

On Geometric Models and Their Accuracy for Extrinsic Sensor Calibration

Kaihong Huang

Cyryll Stachniss

Abstract—Extrinsic sensor calibration is an important task in robotics. There are various ways to perform the calibration task, but it often remains unclear which methods are better than the others. In this paper, we provide a systematic study about the calibration accuracy of three types of calibration methods, each represented by an abstract geometric model based on the sensor configuration and the calibration setup. We discuss the advantages and disadvantages of each model and perform a rigorous study on their noise sensitivity from a geometric perspective. As a result, we can reveal and quantify the relative calibration accuracies of the three models, thus answering the question of “which model is better and why?”. Beside our analytical analysis, we also provide numerical simulation experiments that validate our findings.

I. INTRODUCTION

Most mobile robots perform some form of state estimation such as localization, mapping, simultaneous localization and mapping (SLAM), or exploration. For most of such tasks, it is important to know where the individual sensors are mounted on the robot. The task of determining the relative transformation between the sensors is often referred to as extrinsic calibration. Without this calibration information, a lot of the estimation tasks, especially those related to computing geometric models such as SLAM, do not work properly and/or provide suboptimal results. The extrinsic calibration is also important when the information from multiple sensors have to be fused.

There are various of ways to perform the calibration task. Some methods rely on auxiliary and dedicated markers, while others utilize only the motion information estimated from the sensors. We refer to the former methods as *marker-based* while the later ones as *motion-based*.

Both techniques have been studied for common sensors like cameras, laser-scanners, GPS receivers etc. It remains, however, unclear which methods are better and, more importantly, why they are better than the others.

The main contribution of this work is a systematic study about the calibration accuracy of three kinds of calibration methods, each represented by an abstract geometric model based on the sensor configuration and calibration setup. We discuss the advantages and disadvantages of each model and perform a rigorous study on their noise sensitivity. Theoretical bounds for their calibration accuracy are obtained

All authors are with the University of Bonn, Institute of Geodesy and Geoinformation, Bonn, Germany.

This work has partly been supported by the DFG under the grant number FOR-1015: Mapping on Demand as well as by the European Commission under the Horizon 2020 framework programme under grant agreement no. 645403. Furthermore, we received support by the China Scholarship Council CSC.

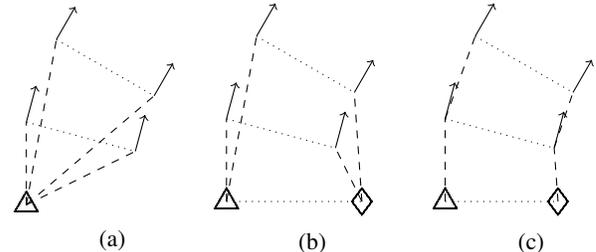


Fig. 1: Three geometric models for calibration problem of a sensor pair. Δ and \diamond refer to the reference frames of the two sensors. (a) Model A uses absolute poses in a common global frame. (b) Model B uses absolute poses in separate global frames. (c) Model C uses individual, relative motions for each sensor. The arrows represent two sensors pose estimate at different timesteps, dotted lines represent the unknown transformations T between the sensors, dashed lines represent pose/motion measurements.

and compared. As a result, we are able to report on the relative calibration accuracy of the three models and are able to answer the question of “which kind is better in which scenario and why?”. Experimental results based on simulation justify our analysis.

II. RELATED WORK

Marker-based calibration methods formed the vast majority of calibration studies. They rely on markers or known environment features and try to estimate the target parameters directly from the sensed features by maximizing a quality measure or the agreement of the sensor data with specific constraints [7], [10], [11], [17].

Typical sensors that fit in this paradigm are cameras and LiDARs. For cameras, usually a set of point correspondences matched from different views of the same scene are used. The sum of squared point re-projection errors are then serves as a cost function for an error minimization in parameter space [8], [18]. For LiDARs, objects with distinguishable shape, like flat surface or round perimeter are used. Taylor et al. [16] introduce a metric called gradient orientation measure that is applicable to different modalities. The underlying geometry of marker-based approach is straight forward and seldom discussed.

An alternative group of calibration methods is motion-based. Motion-based calibration methods are strongly related to the study of the transformation equation $AX=XB$. Early work of Shiu and Ahmad [14] motivates this equation in the context of camera-in-the-hand (aka. hand-eye) calibration.

They provide a closed form solution as well as its uniqueness condition. Following works propose various alternative closed form solutions by using, for example, angle-axis representation [13], or dual-quaternions formulation [3], or the idea of screw motion and screw axis [4]. Despite the simplicity and being fast to compute, these direct approaches do not take measurement uncertainties into full consideration, thus rendering these methods vulnerable to noise.

Strobl and Hirzinger [15] considered the measurement noise, but only as a relative weighting between rotation/translation components in the error metrics for the nonlinear optimization. Brookshire and Teller [2] provide a more formal discussion. They formulate a projected Gaussian noise model for unit dual quaternion SE(3) parameterization. By using unit dual quaternion with 8 parameters, the rotation and translation are jointly represented, including their uncertainty. Maximum likelihood optimization is then carried out under this noise model, which provides a Cramer-Rao bound for the calibrations uncertainty. The aforementioned analysis is rigorous but it covers only one model.

Beside measurement noise, existence of unobservable parameters is another practical problem occurs in motion-based approaches. This happens when the motions experienced by the sensors do not contain enough rotations in 3D, which leads to an incomplete confinement to all the six transformation parameters during the estimation. Brookshire and Teller [1] discuss the parameter observability in an algebraic way by inspecting the rank of the Fisher information matrix, while a more recent work by Maye et al. [12] brings the observability analysis into practical use. They aim to separate the calibration parameters into observable and unobservable part in real-time and update only those parameters that are observable during the optimization. They carry out the analysis in a numerical way by using rank-revealing QR and singular value decompositions of the Fisher information matrix.

Observability analysis can, however, merely state whether the parameter is observable or not. We take a step forward and perform noise sensitivity analysis using the underlying geometric models. This can not only provide the information about observability, but more importantly quantify how the sensor and trajectory configuration relates to the calibration accuracy for three general calibration models.

III. MODELS

In this paper, we consider the calibration problem between two sensors. As proven in [9], problems of that type involving multiple sensors can be formulated in a pair-wise manner without loss of generality.

Consider two sensors a and b that are rigidly attached to a robot. Each sensor is assumed to provide a noisy pose (position and orientation) estimate of itself. Depending on the sensor, the pose estimates could be incremental motions relative to past ego-centric frames of a simultaneous localization and mapping or structure from motion system, or absolute poses with respect to a fixed coordinate system. We denote relative motions as M and absolute poses as P .

More specifically,

$$P, M = \begin{bmatrix} R & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \in \text{SE}(3), \quad (1)$$

with R being a 3x3 rotation matrix and \mathbf{t}_a a 3x1 vector. Subscripts consist of $\{a, b, a', b', w, v\}$ are used to indicate the reference frames. For example, P_{wa} is sensor a 's absolute pose in the fixed coordinate system w , and $M_{b'b}$ is sensor b 's current pose with respect to its previous ego-frame b' .

Since the rotation magnitude of a pose pair plays an important role in our discussion, we denote the rotation magnitude of a rotation matrix R as θ , which is related to the angle-axis representation where a unit vector indicates the rotation axis and $\theta \in [0, \pi]$ describes the magnitude of the rotation about that axis.

Our calibration task is to estimate the relative transformation T_{ab} of sensor b in a 's frame, which contains the relative rotation, denoted as O , and displacement, denoted as ξ

$$T_{ab} = \begin{bmatrix} O_{ab} & \xi_{ab} \\ \mathbf{0} & 1 \end{bmatrix} \in \text{SE}(3). \quad (2)$$

With this transformation, a point \mathbf{p} in sensor b 's frame, which is written as ${}^b\mathbf{p}$, can then be transferred to sensor a by

$${}^a\mathbf{p} = O_{ab} {}^b\mathbf{p} + \xi_{ab}. \quad (3)$$

A. Model A

In this paper, we categorize common extrinsic calibration problems into three types, base on their underlying geometric models. The three models are illustrated in Fig. 1.

In the first model (Fig. 1a), there exists a global reference frame common to both sensors. Each sensor estimates its poses with respect to that global frame. A typical example of this setup is stereo camera calibration, where both cameras look at a checkerboard and build a common reference frame using this checkerboard. As the physical dimension of the checkerboard is known, the camera poses with respect to the checkerboard can be estimated by solving a Perspective N Point problem. Once the camera poses are provided, their relative transformation, i.e. the extrinsic parameter, can then be estimated under the formulation of Model A, which is

$$P_{wa} T_{ab} = P_{wb}, \quad (4)$$

where w is the common reference frame for both sensor a and b . The estimation of T_{ab} with this model is straight forward. A pair of poses is enough to determine the parameter and the estimation can be made explicitly from measurements, i.e. $T_{ab} = (P_{wa})^{-1} P_{wb}$. As we will see in Sec. IV and in the experiments, Model A has a good and stable estimation accuracy among the three models.

However, the setup of a reference frame common to all sensors often requires control points, known landmark, or reference objects with known geometry. Hence, we refer to Model A as marker-based approach. The requirement of a common frame is a major disadvantage as they are hard or even impossible to achieve in some cases. Consider hand-eye calibration problem for example, where the encoder of the robot arm measures nothing other than its own rotation.

Thus a direct shared measurement frame with camera is not possible. Calibration with the wheeled odometry of mobile robot has similar problems. Another generally known example is camera to camera calibration where the cameras have non-overlapping views. In this case, a single checkerboard is not sufficient and a more elaborate infrastructure is required. Beside the landmarks may be hard to setup, we also need to make sure their pose information are error free, because otherwise the estimation result will be biased and may contain systematic errors.

B. Model B

If we allow the sensors to have separated global reference frames, then we can avoid the need for setting up reference objects and thus simplify the calibration process. This relaxation leads us to the Model B, as depicted in Fig. 1b.

In this model, each sensor estimates its poses with respect to a self-defined global reference frame. The global frames are related to each other by a second unknown transformation. The pose pair together with the global frames forms a quadrilateral and the geometric relation becomes

$$P_{wa}T_{ab} = T_{wv}P_{vb}, \quad (5)$$

where v is the reference frame of sensor b . In general, we have $T_{wv} \neq T_{ab}$ and thus two unknown transformations. This model is known as $AX=YB$ in the hand-eye calibration literature.

Since for each sensor, we assume to have a sensor-based odometry or localization system, we can record its trajectory independently (except time synchronization). Therefore, control points or objects for cross reference are no longer needed. This largely simplifies the calibration process, but as we will see in Sec. IV and Fig. 4, this relaxation comes at a price of two “degeneration zones”, where the estimation accuracy will be severely undermined.

C. Model C

Both Model A and Model B utilize absolute pose measurements, but absolute poses are not always available or can be subjected to drift, i.e. measurement errors are not with zero mean. For example, wheeled odometry for ground vehicle provides more accurate relative motion information than absolute ones. And pose estimate from simultaneous localization and mapping algorithm over long trajectories are inevitable to drift often even with loop-closing. In these cases, using relative pose over small time periods is a more attractive option and leads to the use of Model C:

$$M_{a'a}T_{ab} = T_{a'b'}M_{b'b} \quad (6)$$

By definition, $T_{a'a} = T_{ab}$ and thus only one unknown transformation. Its characteristic form $AX=XB$ has been widely studied. Related works such as [4], [13] etc. prove that it requires at least two sets of poses with unparallelled rotation axes to determine a unique solution for T .

Since Model C utilized mainly motion information, we refer to it as *motion-based* method. Geometrically, Model C can be seen as a special case of Model B, but it turns out

that Model C is quite different than Model B in terms of its properties. We will see more on their comparison in Sec. IV-D.

In short, common calibration problems can be categorized into these three models. As we will show later, in general situation, Model A has a good and stable calibration accuracy regardless of the trajectory configuration. On the contrary, the accuracy of Model B and C vary greatly depending on the rotation magnitude θ along the trajectory. However, we discover that the motion-based method (Model C) can in theory have a better estimation accuracy than marker-based approach (Model A) under certain conditions.

IV. NOISE SENSITIVITY ANALYSIS

To understand the estimation accuracy difference between the models, it is necessary to analyze their sensitivity to noise. To achieve this, we rely primarily on the concept of variance propagation in statistics. For readers that are not familiar with variance propagation, we present here a short summary. A more detailed treatment can be found at Förstner and Wrobel [5] page 42-44.

Given a noisy measurement $\tilde{\mathbf{x}} \stackrel{\text{def}}{=} \boldsymbol{\mu}_x + \boldsymbol{\delta}_x$, its expectation and dispersion (variance) is

$$\mathbb{E}(\tilde{\mathbf{x}}) = \boldsymbol{\mu}_x \quad (7)$$

$$\text{ID}(\tilde{\mathbf{x}}) \stackrel{\text{def}}{=} \mathbb{E}(\boldsymbol{\delta}_x \boldsymbol{\delta}_x^T) = \boldsymbol{\Sigma}_{xx}. \quad (8)$$

An estimated random variable $\tilde{\mathbf{y}}$ computed through a nonlinear function $\mathbf{y} = \mathbf{f}(\mathbf{x})$ will approximately have the mean and variance of this form:

$$\mathbb{E}(\tilde{\mathbf{y}}) = \mathbf{f}(\boldsymbol{\mu}_x) \quad (9)$$

$$\text{ID}(\tilde{\mathbf{y}}) = \mathbf{J}\boldsymbol{\Sigma}_{xx}\mathbf{J}^T, \quad (10)$$

where $\mathbf{J} \stackrel{\text{def}}{=} \frac{\partial \mathbf{f}}{\partial \mathbf{x}}$ is the Jacobian of function \mathbf{f} evaluated at $\boldsymbol{\mu}_x$. This is due to the fact that

$$\boldsymbol{\mu}_y + \boldsymbol{\delta}_y \stackrel{\text{def}}{=} \tilde{\mathbf{y}} \stackrel{\text{def}}{=} \mathbf{f}(\tilde{\mathbf{x}}) \quad (11)$$

$$= \mathbf{f}(\boldsymbol{\mu}_x + \boldsymbol{\delta}_x) \quad (12)$$

$$\approx \mathbf{f}(\boldsymbol{\mu}_x) + \left. \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right|_{\mathbf{x}=\boldsymbol{\mu}_x} \boldsymbol{\delta}_x + O(|\boldsymbol{\delta}_x|^2) \quad (13)$$

$$\approx \boldsymbol{\mu}_y + \mathbf{J}\boldsymbol{\delta}_x. \quad (14)$$

Therefore, we have $\boldsymbol{\delta}_y = \mathbf{J}\boldsymbol{\delta}_x$ up to a first order approximation and thus the variance of $\tilde{\mathbf{x}}$ is propagated to $\tilde{\mathbf{y}}$ as

$$\boldsymbol{\Sigma}_{yy} = \mathbb{E}[(\mathbf{J}\boldsymbol{\delta}_x)(\mathbf{J}\boldsymbol{\delta}_x)^T] = \mathbf{J}\boldsymbol{\Sigma}_{xx}\mathbf{J}^T. \quad (15)$$

To analyze the three models, we first identify the nonlinear functions that relate the unknown parameters and the noisy measurements, and then obtain theoretical upper bound of the estimation accuracy by applying variance propagation.

For translation parameter $\boldsymbol{\xi}$, we analyze $\|\boldsymbol{\xi}\|$ instead of carrying out an exhaustive variance propagation for each component. The reason is that vector norms are invariant with rotations. This property allows us to have an intuitive analysis and interpretation, enabling a direct comparison between different models. For the estimation of the orientation

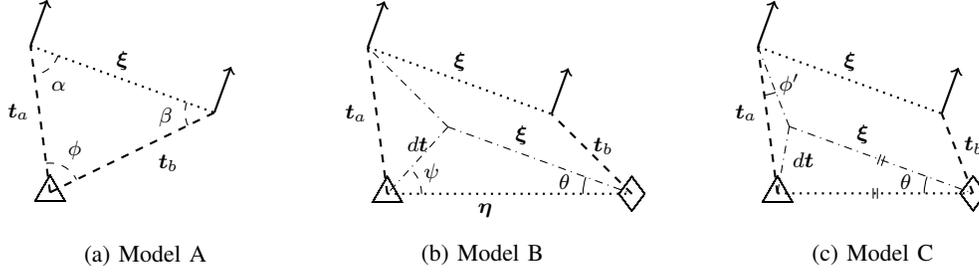


Fig. 2: The three geometric models A, B, and C. Model A uses one global frame, Model B two global frames, and Model C relies on relative motion estimates only.

parameter O , several studies exist and hence will not be covered in this paper. An in depth discussion of such topic can be found at Hartley et al. [6].

In the following discussion, we use tilde accents to denote noisy measurements (e.g. $\tilde{t}_a, \tilde{t}_b, \tilde{\theta}, \dots$) and Σ to denote the corresponding variance of additive noise. The other entities that appear in variance propagation without accents are meant to be noise free latent values. Their values depend on the the physical and spatial configuration of the sensors. As we will see, they play an important rules in determining the calibration accuracy.

A. Model A

We start with the analysis of Model A. Eq. (4) can be expanded to

$$O_{ab} = R_a^T R_b \quad (16)$$

$$\xi_{ab} = R_a^T (t_b - t_a). \quad (17)$$

From Eq. (17) or Fig. 2a, we know that

$$\|\xi\| = \|t_b - t_a\| \quad (18)$$

$$= \sqrt{\|t_b\|^2 + \|t_a\|^2 - 2\|t_a\|\|t_b\|\cos\phi}. \quad (19)$$

Hence, the estimated $\|\tilde{\xi}\|$ is

$$\|\tilde{\xi}\| = \sqrt{\|\tilde{t}_b\|^2 + \|\tilde{t}_a\|^2 - 2\|\tilde{t}_a\|\|\tilde{t}_b\|\cos\phi}, \quad (20)$$

where $\phi \stackrel{\text{def}}{=} \angle(t_a, t_b)$ is the angle between vector t_a and t_b (see Fig. 2a). The uncertainty of $\|\xi\|$ is obtained by applying variance propagation to Eq. (20) and reads to

$$\Sigma_{\|\xi\|}^A = \left[\frac{\|t_a\| - \|t_b\|\cos\phi}{\|\xi\|} \right]^2 \Sigma_{\|t_a\|} + \left[\frac{\|t_b\| - \|t_a\|\cos\phi}{\|\xi\|} \right]^2 \Sigma_{\|t_b\|} + \mathcal{O}(\Sigma_\phi), \quad (21)$$

assuming the noise of $\|\tilde{t}_a\|$ and $\|\tilde{t}_b\|$ are uncorrelated.

Consider the other two angles $\alpha \stackrel{\text{def}}{=} \angle(\xi, -t_a)$ and $\beta \stackrel{\text{def}}{=} \angle(\xi, t_b)$ within the vector triangle, we obtain

$$\|t_a\| = \cos\phi\|t_b\| + \cos\alpha\|\xi\| \quad (22)$$

$$\|t_b\| = \cos\phi\|t_a\| + \cos\beta\|\xi\|, \quad (23)$$

and then Eq. (21) can be simplified to

$$\Sigma_{\|\xi\|}^A = \cos^2\alpha \Sigma_{\|t_a\|} + \cos^2\beta \Sigma_{\|t_b\|}. \quad (24)$$

We can conclude that the more ξ is perpendicular to t_a, t_b , the less sensitive is ξ to noise. $\Sigma_{\|\xi\|}$ is bounded by

$$\Sigma_{\|\xi\|}^A \max = \Sigma_{\|t_a\|} + \Sigma_{\|t_b\|} \quad (25)$$

when $(\alpha, \beta, \phi) = \{(\pi, 0, 0), (0, \pi, 0), (0, 0, \pi)\}$, i.e. t_a, t_b being collinear.

Computing the lower bound for $\Sigma_{\|\xi\|}$ is not as straight forward and it depends on the ratio of $\Sigma_{\|t_a\|}$ and $\Sigma_{\|t_b\|}$. To give a rough idea, we can assume $\Sigma_{\|t_a\|} = \Sigma_{\|t_b\|}$, then

$$\Sigma_{\|\xi\|}^A \min = \frac{1}{2} \frac{\|\xi\|^2}{\|t_a\|^2} \Sigma_{\|t_a\|} \quad (26)$$

which is the case if $\alpha = \beta$. This means the relative uncertainty $r_\xi \stackrel{\text{def}}{=} \frac{\Sigma_{\|\xi\|}}{\|\xi\|^2}$ of the estimation is only a half of $r_t \stackrel{\text{def}}{=} \frac{\Sigma_{\|t_a\|}}{\|t_a\|^2}$ of the measurement.

For example, assume $\|\xi\| = 20 \text{ cm}$, $\|t_a\| \approx 1 \text{ m}$, $\|t_b\|$ varies but $\Sigma_{\|t_a\|} = \Sigma_{\|t_b\|} = 1 \text{ cm}^2$, then the ξ estimated by Model A will have a standard deviation of $0.014 \text{ cm} \sim 1.4 \text{ cm}$ in its length.

B. Model C

We discuss Model C first and leave Model B for last, because the former one is simpler to start with. For Model C, Eq. (6) can be expanded to

$$R_a O_{ab} = O_{ab} R_b \quad (27)$$

$$R_a \xi_{ab} + t_a = O_{ab} t_b + \xi_{ab}. \quad (28)$$

Two pose pairs form a vector quadrilateral. Our target is to find the extrinsic parameters (ξ, O) that ‘‘closes’’ the quadrilateral. Eq. (28) in its original form does not provide much clues for this task. Therefore we introduce an intermediate entity dt and rewrite Eq. (28) as

$$t_a - O_{ab} t_b \stackrel{\text{def}}{=} dt = \xi_{ab} - R_a \xi_{ab}. \quad (29)$$

Fig. 2c illustrated the motivation. Assuming a nonzero rotation and shifting the upper pose pair to the lower pair, we can transform the quadrilateral into two triangles that share one side. These two triangles respectively correspond to the left and right term of Eq. (29). The shared side dt is the translation difference between the sensors a and b , as indicated by the left term¹. On the other hand, the right term $(\xi - R\xi)$ forms an isosceles triangle, which relates the

unknown entity ξ to its two equal sides, with the included angle being the rotation magnitude θ of R .

From this isosceles triangle, we obtain the relation

$$\|dt\| = 2 \sin(\theta/2) \|\xi\|. \quad (30)$$

If $\theta = 0$, Eq. (30) still hold because $\sin(\theta/2) = 0$ and $\|dt\| = 0$, but ξ is no longer unique and could take any values given fixed \tilde{t}_a and \tilde{t}_b .

If $\theta \neq 0$, then $\|\tilde{\xi}\|$ can be estimated from

$$\|\tilde{\xi}\| = \frac{1}{2 \sin(\tilde{\theta}/2)} \|\tilde{dt}\|. \quad (31)$$

By applying variance propagation to Eq. (31), the uncertainty of $\|\tilde{\xi}\|$ turns into

$$\Sigma_{\|\tilde{\xi}\|}^C = \left[\frac{\cos(\theta/2)}{4 \sin^2(\theta/2)} \|dt\| \right]^2 \Sigma_\theta + \left[\frac{1}{2 \sin(\theta/2)} \right]^2 \Sigma_{\|dt\|} \quad (32)$$

or

$$\Sigma_{\|\xi\|}^C = \left[\frac{\cos(\theta/2)}{2 \sin(\theta/2)} \|\xi\| \right]^2 \Sigma_\theta + \left[\frac{1}{2 \sin(\theta/2)} \right]^2 \Sigma_{\|dt\|} \quad (33)$$

due to $\|dt\| = 2 \sin(\theta/2) \|\xi\|$.

Furthermore, as \tilde{dt} is estimated from \tilde{t}_a and \tilde{t}_b with

$$\tilde{dt} = \tilde{t}_a - O_{ab} \tilde{t}_b, \quad (34)$$

we obtain $\Sigma_{\|\tilde{dt}\|}$ similar to Eq. (21) in Model A:

$$\Sigma_{\|\tilde{dt}\|} = \frac{1}{4 \sin^2(\theta/2)} \left(\left[\frac{\|\tilde{t}_a\| - \|\tilde{t}_b\| \cos \phi'}{\|\tilde{\xi}\|} \right]^2 \Sigma_{\|\tilde{t}_a\|} + \left[\frac{\|\tilde{t}_b\| - \|\tilde{t}_a\| \cos \phi'}{\|\tilde{\xi}\|} \right]^2 \Sigma_{\|\tilde{t}_b\|} \right), \quad (35)$$

but with $\phi' \stackrel{\text{def}}{=} \angle(\tilde{t}_a, O_{ab} \tilde{t}_b)$. The leading factor $\frac{1}{4 \sin^2(\theta/2)}$ of Eq. (35) is also due to $\|dt\| = 2 \sin(\theta/2) \|\xi\|$.

From Eq. (32), we can see that $\Sigma_{\|\tilde{\xi}\|}^C$ consists of two parts. One part is the translational ($\|\tilde{dt}\|$) and the other is the rotational one ($\tilde{\theta}$). The relative angle θ of the pose pairs plays an important role in both parts. In situations where θ takes small values (e.g. $\theta < 10^\circ$), the factor $\frac{1}{\sin(\theta/2)}$ (and its power) will be large, meaning any noise in the measurements will be significantly amplified. We refer to this as a ‘‘degeneration zone’’. In extreme cases around $\theta = 0$, the uncertainty (or variance) approaches infinity, because the solution is not unique and can take any values.

In contract to that, if we have a good configuration with a large θ around 180° , the influence of measurement noise will be reduced. The minimum $\Sigma_{\|\xi\|}$ is attained for $\theta = 180^\circ$ with

$$\Sigma_{\|\xi\|}^C \min = \frac{1}{16} \left(\left[\frac{\|\tilde{t}_a\| - \|\tilde{t}_b\| \cos \phi'}{\|\tilde{\xi}\|} \right]^2 \Sigma_{\|\tilde{t}_a\|} + \left[\frac{\|\tilde{t}_b\| - \|\tilde{t}_a\| \cos \phi'}{\|\tilde{\xi}\|} \right]^2 \Sigma_{\|\tilde{t}_b\|} \right). \quad (36)$$

Our simulation experiment in Sec. V confirms this bound.

¹In case the robot pivots around sensor b (i.e. $\tilde{t}_b = 0$), then ξ and O are completely decoupled, and $dt = \tilde{t}_a$ is solely the translation measurement of sensor a .

C. Model B

The analysis of Model B is similar to Model C. Eq. (5) can be expanded to

$$R_{wa} O_{ab} = \tilde{O}_{wv} R_{vb} \quad (37)$$

$$R_{wa} \xi_{ab} + t_{wa} = \tilde{O}_{wv} t_{vb} + \eta_{wv}, \quad (38)$$

where both η_{wv} and ξ_{ab} are unknown translation parameters. We rewrite Eq. (38) as

$$t_{wa} - \tilde{O}_{wv} t_{vb} \stackrel{\text{def}}{=} dt = \eta_{wv} - R_{wa} \xi_{ab}. \quad (39)$$

In this model, the triangle that relates ξ and dt is no longer isosceles and an extra unknown side η is present, as depicted in Fig. 2b. The geometric relation becomes

$$\sin \theta \|\xi\| = \sin \psi \|dt\| \quad (40)$$

with $\psi \stackrel{\text{def}}{=} \angle(\eta, dt)$.

If $\sin \theta = 0$, (i.e. $\theta = 0$ or π , meaning ξ and η are collinear), we have $\sin \psi = 0$. The equation still hold but the system becomes degenerated in the sense that solution to $\tilde{\xi}$ is not unique (so does $\tilde{\eta}$ because $\tilde{\eta} = \tilde{dt} + R\tilde{\xi}$).

Assuming $\sin \theta \neq 0$, the length of $\tilde{\xi}$ relates to $\|\tilde{dt}\|$, $\tilde{\psi}$ and $\tilde{\theta}$ with

$$\|\tilde{\xi}\| = \frac{\sin \tilde{\psi}}{\sin \tilde{\theta}} \|\tilde{dt}\|. \quad (41)$$

By applying variance propagation and omitting the correlation terms, the uncertainty of $\|\tilde{\xi}\|$ becomes

$$\Sigma_{\|\tilde{\xi}\|}^B = \left[\frac{\cos \theta \sin \psi}{\sin^2 \theta} \|dt\| \right]^2 \Sigma_\theta + \left[\frac{\cos \psi}{\sin \theta} \|dt\| \right]^2 \Sigma_\psi + \left[\frac{\sin \psi}{\sin \theta} \right]^2 \Sigma_{\|dt\|} \quad (42)$$

and can be written as

$$\Sigma_{\|\xi\|}^B = \left[\frac{\cos \theta}{\sin \theta} \|\xi\| \right]^2 \Sigma_\theta + \left[\frac{\cos \psi}{\sin \psi} \|\xi\| \right]^2 \Sigma_\psi + \left[\frac{\sin \psi}{\sin \theta} \right]^2 \Sigma_{\|dt\|} \quad (43)$$

due to $\|\xi\| = \frac{\sin \psi}{\sin \theta} \|dt\|$.

The uncertainty $\Sigma_{\|\xi\|}^B$ consists of one translation and two rotation parts. We can identify the two degeneration zones around $\theta = 0$ and $\theta = \pi$ from Eq. (43), which are due to the squared cotangent factors of Σ_ψ and Σ_θ .

The translation part is not as straight forward as in Model C, but we can show that, assuming $\|\eta\| > \|\xi\|$, the factor $\frac{\sin \psi}{\sin \theta}$ is bounded by

$$\frac{\|\xi\|}{\|\eta\| + \|\xi\|} < \frac{\sin \psi}{\sin \theta} < \frac{\|\xi\|}{\|\eta\| - \|\xi\|}, \quad (44)$$

because of Eq. (40) and $\|\eta\| - \|\xi\| < \|dt\| < \|\eta\| + \|\xi\|$. From the perspective of Eq. (44), we can in theory increase the calibration accuracy by making $\|\eta\|$ larger, e.g. separating the global frames further away from each other.

TABLE I: Overview of $\Sigma_{\|\xi\|}$ for the three models.

Model	$\Sigma_{\ \xi\ }$	# of degeneration zones
A (AX=B)	$[2 \sin(\theta/2)]^2 \Sigma_{\ d\mathbf{t}\ }$	0
C (AX=XB)	$\left[\frac{\cos(\theta/2)}{2 \sin(\theta/2)} \ \xi\ \right]^2 \Sigma_\theta + \left[\frac{1}{2 \sin(\theta/2)}\right]^2 \Sigma_{\ d\mathbf{t}\ }$	1
B (AX=YB)	$\left[\frac{\cos \psi}{\sin \psi} \ \xi\ \right]^2 \Sigma_\psi + \left[\frac{\cos \theta}{\sin \theta} \ \xi\ \right]^2 \Sigma_\theta + \left[\frac{\sin \psi}{\sin \theta}\right]^2 \Sigma_{\ d\mathbf{t}\ }$	2

D. Comparison

Given Eq. (21), Eq. (33) and Eq. (43), see also Tab. I, we are now able to compare the three models. For this comparison, we assume the same trajectory and measurement noise.

First, we compare Model B and Model C. From the geometry perspective, Model C can be seen as a special case of Model B with $\boldsymbol{\eta} = \boldsymbol{\xi}$. In this case, we can exploit this equality for the comparison of the models, as it implies

$$\psi = (\pi - \theta)/2. \quad (45)$$

This results in

$$\frac{\sin \psi}{\sin \theta} = \frac{\sin(\pi/2 - \theta/2)}{\sin \theta} = \frac{\cos(\theta/2)}{2 \sin(\theta/2) \cos(\theta/2)} = \frac{1}{2 \sin(\theta/2)}. \quad (46)$$

Given that, the second term of Eq. (33) and the third term of Eq. (43) are equal. Additionally exploiting $\cot \theta = \frac{\cos \theta}{\sin \theta}$ and assuming $\Sigma_\psi = \frac{1}{4} \Sigma_\theta$, we simplify $\Sigma_{\|\xi\|}^B - \Sigma_{\|\xi\|}^C$ for our comparison as follows:

$$\begin{aligned} \Sigma_{\|\xi\|}^B - \Sigma_{\|\xi\|}^C &= \|\xi\|^2 \left[\Sigma_\theta \cot^2 \theta + \Sigma_\psi \cot^2 \psi - \frac{\Sigma_\theta}{4} \cot^2 \frac{\theta}{2} \right] \quad (47) \\ &= \|\xi\|^2 \left[\Sigma_\theta \cot^2 \theta + \frac{\Sigma_\theta}{4} \tan^2 \frac{\theta}{2} - \frac{\Sigma_\theta}{4} \cot^2 \frac{\theta}{2} \right] \quad (48) \\ &= \Sigma_\theta \|\xi\|^2 \underbrace{\left[\cot^2 \theta + \frac{1}{4} \tan^2 \frac{\theta}{2} - \frac{1}{4} \cot^2 \frac{\theta}{2} \right]}_{g(\theta)}. \quad (49) \end{aligned}$$

The term $g(\theta)$ in Eq. (49) provides us the insight that, under which circumstances Model C is better than Model B, or vice versa. We have

$$\begin{cases} g(\theta) < 0 & \text{if } \theta < \pi/2 \\ g(\theta) = 0 & \text{if } \theta = \pi/2 \\ g(\theta) > 0 & \text{otherwise.} \end{cases} \quad (50)$$

See also Fig. 3 for a plot of this term. The term $g(\theta)$ is smaller than zero (between -1/2 and 0) for $\theta \in [0^\circ, 90^\circ)$, meaning Model B is (slightly) better than Model C in that range. For $\theta \in [90^\circ, 180^\circ]$, this term is larger than zero and even approaches infinity, such that Model C is substantially better than Model B here.

Second, we compare Model A to C. Eq. (21) shows that $\Sigma_{\|\xi\|}^A$ is independent of θ , which is an advantage of Model A over Model C and B. Because the terms related to θ can lead to large uncertainty or degeneration zones as we saw before.

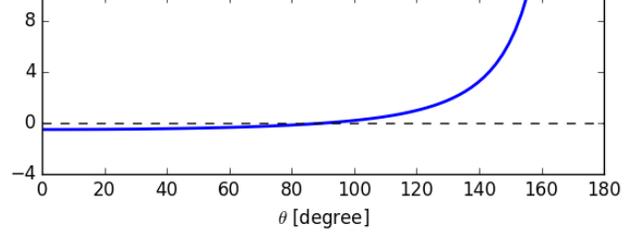


Fig. 3: Illustration of the function $g(\theta)$ in Eq. (49) depending on the value of θ . As can be seen, for $\theta < 90^\circ$, Model B is slightly better than Model C, but for value larger than 90° , Model B degenerate quickly. In contrast, Model C performs well (compare also Fig. 4).

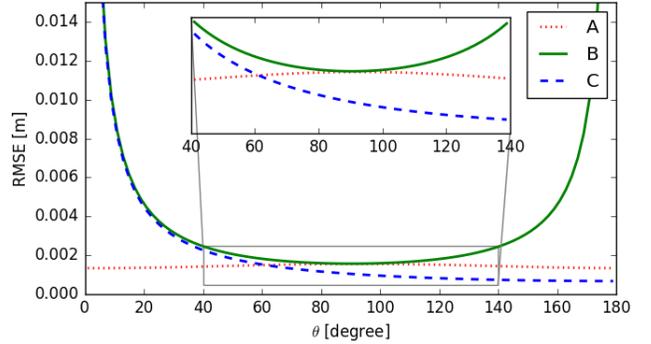


Fig. 4: Result from a simulation experiment for calibrating two sensors with all three models along the same trajectories. The plot shows the RMSE of the three models depending on θ with a close-up view for the interval $\theta \in [40^\circ, 140^\circ]$. As can be seen, Model A performs well, independent of θ . Model C is better than Model A for $\theta > 60^\circ$. Model B never outperforms Model and has practically no advantages over Model C.

By comparing $\Sigma_{\|\xi\|}^A$ to the theoretical minimum value of $\Sigma_{\|\xi\|}^C$ from Eq. (36), we can see that Model C can in theory outperform Model A. As we will, however, observe in the experimental evaluation, Model A is less sensitive to degenerate cases and thus, should be preferred over Model B and Model C in most practical situation. Only for large values of $\theta > 60^\circ$, Model C is better than Model A. Furthermore, Model B never outperforms Model A.

V. EXPERIMENTAL EVALUATION

To validate our analysis, we conduct Monte Carlo simulations. We generate trajectories with controlled configuration parameters for $\|t\|, \theta$, etc. We add zero mean Gaussian noise to the simulated pose measurements and estimate the extrinsic parameters O, ξ with a nonlinear least-squares approach developed in our previous work [9]. This estimation approach is based on the Gauss-Helmert paradigm and able to provide a statistical optimal solution up to the Cramer-Rao bound. Meanwhile, all solutions are initialized with ground-truth values to rule out possible effects of local minimas. In the end, we compute the error as the difference in the length of the vectors ξ and $\tilde{\xi}$, i.e., $\|\xi\| - \|\tilde{\xi}\|$, and calculate the root mean square (RMSE) for each model. The common setup for the simulations are $\|\eta\| = \|\xi\| = 1\text{ m}$, $\|\tilde{t}_a\| = 10\text{ m}$, and noise variance are set to $\Sigma_{\|t_a\|} = \Sigma_{\|t_b\|} = (0.01\text{ m})^2$, $\Sigma_\theta = (0.001\text{ rad})^2$.

We generated 1000 trials/trajectories per value of θ and each trajectory consists of 100 poses, all evaluated for the three models. The result of our Monte Carlo simulations is depicted in Fig. 4. It shows the RMSE for each model with varying values of θ .

For Model A, the RMSE plot is almost straight and with minimal variations as expected. The curve of Model B is dominated by the shape of the squared cotangent function. We can visually identify the two degeneration zones around $\theta = 0^\circ$ and $\theta = 180^\circ$, which are due to the squared cotangent factor of Σ_ψ and Σ_θ , see Eq. (43) in Sec. IV-C. For Model C, the RMSE plot shows almost no difference to that of Model B for values $\theta < 40^\circ$, but it is strictly decreasing and has only one degeneration zone around $\theta = 0$.

Comparing the three models, we see the situation changes as the value of θ increases. Model A's near constant performance remains the best until θ reaches 60° , then Model C becomes the best and eventually its RMSE takes only half of Model A's RMSE when $\theta = 180^\circ$. Model B performs the worst among the three models in this experiment, only in a narrow range (from 80° to 100°) it is close to Model A.

The noticeable deviation of Model B's and C's curve begins at $\theta = 40^\circ$ instead of expected $\theta = 90^\circ$. A possible explanation is because the correlation terms we omitted in Eq. (43) also have a substantial influence on $\Sigma_{\|\xi\|}^B$. But this does not change our conclusion about the accuracy comparison.

To sum up, the results of the simulation experiment support our analytical analysis and suggest that Model A is the preferred solution. For controlled cases with $\theta > 60^\circ$, one can consider Model C instead. The use of Model B however should be avoided.

VI. CONCLUSION

In this paper, we presented a systematic study about the calibration accuracy of three kinds of calibration methods, namely $AX=B$, $AX=XB$ and $AX=YB$. We discussed the advantage and disadvantage of each model and perform a rigorous study on their noise sensitivities. We showed how the sensor configuration and calibration setup influence the

calibration accuracies and answered the question of "which model is better?". Contradict to the common conception that marker-base methods are always superior, we showed that in some cases the motion-based methods can be better than marker-base approaches. In summary, we conclude that if the calibration setup allow for using Model A, i.e., the marker-based approach, it is a good choice and should be used. For controlled settings with $\theta > 60^\circ$, one should also consider Model C as it can provide better estimate of the parameters and typically requires less calibration infrastructure. If Model A cannot be applied, Model C is more appropriate than Model B, but for small values of θ , both degenerate.

REFERENCES

- [1] J. Brookshire and S. Teller. Automatic calibration of multiple coplanar sensors. In *Proc. of Robotics: Science and Systems (RSS)*, volume 33. MIT Press, 2011.
- [2] J. Brookshire and S. Teller. Extrinsic Calibration from Per-Sensor Egomotion. In *Proc. of Robotics: Science and Systems (RSS)*, 2012.
- [3] K. Daniilidis. Hand-eye calibration using dual quaternions. *Intl. Journal of Robotics Research (IJRR)*, 18(3):286–298, 1999.
- [4] I. Fassi and G. Legnani. Hand to sensor calibration: A geometrical interpretation of the matrix equation $AX=XB$. *Journal of Robotic Systems*, 22(9):497–506, 2005.
- [5] W. Förstner and B. Wrobel. *Photogrammetric Computer Vision – Statistics, Geometry, Orientation and Reconstruction*. Springer Verlag, 2016.
- [6] R. Hartley, J. Trumpf, Dai Y, and H. Li. Rotation averaging. *Intl. Journal of Computer Vision (IJCV)*, 103(3):267–305, 2013.
- [7] L. Heng, B. Li, and M. Pollefeys. Camodocal: Automatic intrinsic and extrinsic calibration of a rig with multiple generic cameras and odometry. In *Proc. of the IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2013.
- [8] R. Horaud, G. Csurka, and D. Demirdijian. Stereo calibration from rigid motions. *IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)*, 22(12):1446–1452, Dec 2000.
- [9] K.H. Huang and C. Stachniss. Extrinsic Multi-Sensor Calibration For Mobile Robots Using the Gauss-Helmert Model. In *Proc. of the IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2017.
- [10] R. Kümmerle, G. Grisetti, and W. Burgard. Simultaneous calibration, localization, and mapping. In *Proc. of the IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, pages 3716–3721, 2011.
- [11] P. Lbraly, E. Royer, O. Ait-Aider, C. Deymier, and M. Dhome. Fast calibration of embedded non-overlapping cameras. In *Proc. of the IEEE Intl. Conf. on Robotics & Automation (ICRA)*, pages 221–227, May 2011.
- [12] J. Maye, H. Sommer, G. Agamennoni, R. Siegwart, and P. Furgale. Online self-calibration for robotic systems. *Intl. Journal of Robotics Research (IJRR)*, 35(4):357–380, 2016.
- [13] F.C. Park and B.J. Martin. Robot sensor calibration: solving $AX=XB$ on the Euclidean group. *IEEE Trans. on Robotics and Automation*, 10(5):717–721, Oct 1994.
- [14] Y. C. Shiu and S. Ahmad. Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form $AX=XB$. *IEEE Trans. on Robotics and Automation*, 5(1):16–29, Feb 1989.
- [15] K. H. Strobl and G. Hirzinger. Optimal Hand-Eye Calibration. In *Proc. of the IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, pages 4647–4653, Oct 2006.
- [16] Z. Taylor, J. Nieto, and D. Johnson. Multi-modal sensor calibration using a gradient orientation measure. *Journal of Field Robotics (JFR)*, 32(5):675–695, 2015.
- [17] C. H. Tong and T. D. Barfoot. A self-calibrating 3d ground-truth localization system using retroreflective landmarks. In *Proc. of the IEEE Intl. Conf. on Robotics & Automation (ICRA)*, pages 3601–3606, May 2011.
- [18] A. Zisserman, P. A. Beardsley, and I. D. Reid. Metric calibration of a stereo rig. In *Proc. of the IEEE Workshop on Representation of Visual Scenes*, pages 93–100, Jun 1995.