $\sigma_s = 2.357 \times 10^{-4}$. As seen from the results depicted in Figs. 2 and 3, there is a strong agreement between the estimated rotation error and that predicted from (40).

Lastly, Figs. 2 and 3 confirm that the proposed *direct* LS algorithm achieves performance indistinguishable to that of a linearization-based iterative solver (see Appendix III), without any iterations.

TABLE I: Orientation errors (mdeg)

θ_x	θ_y	θ_z
23.52	-72.82	82.39
23.91	-71.91	75.09
26.48	-72.02	77.28

TABLE II: Scale factor errors (%)

1s_x	1s_y	1s_z	2s_x	2s_y	2s_z
0.01	-0.04	-0.01	-0.01	-0.08	-0.05
-0.01	-0.03	-0.03	0.01	-0.08	0.00
-0.02	0.01	-0.04	0.02	-0.02	0.01

V. EXPERIMENTAL RESULTS

A. Rigid Motion

The proposed algorithm is employed for infield calibration of a device equipped with two IMUs. In this process, we accumulate rotational velocity measurements till we reach an SNR of 4×10^3 per axis. To evaluate performance, we use as ground truth (GT) the parameters determined by a high-accuracy calibration station. The errors in rotation and scale factors are shown in Tables I and II, respectively. In these tests, the global scale is resolved by employing (32) and considering as prior that the x axes's scale factors remain close to 1. These results confirm the proposed infield algorithm's ability to achieve high calibration accuracy.

B. Robustness to Device Flexibility

Under fast motion, some devices may flex which violates the assumption of rigidity between the gyroscopes. Using data from these time instants will negatively impact calibration performance. In what follows, we present our approach for detecting and removing such measurements.

To better understand the effect of device flex we compute the residual, i.e., the difference in the two gyroscopes' measurements after accounting for intrinsics and extrinsics:

$$\mathbf{r}(k) = {}^1\hat{\boldsymbol{\omega}}_m(k) - \hat{\mathbf{S}}_1 \hat{\mathbf{C}} \hat{\mathbf{S}}_2^{-1} {}^2\hat{\boldsymbol{\omega}}_m(k). \quad (42)$$

As shown in Fig. 5 (right), for rigid motions, the true rotational extrinsics are time-invariant, and the residuals resemble the norm of a 3×1 Gaussian noise vector independent of the magnitude of rotation. In contrast, when the device undergoes very fast motion and flexes, the true rotational extrinsics are time-variant, and the residuals become strongly correlated with the rotational velocity norm - see Fig. 5 (left).

Hereafter, we present a simple but effective way to robustify the calibration algorithm against device flexibility. First, we run the infield algorithm once to compute the residuals at each time step and their standard deviation σ_r . Then, we remove gyroscope data whose residuals have magnitude greater than $3\sigma_r$, as well as measurements immediately before and after these time instants to account for hysteresis in the onset/settle of device flex. Finally, we re-run the algorithm using the remaining gyroscope measurements. As evident from the results shown in Table III, the rotation accuracy is greatly improved by following this approach.

VI. CONCLUSION

In this paper, we address the problem of determining the *relative orientation* of two rigidly connected parts of a device that may have undergone plastic deformation. To do so, we considered data available from two 3-axial gyroscopes – each attached on the corresponding part of interest – and introduced a *direct* algorithm for determining the gyroscopes' extrinsics, as well as their scale factors, up to a global scale. The main advantages of our algorithm is that it does not rely on specialized equipment or particular motions. Instead, it can be employed infield without using any additional

⁶The range of values used for skewness as well as for bias drift and additive noise are representative of MEMS gyroscopes' performance.

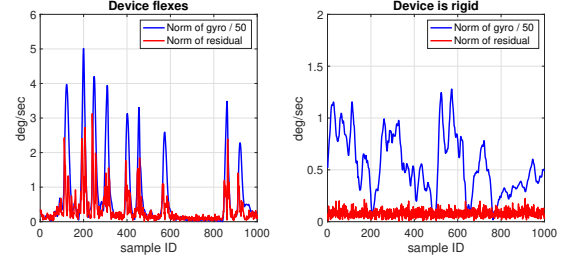


Fig. 5: Norm of residual vs. scaled norm of gyro data.

sensors and delivers high accuracy, comparable to that of a properly initialized iterative solver, in a single step. As part of our main contributions, we analyzed the observability of the system comprising the two gyroscopes and showed that both extrinsics and scale factors (up to global scale) can be determined. Additionally, we assessed the impact of the gyroscopes' skewness error on the extrinsics accuracy. Furthermore, we validated the performance of the proposed algorithm both in simulation and real-world experiments, where we established relations between the expected error and the SNR which is a function of motion profile, gyro noise, and the size of the data used. Moreover, we defined metrics for determining when there is sufficient information for accurate estimation and when calibration should be avoided due to device flexibility. Lastly, we introduced a simple but effective method for determining device flex and removing data that negatively affects estimation accuracy.

TABLE III: Orientation errors mdeg before and after removing measurements via the proposed *mitigation* scheme.

	θ_x	θ_y	θ_z
Before	19.24	985.92	401.91
After	-28.07	-14.05	-3.09

APPENDIX I PROOF OF LEMMA 1

Note that the column vectors of $\boldsymbol{\Omega}_2$ in (17) are *re-centered* measurements ${}^2\hat{\boldsymbol{\omega}}_m(k)$ as in (16). To recover \mathbf{A} from (18), we need $\text{rank}(\boldsymbol{\Omega}_2) = 3$. Suppose only three measurements ${}^2\hat{\boldsymbol{\omega}}_m(k)$ spanning 3 d.o.f., are available. Also, let \mathbf{P}_N be an $N \times N$ matrix whose elements are given by

$$\mathbf{P}_N(i, j) = \begin{cases} (N-1)/N, & i = j \\ -1/N, & \text{otherwise} \end{cases}. \quad (43)$$

Then $\boldsymbol{\Omega}_2$ in (17) for $N = 3$ can be written as

$$\boldsymbol{\Omega}_2 = [{}^2\hat{\boldsymbol{\omega}}_m(1) \quad {}^2\hat{\boldsymbol{\omega}}_m(2) \quad {}^2\hat{\boldsymbol{\omega}}_m(3)] \mathbf{P}_3. \quad (44)$$

Since the elements of every row or column of \mathbf{P}_3 sum up to 0, $\text{rank}(\mathbf{P}) = 2$. This implies $\text{rank}(\boldsymbol{\Omega}_2) \leq 2$. For this reason, a 4-th measurement ${}^2\hat{\boldsymbol{\omega}}_m(4)$, different from the previous linearly independent ones, is necessary.

Since ${}^2\hat{\boldsymbol{\omega}}_m(k)$, $k = 1, 2, 3$, span 3 d.o.f.,

$${}^2\hat{\boldsymbol{\omega}}_m(4) = \alpha_1 {}^2\hat{\boldsymbol{\omega}}_m(1) + \alpha_2 {}^2\hat{\boldsymbol{\omega}}_m(2) + \alpha_3 {}^2\hat{\boldsymbol{\omega}}_m(3) \quad (45)$$

for some α_1, α_2 , and α_3 . Then Ω_2 in (17) for $N = 4$ can be written as

$$\Omega_2 = [{}^2\omega_m(1) \quad {}^2\omega_m(2) \quad {}^2\omega_m(3) \quad {}^2\omega_m(4)] \mathbf{P}_4 \quad (46)$$

$$= \frac{1}{4} [{}^2\omega_m(1) \quad {}^2\omega_m(2) \quad {}^2\omega_m(3)] [\mathbf{D} \quad \mathbf{d}] \quad (47)$$

where $\mathbf{d} = [-1 + 3\alpha_1 \quad -1 + 3\alpha_2 \quad -1 + 3\alpha_3]^\top$ and

$$\mathbf{D} = \begin{bmatrix} 3 - \alpha_1 & -1 - \alpha_1 & -1 - \alpha_1 \\ -1 - \alpha_2 & 3 - \alpha_2 & -1 - \alpha_2 \\ -1 - \alpha_3 & -1 - \alpha_3 & 3 - \alpha_3 \end{bmatrix}. \quad (48)$$

The *re-centered* measurements are linearly dependent since

$${}^2\check{\omega}_m(1) + {}^2\check{\omega}_m(2) + {}^2\check{\omega}_m(3) + {}^2\check{\omega}_m(4) = \mathbf{0}_{3 \times 1}. \quad (49)$$

Thus, to have $\text{rank}(\Omega_2) = 3$, its first three columns ${}^2\check{\omega}_m(k)$, $k = 1, 2, 3$, need to span 3 d.o.f., which requires \mathbf{D} in (47) to be invertible. Since

$$\det(\mathbf{D}) = 16 \cdot (1 - \alpha_1 - \alpha_2 - \alpha_3) \quad (50)$$

$\alpha_1 + \alpha_2 + \alpha_3 \neq 1$ is a necessary and sufficient condition for \mathbf{D} to be invertible.

APPENDIX II

LOWER BOUND ON ERROR COVARIANCE

Consider the simplified case of zero bias and perfect intrinsics where only rotation is estimated. The residual for the k th pair of measurements is:

$$\mathbf{r}(k) = {}^1\omega_m(k) - \hat{\mathbf{C}}^2\omega_m(k) \quad (51)$$

$$\begin{aligned} &\simeq {}^1\omega(k) + \mathbf{n}_1(k) - (\mathbf{I}_3 + [\tilde{\boldsymbol{\theta}} \times])\mathbf{C}({}^2\omega(k) + \mathbf{n}_2(k)) \\ &\simeq [{}^1\omega(k) \times] \tilde{\boldsymbol{\theta}} + \mathbf{n}_1(k) + \mathbf{C} \mathbf{n}_2(k) \end{aligned}$$

where $\hat{\mathbf{C}} \simeq (\mathbf{I}_3 + [\tilde{\boldsymbol{\theta}} \times])\mathbf{C}$ with $\tilde{\boldsymbol{\theta}}$ being the rotation error. Furthermore, let $\text{Cov}\{\mathbf{n}_{12}\} = \sigma^2\mathbf{I}_3$. Then, the information matrix from the k th measurement is:

$$\begin{aligned} \mathbf{H}_k &= \frac{1}{\sigma^2} [\omega(k) \times]^\top [\omega(k) \times] \\ &= \frac{1}{\sigma^2} \begin{bmatrix} \omega_{k,2}^2 + \omega_{k,3}^2 & -\omega_{k,1}\omega_{k,2} & -\omega_{k,1}\omega_{k,3} \\ -\omega_{k,1}\omega_{k,2} & \omega_{k,1}^2 + \omega_{k,3}^2 & -\omega_{k,2}\omega_{k,3} \\ -\omega_{k,1}\omega_{k,3} & -\omega_{k,2}\omega_{k,3} & \omega_{k,1}^2 + \omega_{k,2}^2 \end{bmatrix} \end{aligned} \quad (52)$$

where we have dropped the superscript 1 for clarity, and $\omega_{k,i}$ is the i th element of $\omega(k)$.

Let \mathbf{P} and \mathbf{H} be the 3×3 error covariance and information matrix, respectively. We have

$$\mathbf{P} = \mathbf{H}^{-1} = \left(\sum_{k=1}^N \mathbf{H}_k \right)^{-1}. \quad (53)$$

Furthermore, suppose the rotation velocities are zero mean and uncorrelated across axes. Then, from (53) and (52), and using Jensen's inequality [32], we have

$$\mathbb{E}\{\mathbf{P}\} \succeq \left(\sum_{k=1}^N \mathbb{E}\{\mathbf{H}_k\} \right)^{-1} = \sigma^2 \text{diag}\{E_{23}^{-1}, E_{13}^{-1}, E_{12}^{-1}\}$$

where $\mathbf{A} \succeq \mathbf{B}$ means $\mathbf{A} - \mathbf{B}$ is positive semi-definite, and $E_{ij} := \sum_{k=1}^N \mathbb{E}\{\omega_{k,i}^2 + \omega_{k,j}^2\}$. Therefore:

$$\text{Tr}\{\mathbb{E}\{\mathbf{P}\}\} \geq \sigma^2(E_{23}^{-1} + E_{13}^{-1} + E_{12}^{-1}) \quad (54)$$

$$\geq 9\sigma^2(E_{23} + E_{13} + E_{12})^{-1} \quad (55)$$

$$= 4.5(\mathbb{E}\{\text{SNR}^2\})^{-1} \quad (56)$$

where (56) uses (36), and (55) uses Jensen's inequality and that $f(x) = 1/x$ is a convex function.

APPENDIX III ITERATIVE SOLVER

The intrinsic error is modeled as

$$\mathbf{S}_i = \hat{\mathbf{S}}_i + \tilde{\mathbf{S}}_i = \hat{\mathbf{S}}_i(\mathbf{I}_3 + \tilde{\mathbf{S}}_i') \quad (57)$$

in which $\hat{\mathbf{S}}_i$ is the current estimate of \mathbf{S}_i , $\tilde{\mathbf{S}}_i$ is the estimation error, and $\tilde{\mathbf{S}}_i' = \hat{\mathbf{S}}_i^{-1}\tilde{\mathbf{S}}_i$ is the percentage formulation of the error estimated by the iterative solver.

Multiplying both sides of (1) with $\hat{\mathbf{S}}_i^{-1}$, we have

$${}^i\omega_m'(k) := \hat{\mathbf{S}}_i^{-1}{}^i\omega_m(k) = (\mathbf{I}_3 + \tilde{\mathbf{S}}_i'){}^i\omega(k) + \mathbf{b}_i' + \mathbf{n}_i'(k)$$

where $\mathbf{b}_i' := \hat{\mathbf{S}}_i^{-1}\mathbf{b}_i$ and $\mathbf{n}_i'(k) := \hat{\mathbf{S}}_i^{-1}\mathbf{n}_i(k)$. Furthermore, the true rotation is approximated as $\mathbf{C} \simeq (\mathbf{I}_3 - [\tilde{\boldsymbol{\theta}} \times])\hat{\mathbf{C}}$. Then the residual can be defined as

$$\mathbf{r}(k) = {}^1\omega_m'(k) - \hat{\mathbf{C}}^2\omega_m'(k) \quad (58)$$

$$\begin{aligned} &\simeq (\mathbf{I}_3 + \tilde{\mathbf{S}}_1'){}^1\omega(k) + \mathbf{b}_1' + \mathbf{n}_1'(k) \\ &\quad - (\mathbf{I}_3 + [\tilde{\boldsymbol{\theta}} \times])\mathbf{C}((\mathbf{I}_3 + \tilde{\mathbf{S}}_2'){}^2\omega(k) + \mathbf{b}_2' + \mathbf{n}_2'(k)) \\ &= [{}^1\omega(k) \times] \tilde{\boldsymbol{\theta}} + \mathbf{\Gamma}^1\omega(k) + \mathbf{b}_{12} + \mathbf{n}_{12}(k) \end{aligned}$$

where $\mathbf{b}_{12} := \mathbf{b}_1' - (\mathbf{I}_3 + [\tilde{\boldsymbol{\theta}} \times])\hat{\mathbf{C}}\mathbf{b}_2'$, $\mathbf{n}_{12}(k) := \mathbf{n}_1'(k) - \hat{\mathbf{C}}\mathbf{n}_2'(k)$, and

$$\mathbf{\Gamma} := \tilde{\mathbf{S}}_1' - \mathbf{C}\tilde{\mathbf{S}}_2'\mathbf{C}^\top \quad (59)$$

which is symmetric since $\tilde{\mathbf{S}}_1'$ and $\tilde{\mathbf{S}}_2'$ are diagonal. Let

$$\boldsymbol{\gamma} := [\gamma_{11} \quad \gamma_{12} \quad \gamma_{13} \quad \gamma_{22} \quad \gamma_{23} \quad \gamma_{33}]^\top \quad (60)$$

with γ_{ij} being the (i, j) th element of $\mathbf{\Gamma}$, We have

$$\begin{aligned} \mathbf{r}(k) &= [{}^1\omega(k) \times] \tilde{\boldsymbol{\theta}} + \mathbf{\Omega}(k)\boldsymbol{\gamma} + \mathbf{b}_{12} + \mathbf{n}_{12}(k) \\ &= \mathbf{J}(k)\mathbf{x} + \mathbf{b}_{12} + \mathbf{n}_{12}(k) \end{aligned} \quad (61)$$

where

$$\mathbf{\Omega}(k) = \begin{bmatrix} \omega_{1,k} & \omega_{2,k} & \omega_{3,k} & 0 & 0 & 0 \\ 0 & \omega_{1,k} & 0 & \omega_{2,k} & \omega_{3,k} & 0 \\ 0 & 0 & \omega_{1,k} & 0 & \omega_{2,k} & \omega_{3,k} \end{bmatrix} \quad (62)$$

$$\mathbf{J}(k) = [[{}^1\omega(k) \times] \quad \mathbf{\Omega}(k)] \quad (63)$$

$$\mathbf{x} = [\tilde{\boldsymbol{\theta}}^\top \quad \boldsymbol{\gamma}^\top]^\top \quad (64)$$

in which $\omega_{1,k}$, $\omega_{2,k}$, and $\omega_{3,k}$ are the 3 elements of ${}^1\omega(k)$. Furthermore, since $\hat{\mathbf{S}}_i \simeq \mathbf{I}_3$ and $\hat{\mathbf{C}}$ is an orthonormal matrix, we have $\text{Cov}\{\mathbf{n}_{12}\} \simeq \sigma^2\mathbf{I}_3$ with $\sigma^2 = \sigma_1^2 + \sigma_2^2$.

Using (61), a least-squares problem can be formulated with objective function given by

$$f(\mathbf{x}, \mathbf{b}_{12}) = \frac{1}{2\sigma^2} \sum_{k=1}^N \|\mathbf{r}(k) - \mathbf{J}(k)\mathbf{x} - \mathbf{b}_{12}\|^2. \quad (65)$$

To find the optimal solution, first solve for \mathbf{b}_{12} as

$$\frac{\partial f}{\partial \mathbf{b}_{12}} = \mathbf{0} \Rightarrow \mathbf{b}_{12}^*(\mathbf{x}) = \frac{1}{N} \sum_{k=1}^N (\mathbf{r}(k) - \mathbf{J}(k)\mathbf{x}) = \bar{\mathbf{r}} - \bar{\mathbf{J}}\mathbf{x}.$$

Substituting $\mathbf{b}_{12} = \mathbf{b}_{12}^*(\mathbf{x})$ into (65), we obtain an objective function about \mathbf{x} as

$$f'(\mathbf{x}) = \frac{1}{2\sigma^2} \sum_{k=1}^N \|\check{\mathbf{r}}(k) - \check{\mathbf{J}}(k)\mathbf{x}\|^2 \quad (66)$$

where

$$\check{\mathbf{r}}(k) = \mathbf{r}(k) - \bar{\mathbf{r}} = {}^1\check{\omega}_m'(k) - \hat{\mathbf{C}}^2\check{\omega}_m'(k) \quad (67)$$

$$\check{\mathbf{J}}(k) = \mathbf{J}(k) - \bar{\mathbf{J}} = [[{}^1\check{\omega}(k) \times] \quad \check{\mathbf{\Omega}}(k)] \quad (68)$$

in which (67) uses (58),

$$\check{\Omega}(k) = \begin{bmatrix} \check{\omega}_{1,k} & \check{\omega}_{2,k} & \check{\omega}_{3,k} & 0 & 0 & 0 \\ 0 & \check{\omega}_{1,k} & 0 & \check{\omega}_{2,k} & \check{\omega}_{3,k} & 0 \\ 0 & 0 & \check{\omega}_{1,k} & 0 & \check{\omega}_{2,k} & \check{\omega}_{3,k} \end{bmatrix}$$

where $\check{\omega}_{1,k}$, $\check{\omega}_{2,k}$, and $\check{\omega}_{3,k}$ are the 3 elements of ${}^1\check{\omega}(k)$. Note that the Jacobian $\check{\mathbf{J}}(k)$ in (68) can be computed using the measurements.

Then, the iterative solver using, e.g., Gauss-Newton's method, solves a least squares problem for \mathbf{x} in each iteration with objective function given by (66). After that, the diagonal elements of $\check{\mathbf{S}}'_1$ and $\check{\mathbf{S}}'_2$, ${}^1\check{s}'$ and ${}^2\check{s}'$, can be recovered from (59) for $\mathbf{C} = \check{\mathbf{C}}$. Since Γ is symmetric, (59) provides 6 linear equations for \check{s}'_1 and \check{s}'_2

$$\begin{bmatrix} \gamma_{11} \\ \gamma_{12} \\ \gamma_{13} \\ \gamma_{22} \\ \gamma_{23} \\ \gamma_{33} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & -\mathbf{c}_1 \odot \mathbf{c}_1 \\ 0 & 0 & 0 & -\mathbf{c}_1 \odot \mathbf{c}_2 \\ 0 & 0 & 0 & -\mathbf{c}_1 \odot \mathbf{c}_3 \\ 0 & 1 & 0 & -\mathbf{c}_2 \odot \mathbf{c}_2 \\ 0 & 0 & 0 & -\mathbf{c}_2 \odot \mathbf{c}_3 \\ 0 & 0 & 1 & -\mathbf{c}_3 \odot \mathbf{c}_3 \end{bmatrix}}_{\Phi} \begin{bmatrix} {}^1\check{s}_x \\ {}^1\check{s}_y \\ {}^1\check{s}_z \\ {}^2\check{s}_x \\ {}^2\check{s}_y \\ {}^2\check{s}_z \end{bmatrix} \quad (69)$$

where \mathbf{c}_i is the i th row of \mathbf{C} , and \odot means element-wise multiplication. Since $\mathbf{C} \in \text{SO}(3)$, the sum of the 3 elements of $\mathbf{c}_i \odot \mathbf{c}_j$ is 1 for $i = j$, and 0 otherwise. As a result, the column vectors of Φ are linearly dependent, and at most 5 d.o.f. of the scale factors can be recovered, which agrees with Theorem 2. Thus, we fix the 1st element of ${}^1\check{s}'$ as zero and solve for the remaining 5 scale factor errors. Lastly, for the singular cases of Lemmas 2 and 3, we need to fix the unobservable elements in ${}^1\check{s}'$ and ${}^2\check{s}'$ as zero, and solve for the remaining observable d.o.f. of scale factor errors. The iterative algorithm can be initialized using the solution from the direct solver. The computational complexity of one iteration is $O(N)$, same as that for the direct solver.

REFERENCES

- [1] A. I. Mourikis and S. I. Roumeliotis, "A multi-state constraint Kalman filter for vision-aided inertial navigation," in *Proc. of IEEE International Conference on Robotics and Automation*, Rome, Italy, 2007.
- [2] M. Li and A. I. Mourikis, "High-precision, consistent EKF-based visual-inertial odometry," *International Journal of Robotics Research*, vol. 32, no. 6, pp. 690–711, May 2013.
- [3] S. Leutenegger, S. Lynen, M. Bosse, R. Siegwart, and P. Furgale, "Keyframe-based visual-inertial odometry using nonlinear optimization," *International Journal of Robotics Research*, vol. 34, no. 3, pp. 314–334, 2015.
- [4] K. Wu, A. M. Ahmed, G. A. Georgiou, and S. I. Roumeliotis, "A square root inverse filter for efficient vision-aided inertial navigation on mobile devices," in *Proc. of Robotics: Science and Systems*, vol. 2, Rome, Italy, July 2015.
- [5] T. Do, L. C. Carrillo-Arce, and S. I. Roumeliotis, "High-speed autonomous quadrotor navigation through visual and inertial paths," *International Journal of Robotics Research*, vol. 38, no. 4, pp. 486–504, 2019.
- [6] T. Qin, P. Li, and S. Shen, "VINS-Mono: A robust and versatile monocular visual-inertial state estimator," *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 1004–1020, Aug. 2018.
- [7] P. Geneva, K. Eickenhoff, W. Lee, Y. Yang, and G. Huang, "OpenVINS: a research platform for visual-inertial estimation," in *Proc. of the IEEE International Conference on Robotics and Automation*, Paris, France, 2020.
- [8] W. Lee, Y. Yang, and G. Huang, "Efficient multi-sensor aided inertial navigation with online calibration," in *Proc. of the IEEE International Conference on Robotics and Automation*, Xi'an, China, 2021.
- [9] S. Rahman, A. Quattrini Li, and I. Rekleitis, "SVIn2: A multi-sensor fusion-based underwater SLAM system," *International Journal of Robotics Research*, vol. 41, no. 11–12, pp. 1022–1042, July 2022.
- [10] X. Zuo, P. Geneva, W. Lee, Y. Liu, and G. Huang, "LIC-Fusion: Lidar-inertial-camera odometry," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, Macau, China, Nov. 2019.
- [11] X. Zuo, Y. Yang, P. Geneva, J. Lv, Y. Liu, G. Huang, and M. Pollefeys, "LIC-Fusion 2.0: Lidar-inertial-camera odometry with sliding-window plane-feature tracking," in *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Las Vegas, NV, 2020.
- [12] S. Zheng, Z. Li, Y. Liu, H. Zhang, P. Zheng, X. Liang, X. Bu, and X. Zou, "UWB-VIO fusion for accurate and robust relative localization of round robotic teams," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, p. 11950–11957, Oct. 2022.
- [13] B. Gao, B. Lian, D. Wang, and C. Tang, "Low drift visual inertial odometry with UWB aided for indoor localization," *IET Communications*, vol. 16, no. 10, p. 1083–1093, June 2022.
- [14] K. Eickenhoff, P. Geneva, and G. Huang, "MIMC-VINS: A versatile and resilient multi-imu multi-camera visual-inertial navigation system," *IEEE Transactions on Robotics*, vol. 37, no. 5, pp. 1360–1380, Oct. 2021.
- [15] M. Zhang, X. Xu, Y. Chen, and M. Li, "A lightweight and accurate localization algorithm using multiple inertial measurement units," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1508–1515, Apr. 2020.
- [16] Y. Yang, P. Geneva, and G. Huang, "Multi-visual-inertial system: Analysis, calibration and estimation," *arXiv preprint arXiv:2308.05303*, 2023.
- [17] R. Zhang, F. Hoffinger, and L. M. Reind, "Calibration of an IMU using 3D rotation platform," *IEEE Sensors Journal*, vol. 14, no. 6, pp. 1778–1787, June 2014.
- [18] D. Tedaldi, A. Pretto, and E. Menegatti, "A robust and easy to implement method for IMU calibration without external equipments," in *Proc. of IEEE International Conference on Robotics and Automation*, Hong Kong, China, 2014.
- [19] Y. Yang, P. Geneva, X. Zuo, and G. Huang, "Online self-calibration for visual-inertial navigation: Models, analysis, and degeneracy," *IEEE Transactions on Robotics*, vol. 39, no. 5, pp. 3479–3498, Oct. 2023.
- [20] S. Y. Cho and C. G. Park, "A calibration technique for a redundant imu containing low-grade inertial sensors," *ETRI journal*, vol. 27, no. 4, pp. 418–426, Aug. 2005.
- [21] J. Rohac, M. Sipos, and J. Simanek, "Calibration of low-cost triaxial inertial sensors," *IEEE Instrumentation & Measurement Magazine*, vol. 18, no. 6, pp. 32–38, Dec. 2015.
- [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 *Proc. of IEEE International Conference on Robotics and Automation*, Stockholm, Sweden, 2016.
- [23] A. Jadid, L. Rudolph, F. Pankratz, and G. Klinker, "Utilizing multiple calibrated imus for enhanced mixed reality tracking," in *Proc. of IEEE International Symposium on Mixed and Augmented Reality Adjunct*, Beijing, China, 2019, pp. 384–386.
- [24] F. L. Kooi and A. Toet, *Display Systems: Binocular and 3-D, Visual Comfort*. Encyclopedia of Optical and Photonic Engineering CRC Press, Sept. 2015.
- [25] H. Self, "Optical tolerances for alignment and image differences for binocular helmet-mounted displays," Armstrong Aerospace Medical Research Laboratory, 86-019, Tech. Rep., 1986.
- [26] J. Lee, D. Hanley, and T. Bretl, "Extrinsic calibration of multiple inertial sensors from arbitrary trajectories," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 2055–2062, Apr. 2022.
- [27] Y. Yu, Y. Liu, F. Fu, S. He, D. Zhu, L. Wang, X. Zhang, and J. Li, "Fast extrinsic calibration for multiple inertial measurement units in visual-inertial system," in *Proc. of IEEE International Conference on Robotics and Automation*, ExCeL London, 2023.
- [28] J. J. Craig, *Introduction to Robotics*. Pearson Education, 2006.
- [29] B. K. Horn, H. M. Hilden, and S. Negahdaripour, "Closed-form solution of absolute orientation using orthonormal matrices," *Journal of the Optical Society*, vol. 5, no. 7, pp. 1127–1135, July 1988.
- [30] S. M. Kay, *Fundamentals of statistical signal processing: estimation theory*. Prentice-Hall, Inc., 1993.
- [31] T. Wang and S. I. Roumeliotis, "A direct algorithm for multi-gyroscope infield calibration," *arXiv preprint arXiv:2403.08177*, 2024.
- [32] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.