Multiscale and Uncertainty-aware Targetless Hand-Eye Calibration via the Gauss-Helmert Model

Marta Čolaković-Bencerić, Juraj Peršić, Ivan Marković, Ivan Petrović

***Abstract*—The operational reliability of an autonomous robot depends crucially on extrinsic sensor calibration as a prerequisite for precise and accurate data fusion. Exploring the calibration of unscaled sensors (e.g., monocular cameras) and the effective utilization of uncertainties are difficult and often overlooked. The development of a solution for the simultaneous calibration of hand-eye sensors and scale estimation based on the Gauss-Helmert model aims to utilize the valuable information contained in the uncertainty of odometry. In this work, we propose a versatile and robust solution for batch calibration based on the analytical on-manifold approach for estimation. The versatility of our method is demonstrated by its ability to calibrate multiple unscaled and metric-scaled sensors while dealing with odometry failures and reinitializations. Importantly, all estimated parameters are provided with their corresponding uncertainties. The validation of our method and its comparison with five competing state-of-the-art calibration methods in both simulations and real-world experiments show its superior accuracy, with particularly promising results observed in high-noise scenarios.**

***Index Terms*—extrinsic calibration, multisensor systems, autonomous robots, Gauss-Helmert model.**

I. Introduction

MOBILE robotic autonomy hinges on accurate perception of the environment. To achieve a robust perception pipeline, mobile robotic systems often feature a heterogeneous sensor suite. Extrinsic calibration is tasked with determining spatial relations between the coordinate systems associated with various sensors and robot elements to facilitate sensor fusion. However, achieving accurate calibration is challenging, especially in dynamic or outdoor environments, and is further complicated by the use of diverse sensors like cameras, lidars, radars, and IMUs, each with different sensing modalities and noise characteristics.

Extrinsic calibration methods vary based on their data correspondence strategy, using either manufactured targets or environmental information. The use of targets characterizes target-based methods, [1]–[3], which lack generality in terms of sensor use and require an engineered target for their operation. This effectively prevents the operator from altering the sensor suite without significant adjustments to the calibration procedure. In contrast to targetless calibration methods, target-based methods find limited use in continuous decalibration monitoring and recalibration during robot deployment, especially in outdoor environments.

Marta Čolaković-Bencerić, Ivan Marković and Ivan Petrović are with the University of Zagreb Faculty of Electrical Engineering and Computing, Laboratory for Autonomous Systems and Mobile Robotics, Croatia. E-mail: marta.colakovic@fer.hr, ivan.markovic@fer.hr, ivan.petrovic@fer.hr

Juraj Peršić is with Calirad, Croatia. E-mail: juraj.persic@calirad.net

Targetless calibration methods can be subdivided into feature-based and odometry-based approaches. Feature-based calibration uses features extracted from observations of the natural environment to find data correspondences, [4]–[7], while odometry-based extrinsic calibration leverages sensor odometry, [8]–[10]. While feature-based methods are often tailored to specific sensor combinations, making calibration of large multisensor systems difficult [4], odometry-based methods make fewer assumptions about the environment, often requiring only temporal synchronization [8], [11] or metrically scaled 6DoF motion [12]. Odometry-based calibration offers flexibility for use with a wide range of sensors but typically suffers from lower accuracy and increased sensitivity to errors in scale estimation.

One common formulation of the odometry-based calibration problem is the *hand-eye* problem, expressed by the equation:

$$\mathbf{AX} = \mathbf{XB}, \tag{1}$$

where $\mathbf{A}$ and $\mathbf{B}$ denote sensors' relative motion, and $\mathbf{X}$ the rigid transformation between their coordinate frames. Although this formulation originated in robotic manipulator calibration [13], it can be applied to any sensor that can determine its ego-motion, including 3D lidars, stereo cameras, GNSS-INS devices, and wheel odometers. The hand-eye problem can be framed as a standalone procedure [14] or as part of a larger calibration pipeline [9], [15], [16]. Unfortunately, calibrating unscaled sensors, namely monocular cameras, in the context of odometry-based calibration is a significant challenge due to the inherent limitations of these sensors in capturing depth information without the use of engineered targets. Unlike metrically scaled sensors, such as lidars or stereo cameras, monocular cameras only provide relative motion in the projective space, lacking the necessary scale factor of the translational component of motion. Since cameras are among the most commonly used sensors in robotics due to their affordability and versatility, the issue of scale estimation remains a relevant issue in the field of odometry-based robotic calibration.

Given the complexity of multisensor systems and a demand for flexible calibration procedures, we focus on odometry-based calibration, despite odometry-based calibration often suffering from lower accuracy than target or feature-based calibration methods. Therefore, effective calibration methods must account for uncertainty in odometry data by incorporating robust weighting strategies and uncertainty-aware optimization techniques. Our method addresses these issues using uncertainty-aware optimization and the Gauss-Helmert

model, ensuring robust calibration despite high noise and scale ambiguities. Our key scientific contributions are as follows:

- A multiscale, multisensor calibration framework based on the Gauss-Helmert model, supporting both scaled and unscaled sensors with asynchronous changes in translational scale.
- Analytical derivation of all optimization components, including closed-form Jacobians for robust and efficient optimization.
- Guaranteed positive scale factor estimates, enhancing convergence properties in noisy environments.
- Uncertainty-driven weighting, improving calibration accuracy and estimating both calibration parameters and scale factor uncertainty.
- Uncertainty-aware trajectory averaging through model-constrained, uncertainty-weighted corrections.

These contributions provide a noise-robust solution to the hand-eye calibration problem, effectively addressing scale ambiguity and measurement uncertainty, and are broadly applicable to diverse sensors by relying exclusively on ego-motion data.

The remainder of this paper is organized as follows: Section II reviews existing solutions to the hand-eye calibration problem. Section III introduces the unscaled hand-eye calibration problem, details the necessary mathematical background, and presents an adaptable multisensor calibration method based on the Gauss-Helmert model. Section IV demonstrates the method's performance on a 3D simulated dataset and compares it to five state-of-the-art solutions. The capabilities of the proposed calibration method are further validated on real-world data in Sec. V, while Sec. VI concludes our work.

II. Related work

Early solutions to the hand-eye problem relied on the on-manifold *SO(3)* representation [13], [17], [18] or on quaternion representation of rotation [19]. Most of these early solutions feature decoupled hand-eye calibration through closed-form solutions [13], [17], [19], with few moving towards iterative procedures, [20]. Decoupled solutions ignore the non-linear coupling of the translational and rotational components, further propagating the rotational error to the translational calibration parameter estimate, as noted by [20]. Subsequent simultaneous hand-eye solutions gave rise to the use of dual quaternion representation of poses, [12], [21], [22]. In [21] the solution was found through the used of Singular Value Decomposition (SVD), while iterative approaches, such as in [12], [20], use a non-linear least-squares solver. Later, in [23] and [22] the authors obtain a globally optimal solution through the Lagrangian dual problem and semidefinite programming (SDP) of quadratically constrained quadratic programs (QCQPs) using quaternion and dual-quaternion representation, respectively.

However, very few hand-eye calibration solutions include a systematic treatment of measurement uncertainty. One of these solutions, presented in [24], uses the Gauss-Helmert (GH) model to formulate the optimization problem on the $SO(3)\times\mathbb{R}^3$ group. The GH model assumes imperfect measurements, correcting them along with the calibration parameters during the optimization procedure. The proposed solution, while effective, does not extend to a fully analytical formulation and is limited to the calibration of metrically scaled sensors only. In [25]–[27] the authors formulate the hand-eye problem on the *SE(3)* group by describing the batch ego-motion measurements as distributions using their associated mean and covariance. A closed-form batch solution that does not require known data correspondences is found in [25] by assuming highly focused ego-motion distribution functions. Continuing, in [26], the authors improve on the batch method by formulating the hand-eye problem as a minimization problem of the Kullback-Leibler divergence. Finally, in [27] the authors alleviate the restrictions on data distribution tightness and significantly improve algorithm performance. While the solutions proposed in [25]–[27] utilize a probabilistic description of measurement data, they do not estimate the uncertainty of the resulting calibration parameters. A disciplined approach to extrinsic hand-eye calibration and uncertainty estimation is presented in [28]. The authors formulated the problem on the $SO(3)\times\mathbb{R}^3$ group and jointly estimated the calibration parameters and their uncertainty using the tools of on-manifold differentiation and uncertainty propagation.

All of the hand-eye self-calibration solutions discussed above assume metrically scaled motion. Unfortunately, monocular camera measurements lack depth information, making it impossible to discern the translational scale of the camera's motion without the use of engineered targets. Early work on self-calibration that includes scale factor estimation, solved the unscaled hand-eye problem by using homogeneous matrices for rotation estimation formulated using the Kronecker product, followed by estimation of translational calibration parameters and scale, [29]. Another decoupled method was presented in [30], where authors formulate a second-order cone programming (SOCP) estimator for translational parameters and scale, with given rotational calibration parameters. A similar solution was proposed in [31], where the sensor ego-motion was represented by dual-quaternions, with scale and translational parameters estimated after the rotational component of calibration, using a SOCP-based algorithm. Further decoupled solutions to the unscaled hand-eye problem were presented in [9] and [15] as a part of a wider calibration framework. Odometry failure and the possibility of multiple scale factors within one sequence were first addressed in [16], with a simultaneous estimator specifically designed for 2D sequences. Multiple scale factors have been introduced to account for possible odometry failure in low texture environments or during quick rotation. Joint estimation of extrinsic calibration parameters for 3D motion and scale is first featured in [32]. The authors formulated two non-linear cost functions, one relying on the homogeneous matrix formulation and the other on dual-quaternion representation of data, both jointly estimating all six calibration parameters and a single scale factor. More recently, in [33] the authors expanded the method proposed in [23] and included scale estimation in the QCQP formulation of the hand-eye self-calibration problem, again, proposing a globally optimal solution. Similarly, in [34] we see the solution proposed in [22] expanded to include multiple scale factors per sensor.

The GH model was developed and investigated for geodesic and scan matching applications, [35], [36] and, finally, stereo-camera calibration [37], with odometry-based self-calibration remaining woefully under-examined. In [38] a GH-based estimator and outlier removal scheme was proposed as a precursor to feature-based calibration. Unfortunately, the calibration scheme accommodates only two sensors and a single scale factor estimate per dataset, with neither the performance nor derivation of the method thoroughly demonstrated. Additionally, the method does not ensure positive scale, which can affect the convergence properties of the calibration procedure. In an effort to address calibration uncertainty of vision-guided robots, Ulrich and Hillemann developed an estimator based on the Gauss-Markov (GM) model with added fictitious observations, [39], [40]. In a self-calibration scenario, the GM-based estimator minimizes a cost function that combines reprojection errors with errors from fictitious robot pose observations. Furthermore, the approach is limited to vision-based sensor systems and does not provide the flexibility that a purely odometry-based calibration offers. In addition, self-calibration significantly increases the number of reconstructed 3D points, which leads to a significant increase in execution time. In contrast to the GH-based estimator, the extended GM-based estimator does not enforce any constraints, which makes it unsuitable for averaging trajectories.

III. Proposed method

In this section, we propose an odometry-based unscaled calibration method. The unscaled hand-eye calibration problem is presented in Subsec. III-A and the necessary mathematical background and the on-manifold calculus toolset in Subsec. III-B. Then we describe the chosen optimization procedure in Subsec. III-C and propose the analytical multiscale multisensor calibration solution in Subsec. III-D.

A. Problem formulation

The objective of odometry-based extrinsic calibration is to estimate the spatial relationship between coordinate systems associated with different sensors of a robotic system. This versatile class of calibration methods only requires the ability to determine the sensor's ego-motion. Most previously presented odometry-based methods assume that the calibrated sensors are metrically scaled and, thus, estimate only the rigid transformation between their coordinate systems. It follows that the aforementioned family of calibration methods is capable of calibrating sensors like lidars and stereo-cameras, critically neglecting a very common robotic sensor, the monocular camera. Monocular camera odometry can only be estimated up to the scale factor of its translational component. Therefore, we propose an extended hand-eye problem formulation that jointly estimates extrinsic calibration parameters and multiple scale factors per sensor trajectory. The extended hand-eye formulation retains the familiar form (1), where $\mathbf{A}$ and $\mathbf{B}$ represent sensors' relative metrically unscaled motion calculated from the estimated odometry. Both $\mathbf{A}$ and $\mathbf{B}$ are represented in an ego-centric frame of the associated sensor at the previous timestep. $\mathbf{X}$ then represents a similarity transformation connecting these motion estimates. Using the groundwork set by [24], we propose a fully analytical multiscale multisensor solution to the unscaled hand-eye problem. Multiscale estimation allows for odometry-based calibration even in sequences with odometry failure. Additionally, it allows us to link short odometry sequences, as long as the data was collected with unchanged extrinsic calibration, into one larger dataset to circumvent data sparsity that can occur in challenging or asynchronous sequences. Via the methods of on-manifold calculus, we have analytically derived all necessary components of the used optimization scheme based on the GH model. Finally, through careful consideration of measurement uncertainty, we have illustrated the importance of utilizing this often untapped source of information.

B. Mathematical background

Metrically scaled ego-motion, i.e., rigid transformations, are represented by $\mathbf{T} \in \textit{SE(3)}$ – an object composed of a rotation matrix $\mathbf{R} \in \textit{SO(3)}$ and a translation vector $\mathbf{t} \in \mathbb{R}^3$,

$$\textit{SE(3)} : \mathbf{T} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix} \\ \mathbf{T} \in \textit{SE(3)} \subset \mathbb{R}^{4\times4} \mid \mathbf{R} \in \textit{SO(3)}, \mathbf{t} \in \mathbb{R}^3. \quad (2)$$

Rigid transformations as members of the *SE(3)* group make up a smooth differential manifold, which allows for straightforward analytical differentiation. Critically, they can be expressed by an element of the vector space isomorphic to the tangent Lie algebra of the *SO(3)*, $\mathbf{r} = \mathrm{Log}(\mathbf{R}) \in \mathbb{R}^3 \cong \mathfrak{so}(3) = \textit{TSO(3)}$, and a vector $\mathbf{t} \in \mathbb{R}^3$. The group $SO(3)$ represents the group of all 3×3 orthogonal matrices with determinant 1, which corresponds to the group of rotations in three-dimensional space. The Lie algebra of $SO(3)$, denoted as $\mathfrak{so}(3)$ is the set of 3×3 skew-symmetric matrices, which can be mapped to the space of infinitesimal rotations. Retaining the on-manifold representation of the rotational component of a rigid transform preserves the analytical differentiability, all the while maintaining clarity of representation when transitioning to unscaled motion. The logarithmic mapping, Log, connects elements of the group to the elements of a vector space isomorphic to the tangent Lie algebra, while the exponential mapping, Exp, does the inverse. The aforementioned *SO(3)* logarithmic and exponential mapping are defined by

$$\mathrm{Log} : \mathbf{R} \longrightarrow \mathbf{r}, \quad (3)$$

$$\mathrm{Exp} : \mathbf{r} \longrightarrow \mathbf{R}, \quad (4)$$

$$\mathbf{R} \in \textit{SO(3)}, \mathbf{r} \in \mathbb{R}^3 \cong \mathfrak{so}(3) = \textit{TSO(3)}$$

and equal to

$$\mathrm{Log}(\mathbf{R}) = \frac{r(\mathbf{R} - \mathbf{R}^\top)^\vee}{2\sin(r)}, \quad (5)$$

$$\mathrm{Exp}(\mathbf{r}) = \exp(\mathbf{r}^\wedge) = \sum_{n=0}^{\infty} \frac{1}{n!}(\mathbf{r}^\wedge)^n, \quad (6)$$

where

$$\mathbf{r}^\wedge = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}^\wedge = \begin{bmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{bmatrix} \in \mathfrak{so}(3), \quad (7)$$

$(\mathbf{r}^\wedge)^\vee = \mathbf{r}$ and $r = |\mathbf{r}|$, [41].

Two other important operators in the toolset of on-manifold calculus are the left *SO(3)* group Jacobian and its inverse, i.e. $\mathbf{J}_l(\mathbf{r})$ and $\mathbf{J}_l^{-1}(\mathbf{r})$. In general, the left group Jacobian maps small perturbations of the parameter vector to the corresponding perturbations in the global Lie algebra. For the rotation group $SO(3)$, the left Jacobian $\mathbf{J}_l(\mathbf{r})$ is equal to

$$\mathbf{J}_l(r) = \mathbf{I}_3 + \frac{1-\cos(r)}{r^2}\mathbf{r}^\wedge + \frac{r-\sin(r)}{r^3}\mathbf{r}^{\wedge 2}, \tag{8}$$

where $\mathbf{r}^\wedge \in \mathfrak{so}(3)$ and $r = |\mathbf{r}|$. The inverse of the left Jacobian $\mathbf{J}_l(\mathbf{r})$ is given by

$$\mathbf{J}_l^{-1}(\mathbf{r}) = \mathbf{I}_3 - \frac{1}{2}\mathbf{r}^\wedge + \left(\frac{1}{r^2} - \frac{1+\cos(r)}{2r\sin(r)}\right)\mathbf{r}^{\wedge 2}. \tag{9}$$

We define the incrementation strategy on the *SO(3)* group as multiplicative and establish an additive equivalent

$$\delta\mathbf{R}\cdot\mathbf{R} = \mathrm{Exp}(\delta\mathbf{r})\cdot\mathrm{Exp}(\mathbf{r}) \simeq \mathrm{Exp}\left(\mathbf{r} + \mathbf{J}_l^{-1}(\mathbf{r})\delta\mathbf{r}\right). \tag{10}$$

As a result, we are able to maintain simplicity of analytical differentiation and solution expressions, while taking into account rotational multiplicity. Finally, we can combine the incrementation of the rotation parameters $\mathbf{r}$ on the manifold and the vector addition of the translation parameters $\mathbf{t}$ into a joint operator $\boxplus$ that describes the perturbation

$$\begin{bmatrix} \Delta\mathbf{t} \\ \Delta\mathbf{r} \end{bmatrix} \boxplus \begin{bmatrix} \mathbf{t} \\ \mathbf{r} \end{bmatrix} = \begin{bmatrix} \Delta\mathbf{t}+\mathbf{t} \\ \mathrm{Log}(\mathrm{Exp}(\Delta\mathbf{r})\mathrm{Exp}(\mathbf{r})) \end{bmatrix} \tag{11}$$

$$\simeq \begin{bmatrix} \Delta\mathbf{t}+\mathbf{t} \\ \mathbf{r}+\mathbf{J}_l^{-1}(\mathbf{r})\delta\mathbf{r} \end{bmatrix}. \tag{12}$$

In the composite space $\langle\mathbb{R}^3, \mathrm{SO}(3)\rangle$, rotations and translations do not interact and can only be use to model perturbations, perform differentiations and calculate covariances [42].

Information gathered by certain sensors, namely monocular cameras in an unmapped environment, cannot be used to reliably infer scale. A classical way to expand rigid transform representation to include scale is to expand the $\langle\mathbb{R}^3, \mathrm{SO}(3)\rangle$ group with the group of scale factors, $s \in \mathbb{R}$. Unscaled ego-motion or similarity transforms are then represented by

$$\begin{aligned} \textit{Sim(3)} : \mathbf{T}^{Sim3} &= \begin{bmatrix} s\mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \\ &\mathbf{T} \in \textit{Sim(3)} \subset \mathbb{R}^{4\times 4} \\ &\mathbf{R} \in \textit{SO(3)},\ \mathbf{t} \in \mathbb{R}^3,\ s \in \mathbb{R}, \end{aligned} \tag{13}$$

where s signifies the scale factor. Similarly, the vector representation becomes $\mathbf{v} = \begin{bmatrix} \mathbf{t} & \mathbf{r} & s \end{bmatrix}^T$, where $\mathbf{t}$ denotes translation, $\mathbf{r}$ denotes rotational parameters and s denotes the scale factor. For the sake of simplicity, the perturbation operator adopts the notation of the perturbation of the rigid transformation and is equal to

$$\begin{bmatrix} \Delta\mathbf{t} \\ \Delta\mathbf{r} \\ \Delta s \end{bmatrix} \boxplus \begin{bmatrix} \mathbf{t} \\ \mathbf{r} \\ s \end{bmatrix} = \begin{bmatrix} \Delta\mathbf{t}+\mathbf{t} \\ \mathrm{Log}(\mathrm{Exp}(\Delta\mathbf{r})\mathrm{Exp}(\mathbf{r})) \\ \Delta s + s \end{bmatrix}. \tag{14}$$

C. Gauss-Helmert optimization

In real-world systems and data, measurement uncertainty is an ever-present problem. Many least-squares (LSQ) estimation solutions tend to overlook the critical issue of appropriate weighting, often neglecting available covariance information in favor of simpler heuristic approaches. To fully leverage the available uncertainty data, we implement an iterative procedure akin to total-least-squares regression, utilizing the Gauss-Helmert (GH) model with a covariance-based weighting scheme. Within the GH model framework we estimate the sought parameters, while simultaneously correcting the observations. By employing uncertainty information, if available, to correct observations, one can significantly mitigate the impact of noise present in the data.

The sensor system is assumed to consist of m sensors, where sensor A is designated as the 'central' sensor, and the remaining sensors are labeled B through M. For a system with m sensors and a dataset comprising odometry segments S_j for each sensor j ($j = B, \ldots, M$), the unknown calibration parameters are denoted as follows

$$\boldsymbol{\xi} = \begin{bmatrix} \mathbf{t}_B & \mathbf{r}_B & s_1^B & \ldots & s_{S_B}^B & \ldots & \mathbf{t}_M & \mathbf{r}_M & s_1^M & \ldots & s_{S_M}^M \end{bmatrix}^\top. \tag{15}$$

Translational and rotational components of the similarity transformation between coordinate frames of sensor A and any other sensor j are denoted by $\mathbf{t}_j$ and $\mathbf{r}_j$. Associated scale factors are denoted by $s_1^j \ldots s_{S_j}^j$. Finally, temporally synchronized observations are represented as

$$\mathbf{l}_i = \begin{bmatrix} \mathbf{t}_{Ai} & \mathbf{r}_{Ai} & \mathbf{t}_{Bi} & \mathbf{r}_{Bi} & \ldots & \mathbf{t}_{Mi} & \mathbf{r}_{Mi} \end{bmatrix}^\top \tag{16}$$

where $i = 1, \ldots, N$, with N representing the number of odometry data points. Here, $\mathbf{t}_{Ai}$ denotes the translational component of the ego-motion of sensor A during the time interval $t_{i+1} - t_i$, and $\mathbf{r}_{Ai}$ represents the corresponding rotational component. Similarly, $\mathbf{t}_{Bi}$ and $\mathbf{r}_{Bi}$ denote the translational and rotational components of ego-motion for sensor B, and so forth for other sensors M. All ego-motion measurements are given in the ego-frame of the sensor at the previous timestep. The relationship between sensor poses at times t_i and t_{i+1}, the ego-motion measurements and the transformation between then is demonstrated in Fig. 1. The total corrections to observations at k-th iteration are defined as $\boldsymbol{\epsilon}_i^{(k)} = \Delta\mathbf{l}_i^{(k)} \boxplus \Delta\mathbf{l}_i^{(k-1)} \boxplus \ldots \Delta\mathbf{l}_i^{(0)}$ and equal to

$$\boldsymbol{\epsilon}_i = \begin{bmatrix} \boldsymbol{\epsilon}_{\mathbf{t}Ai} & \boldsymbol{\epsilon}_{\mathbf{r}Ai} & \ldots & \boldsymbol{\epsilon}_{\mathbf{t}Mi} & \boldsymbol{\epsilon}_{\mathbf{r}Mi} \end{bmatrix}^\top. \tag{17}$$

The unknown parameters are estimated by minimizing the sum of squares of observation corrections $\boldsymbol{\epsilon}_i$ weighted by the inverse of their covariance matrix $\Sigma_{\mathbf{ll}_i}^{-1}$, while strictly enforcing constraints $\mathbf{g}_i(\boldsymbol{\xi}, \mathbf{l}_i) = \mathbf{0}$ informed by the problem at hand. The constraint equations, derived from the unscaled hand-eye problem formulation, are formally defined in Sec. III-D with Eq. (29). Additionally, to avoid physically impossible solutions

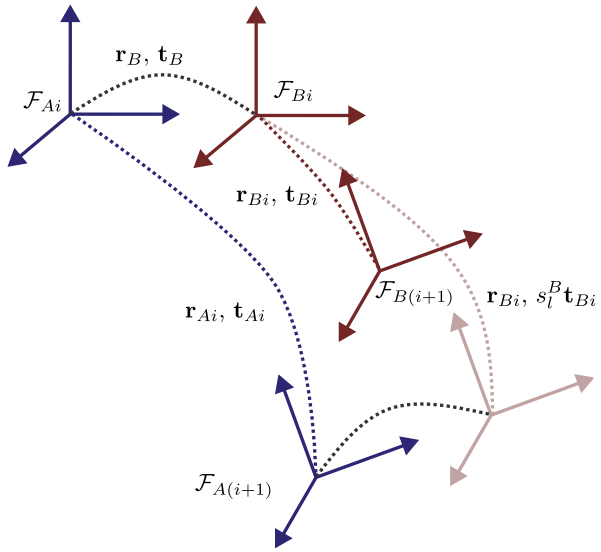

Fig. 1. The proposed method calculates the transformation between frames of sensors A and B, denoted by $\mathbf{r}_B$, $\mathbf{t}_B$, and the scale factor of that odometry segment s_l^B. The measurements or ego-motions are symbolized by $\mathbf{r}_{Ai}$, $\mathbf{t}_{Ai}$ and $\mathbf{r}_{Bi}$, $\mathbf{t}_{Bi}$. The ego-motions denote the relative motion of sensors between timestamps $i+1$ and i.

($s \leq 0$) and allow for a larger basin of convergence, we maintain a strictly positive scale factor.

$$\begin{aligned} \underset{\Delta\boldsymbol{\xi},\, \Delta\mathbf{l}_i}{\operatorname{argmin}} & \frac{1}{2} \sum_{i=1}^{N} \boldsymbol{\epsilon}_i^\top \Sigma_{\mathbf{ll}\,i}^{-1} \boldsymbol{\epsilon}_i \\ \text{s.t. } & \mathbf{g}_i(\boldsymbol{\xi}, \mathbf{l}_i) = \mathbf{0}, \\ & s > 0. \end{aligned} \tag{18}$$

To find the solution we performed a procedure similar to classic derivations described in [36] and [38]. By expressing the total correction to observations at the $(k+1)$-th iteration as $\boldsymbol{\epsilon}_i^{(k+1)} = \Delta\mathbf{l}_i^{(k)} \oplus\!\!\!\!\ominus \boldsymbol{\epsilon}_i^{(k)}$ we reformulate the minimization problem as

$$\begin{aligned} \underset{\Delta\boldsymbol{\xi},\, \Delta\mathbf{l}_i}{\operatorname{argmin}} & \frac{1}{2} \sum_{i=1}^{N} (\Delta\mathbf{l}_i^{(k)} \oplus\!\!\!\!\ominus \boldsymbol{\epsilon}_i^{(k)})^\top \Sigma_{\mathbf{ll}\,i}^{-1} (\Delta\mathbf{l}_i^{(k)} \oplus\!\!\!\!\ominus \boldsymbol{\epsilon}_i^{(k)}) \\ \text{s.t. } & \mathbf{g}_i(\Delta\boldsymbol{\xi}^{(k)} \oplus\!\!\!\!\ominus \boldsymbol{\xi}^{(k)}, \Delta\mathbf{l}_i^{(k)} \oplus\!\!\!\!\ominus \mathbf{l}_i^{(k)}) = 0, \\ & \Delta s_l + s_l^{(k)} > 0 \quad s_l = 1, \ldots, S_j. \end{aligned} \tag{19}$$

Here $\Delta\mathbf{l}_i^{(k)}$ represents optimal correction to observation i estimated in the k-th iteration. These corrections allow us to calculate the updated measurements $\mathbf{l}_i^{(k+1)} = \Delta\mathbf{l}_i^{(k)} \oplus\!\!\!\!\ominus \mathbf{l}_i^k$ and the updated total accumulated corrections to observations $\boldsymbol{\epsilon}_i^{(k+1)}$. The hidden assumption used in defining the estimation problem is that the ego-motions as measurements are uncorrelated at any timestep.

By exploiting the equivalence (10) we reformulate $\Delta\mathbf{l}_i^{(k)} \oplus\!\!\!\!\ominus \boldsymbol{\epsilon}_i^{(k)}$ as $\mathbf{P}_i^{(k)} \Delta\mathbf{l}_i^{(k)} + \boldsymbol{\epsilon}_i^{(k)}$, with

$$\mathbf{P}_i^{(k)} = \operatorname{diag}\begin{pmatrix} \mathbf{I}_3 & \mathbf{J}_l^{-1}(\boldsymbol{\epsilon}_{\mathbf{r}Ai}^{(k)}) & \mathbf{I}_3 & \mathbf{J}_l^{-1}(\boldsymbol{\epsilon}_{\mathbf{r}Bi}^{(k)}) & \ldots \end{pmatrix} \tag{20}$$

and

$$\mathbf{P}_i^{-1(k)} = \operatorname{diag}\begin{pmatrix} \mathbf{I}_3 & \mathbf{J}_l(\boldsymbol{\epsilon}_{\mathbf{r}Ai}^{(k)}) & \mathbf{I}_3 & \mathbf{J}_l(\boldsymbol{\epsilon}_{\mathbf{r}Bi}^{(k)}) & \ldots \end{pmatrix}. \tag{21}$$

The constraints are then linearized at the k-th iteration

$$\mathbf{Y}_i^{(k)} \Delta\boldsymbol{\xi}^{(k)} + \mathbf{Z}_i^{(k)} \Delta\mathbf{l}_i^{(k)} + \mathbf{g}_i(\boldsymbol{\xi}^{(k)}, \mathbf{l}_i^{(k)}) = 0, \tag{22}$$

with $\mathbf{Y}_i^{(k)}$ and $\mathbf{Z}_i^{(k)}$ denoting Jacobians of the constraints with respect to the unknown parameters and the observations evaluated at the current estimate. For brevity, the constraint function evaluated at the current estimate $\mathbf{g}_i(\boldsymbol{\xi}^{(k)}, \mathbf{l}_i^{(k)})$, is denoted as $\mathbf{g}_i^{(k)}$ in the remainder of the manuscript. Finally, we use the Karush–Kuhn–Tucker (KKT) multipliers [43] to formulate the Lagrange function,

$$\begin{aligned} \mathcal{L} = & \frac{1}{2} \sum_{i=1}^{N} (\mathbf{P}_i^{(k)} \Delta\mathbf{l}_i^{(k)} + \boldsymbol{\epsilon}_i^{(k)})^\top \Sigma_{\mathbf{ll}\,i}^{-1} (\mathbf{P}_i^{(k)} \Delta\mathbf{l}_i^{(k)} + \boldsymbol{\epsilon}_i^{(k)}) \\ & + \sum_{i=1}^{N} \mathbf{L}_i^\top \left(\mathbf{Y}_i^{(k)} \Delta\boldsymbol{\xi}^{(k)} + \mathbf{Z}_i^{(k)} \Delta\mathbf{l}_i^{(k)} + \mathbf{g}_i^{(k)} \right) \\ & + \sum_{j=B}^{M} \sum_{l=0}^{S_j} K_l^j \left(\Delta s_l^{j(k)} + s_l^{j(k)} - \beta_l^{j2} \right), \end{aligned} \tag{23}$$

where the KKT multipliers that ensure strict adherence to the hand-eye constraints, defined formally in Sec. III-D via (29), are denoted by $\mathbf{L}_i^\top$. To ensure a positive scale, we introduced the strictly positive slackness term β_l^{j2} and the multiplier K_l^j, with the multiplier being active only when $s_l^j \leq 0$. Conversely, when $s_l^j > 0$ parameter K_l^j equals 0. By setting the derivatives of $\mathcal{L}$ to zero and utilizing, now linear, constraints, we reach the optimal corrections to the calibration parameters and observations

$$\begin{aligned} \Delta\boldsymbol{\xi}^{(k)} = & \left(\sum_{i=1}^{N} \mathbf{Y}_i^{(k)\top} \mathbf{W}_i \mathbf{Y}_i^{(k)} \right)^{-1} \\ & \cdot \left(\sum_{i=1}^{N} \mathbf{Y}_i^{(k)\top} \mathbf{W}_i \mathbf{c}_i^{(k)} - \sum_{j=B}^{M} \sum_{l=1}^{S_j} \begin{pmatrix} 0 \\ \vdots \\ K_l^j \\ \vdots \end{pmatrix} \right), \end{aligned} \tag{24}$$

$$\begin{aligned} \Delta\mathbf{l}_i^{(k)} = & \mathbf{P}_i^{-1(k)} \Big(\Sigma_{\mathbf{ll}\,i} \left(\mathbf{Z}_i^{(k)} \mathbf{P}_i^{-1(k)} \right)^\top \mathbf{W}_i \\ & \cdot \left(\mathbf{c}_i^{(k)} - \mathbf{Y}_i^{(k)} \Delta\boldsymbol{\xi}^{(k)} \right) - \boldsymbol{\epsilon}_i^{(k)} \Big). \end{aligned} \tag{25}$$

In (24) and (25) $\mathbf{W}_i$ represents the weighting matrices equal to $\left(\mathbf{Z}_i^{(k)} \mathbf{P}_i^{-1(k)} \Sigma_{\mathbf{ll}\,i} \left(\mathbf{Z}_i^{(k)} \mathbf{P}_i^{-1(k)} \right)^\top \right)^{-1}$. Vector $\mathbf{c}_i$ equals $\mathbf{c}_i^{(k)} = \mathbf{Z}_i^{(k)} \mathbf{P}_i^{-1(k)} \boldsymbol{\epsilon}_i^{(k)} - \mathbf{g}_i^{(k)}$, while

$$\begin{aligned} K_l^j = & \Sigma_{\boldsymbol{\xi}\boldsymbol{\xi}\; s_l^j, s_l^j}^{-1} \Bigg(-\beta_l^{j2} + s_l^{j(k)} \\ & + \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \end{pmatrix}^\top \Sigma_{\boldsymbol{\xi}\boldsymbol{\xi}} \sum_{i=1}^{N} \mathbf{Y}_i^{(k)\top} \mathbf{W}_i \mathbf{c}_i^{(k)} \Bigg), \end{aligned} \tag{26}$$

with

$$\Sigma_{\boldsymbol{\xi\xi}} = \left(\sum_{i=1}^{N} \mathbf{Y}_i^{(k)\top} \mathbf{W}_i \mathbf{Y}_i^{(k)} \right)^{-1}. \tag{27}$$

The symbol $\Sigma^{-1}_{\boldsymbol{\xi\xi} s_l^j, s_l^j}$ denotes the diagonal entry of the $\Sigma^{-1}_{\boldsymbol{\xi\xi}}$ matrix corresponding to the precision of the s_l^j scale parameter. In the proposed solution we have heuristically set $\beta_l^{j\,2}$ to $-s_l^{j\,(k)}$. That ensures that in the iteration following a change of scale into negative values $s_l^{j\,(k+1)} \leftarrow |s_l^{j\,(k)}|$. To reach the theoretical precision at convergence of estimated parameters we calculate the *a posteriori* variance factor

$$\hat{\sigma}_0^2 = \frac{\sum_{i=1}^{N} \boldsymbol{\epsilon}_i^\top \Sigma_{\mathbf{ll}\,i}^{-1} \boldsymbol{\epsilon}_i}{6 \cdot N \cdot (m-1) - \sum_{j=B}^{M} (6 + S_j)}, \tag{28}$$

where m denotes the number of sensors and S_j the number of estimated scale factors associated with sensor j. The covariance at convergence then equals $\hat{\sigma}_0^2 \Sigma_{\boldsymbol{\xi\xi}}$. Variance components of ego-motion measurements of multi-sensor systems can be estimated using least-squares variance component estimation (LS-VCE) [44], if a sufficient number of independent constraints are given, [45].

D. Multiscale multisensor solution

Assuming a multisensor system comprised of any number of monocular cameras and metrically scaled sensors, we present an analytically derived iterative non-linear LSQ solution. We adopted the Gauss-Helmert optimization paradigm, previously presented in [24], and expanded it by including scale estimation in the hand-eye calibration framework. Moreover, we have analytically derived all necessary Jacobians used in the final calibration solution (24) and measurement correction (25).

In real-world applications, odometry estimation often fails due to a lack of visual texture or rapid rotational movements. As a consequence of odometry failures, the continuous data collected by a sensor can result in several discontinuous segments. To avoid calibration errors due to limited data, it is useful to collate these segments by including multiple scale factors into the estimation procedure, i.e., a separate scale factor for each segment. The sensors are connected in a star-shaped formation, where the set of calibration parameters $\mathbf{r}_j$, $\mathbf{t}_j$ and the scale factor s_k^j define the similarity transformation between the estimated ego-motion of sensors A and j for a particular odometry segment denoted by k. If a system contains a correctly scaled sensor it is set as the reference sensor A. Any additional metrically scaled sensor is included by omitting the associated scale variable from the calibration vector $\boldsymbol{\xi}$. As it is unlikely that the scale factor will change drastically within a continuous odometry segment, we do not assign a separate scale factor variable to each data point to avoid overfitting. The number of scale factor variables per sensor is determined by the number of discontinuous odometry segments. It is assumed that odometry failure of each of the sensors and, therefore, the time of failure are previously detected, either during odometry estimation or post-processing, which lies outside of the scope of this paper. Finally, we modify the familiar form, [24], of the hand-eye problem constraints by including the scale variables associated with different odometry segments the datapoint belongs to

$$\mathbf{g}_i(\boldsymbol{\xi}, \mathbf{l}_i) = \begin{bmatrix} (\mathbf{R}_{Ai} - \mathbf{I}_3)\mathbf{t}_B - s_k^B \mathbf{R}_B \mathbf{t}_{Bi} + \mathbf{t}_{Ai} \\ \mathbf{r}_{Ai} - \mathbf{R}_B \mathbf{r}_{Bi} \\ \vdots \\ (\mathbf{R}_{Ai} - \mathbf{I}_3)\mathbf{t}_M - s_l^M \mathbf{R}_M \mathbf{t}_{Mi} + \mathbf{t}_{Ai} \\ \mathbf{r}_{Ai} - \mathbf{R}_M \mathbf{r}_{Mi} \end{bmatrix}. \tag{29}$$

In order to find the optimal corrections to the calibration parameters $\Delta\boldsymbol{\xi}$ and observations $\Delta\mathbf{l}_i$, we needed to calculate the Jacobians of the constraints with respect to the unknown parameters and observations, namely $\mathbf{Y}_i$ and $\mathbf{Z}_i$. As previously stated, vectors describing rotation $\mathbf{r}$ are incremented as described in (10), the translational calibration parameters are incremented as simple $\mathbb{R}^3$ vectors $\mathbf{t} \rightarrow \mathbf{t} + \delta\mathbf{t}$, while the scale factor is incremented as a scalar quantity, $s \rightarrow s + \delta s$. We utilized the toolset of on-manifold differentiation to calculate analytical partial Jacobians of the constraints $\mathbf{g}_i$ in Appendix A. The Jacobian of the proposed constraints with regards to the unknown parameters was found to be equal to (30), where $\delta_{k'k}$ denotes the Kronecker delta function. Similarly, the Jacobian of constraints with respect to the observations equals (31), where $\mathbf{J}_l$ is the aforementioned left *SO(3)* group Jacobian, and $\mathbf{J}_l^{-1}$ its inverse.

E. Ego-motion uncertainty

The presented GH-based calibration solution hinges on the accurate knowledge of measurement uncertainty. It is sometimes overlooked that the uncertainty provided with odometry data is not the required measurement uncertainty. The observations of the hand-eye problem are, in fact, ego-motions. To reach the correct expression for ego-motion covariance matrices $\Sigma_{\mathbf{ll}}$ we need to apply forward propagation of uncertainty, given by

$$\Sigma_{\mathbf{f}} = \boldsymbol{\mathcal{J}} \Sigma_{\mathbf{x}} \boldsymbol{\mathcal{J}}^\top, \tag{32}$$

with $\mathbf{f}$ being the function of $\mathbf{x}$ and $\boldsymbol{\mathcal{J}} = \dfrac{\partial \mathbf{f}}{\partial \mathbf{x}}$.

Ego-motion, or change of pose, of sensor A between time instances t_i and t_{i+1} is defined as

$$\mathbf{A}_i = (\mathbf{M}_{t,i}^A)^{-1} \mathbf{M}_{t,i+1}^A, \quad \mathrm{Log}(\mathbf{A}_i) = \mathbf{l}_i^A, \tag{33}$$

where $\mathbf{M}_{t,i}^A$ denotes the pose of sensor A at time instance t_i. Given the chosen pose representation, $\mathbf{x}_i^A \in$ *TSO(3)* $\times \mathbb{R}^3$, and pose uncertainty, $\Sigma_{\mathbf{x},i}$, we express ego-motion $\mathbf{l}_i^A$ as a function of $\mathbf{x}_i^A$ and $\mathbf{x}_{i+1}^A$

$$\mathbf{l}_i^A = \begin{bmatrix} \mathbf{R}_i^\top (\mathbf{t}_{i+1} - \mathbf{t}_i) \\ \mathrm{Log}(\mathbf{R}_i^\top \mathbf{R}_{i+1}) \end{bmatrix}. \tag{34}$$

Finally, by employing (32) we calculate the ego-motion covariance matrix

$$\begin{aligned} \Sigma_{\mathbf{ll}\,i^A} = & \frac{\partial \mathbf{l}_i^A}{\partial \mathbf{x}_i^A} \Sigma_{\mathbf{x},i} \left(\frac{\partial \mathbf{l}_i^A}{\partial \mathbf{x}_i^A} \right)^\top \\ & + \frac{\partial \mathbf{l}_i^A}{\partial \mathbf{x}_{i+1}^A} \Sigma_{\mathbf{x},i+1} \left(\frac{\partial \mathbf{l}_i^A}{\partial \mathbf{x}_{i+1}^A} \right)^\top. \end{aligned} \tag{35}$$

$$\mathbf{Y}_i = \frac{\partial \mathbf{g}(\boldsymbol{\xi}, \mathbf{l}_i)}{\partial \boldsymbol{\xi}}$$
$$= \begin{bmatrix} \mathbf{R}_{Ai} - \mathbf{I}_3 & s_k^B(\mathbf{R}_B\mathbf{t}_{Bi})^\wedge & \mathbf{0}_{3\times1} & \cdots & -\delta_{k'k}\mathbf{R}_B\mathbf{t}_{Bi} & \cdots & & & & & & \\ \mathbf{0}_{3\times3} & (\mathbf{R}_B\mathbf{r}_{Bi})^\wedge & \mathbf{0}_{3\times1} & \cdots & \mathbf{0}_{3\times1} & \cdots & & & & & & \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times1} & \cdots & \mathbf{0}_{3\times1} & \cdots & \mathbf{R}_{Ai} - \mathbf{I}_3 & s_l^C(\mathbf{R}_C\mathbf{t}_{Ci})^\wedge & \mathbf{0}_{3\times1} & \cdots & -\delta_{l'l}\mathbf{R}_C\mathbf{t}_{Ci} & \cdots \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times1} & \cdots & \mathbf{0}_{3\times1} & \cdots & \mathbf{0}_{3\times3} & (\mathbf{R}_c\mathbf{r}_{Ci})^\wedge & \mathbf{0}_{3\times1} & \cdots & \mathbf{0}_{3\times1} & \cdots \\ \vdots & \ddots & & & & & & & & & & \end{bmatrix} \tag{30}$$

$$\mathbf{Z}_i = \frac{\partial \mathbf{g}(\boldsymbol{\xi}, \mathbf{l}_i)}{\partial \mathbf{l}_i}$$
$$= \begin{bmatrix} \mathbf{I}_3 & -(\mathbf{R}_{Ai}\mathbf{t}_B)^\wedge & -s_k^B\mathbf{R}_B & \mathbf{0}_{3\times3} & \cdots & & \\ \mathbf{0}_{3\times3} & \mathbf{J}_l^{-1}(\mathbf{r}_{Ai}) & \mathbf{0}_{3\times3} & -\mathbf{R}_B\mathbf{J}_l^{-1}(\mathbf{r}_{Bi}) & \mathbf{0}_{3\times3} & \cdots & \\ \mathbf{I}_3 & -(\mathbf{R}_{Ai}\mathbf{t}_C)^\wedge & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & -s_l^C\mathbf{R}_C & \mathbf{0}_{3\times3} & \cdots \\ \mathbf{0}_{3\times3} & \mathbf{J}_l^{-1}(\mathbf{r}_{Ai}) & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & -\mathbf{R}_C\mathbf{J}_l^{-1}(\mathbf{r}_{Ci}) & \mathbf{0}_{3\times3} & \cdots \\ & \cdots & & & & & & \\ \vdots & \ddots & & & & & & \end{bmatrix} \tag{31}$$

The analytic derivation of the necessary Jacobian blocks is presented in the Appendix, with full Jacobians being equal to

$$\frac{\partial \mathbf{l}_i^A}{\partial \mathbf{x}_i^A} = \begin{bmatrix} -\mathbf{R}_i^\top & \mathbf{R}_i^\top(\mathbf{t}_{i+1} - \mathbf{t}_i)^\wedge \\ \mathbf{0}_{3\times3} & -\mathbf{J}_l^{-1}(\mathrm{Log}(\mathbf{R}_i^\top\mathbf{R}_{i+1}))\mathbf{R}_i^\top \end{bmatrix}, \tag{36}$$

$$\frac{\partial \mathbf{l}_i^A}{\partial \mathbf{x}_{i+1}^A} = \begin{bmatrix} \mathbf{R}_i^\top & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{J}_l^{-1}(\mathrm{Log}(\mathbf{R}_i^\top\mathbf{R}_{i+1}))\mathbf{R}_i^\top \end{bmatrix}. \tag{37}$$

F. GH-based algorithm for hand-eye calibration

In this section, we present the proposed GH-based algorithm for hand-eye calibration outlined in Algorithm 1 and visually represented in Fig. 2. The proposed algorithm is formulated as a minimization problem, where the corrections to observations are minimized, hand-eye constraints are strictly enforced, and scale factors are kept positive. The required input consists of synchronized ego-motion measurements, associated uncertainties, and timestamps of odometry failure indicating a change of relative scale. To explore the robustness of the algorithm to initial conditions and demonstrate its convergence properties, we initialize the calibration parameters at zero and the scale parameters at one. However, in practical applications, it is advisable to initialize these parameters using a linear solution or a reasonable estimate. After initialization, the main loop of the algorithm is repeated until calibration and scale parameter convergence.

The main loop of the algorithm consists of three major components. Firstly, Jacobians (30) and (31), auxiliary quantities $\mathbf{c}_i$ and $\mathbf{P}_i$, and weights $\mathbf{W}_i$ are analytically calculated for each data point. Then, the current calibration and scale uncertainty estimate (27) and increment (24) are calculated. Here, K_l^j is set to zero for all positive scale estimates and it is equal to (26) for negative scale estimates. Finally, the corrections to each data point are calculated and applied, the total measurement correction vector is updated, and the calibration and scale vector is incremented.

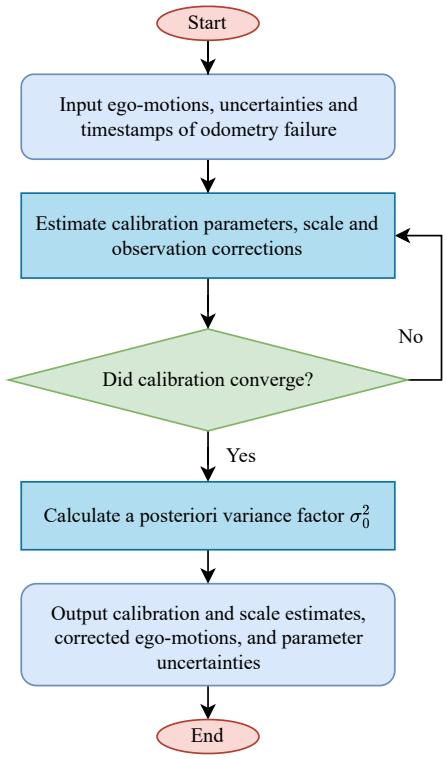

Fig. 2. Flowchart of the proposed uncertainty-aware multiscale calibration solution.

IV. Simulated Experiments

In this section we carried out a systematic analysis on a realistic simulated dataset in order to demonstrate the capabilities of the proposed calibration method. Firstly, we have conducted experiments over a large range of odometry uncertainties and relative scale factors. These experiments are used to illustrate the method's accuracy in estimating calibration parameters and parameter uncertainties. Furthermore, we demonstrated how the estimation of multiple scale factors allows us to benefit from aggregating visual odometry

Algorithm 1 GH-based algorithm for hand-eye calibration

Input:
Ego-motions $\mathbf{l}_i^{(0)}$ and uncertainties $\Sigma_{\mathbf{ll}\,i}$, $i \leftarrow 1$ to N,
odometry failure timestamps

Output:
Calibration and scale vector $\boldsymbol{\xi}$,
associated uncertainty $\Sigma_{\boldsymbol{\xi}\boldsymbol{\xi}}$,
corrected ego-motion measurements $\mathbf{l}_i$, $i \leftarrow 1$ to N

procedure GH CALIBRATION
  initialize $\boldsymbol{\xi} \leftarrow \mathbf{r}_j$, $\mathbf{t}_j = \mathbf{0}$, $s_l^j = 1 \quad \forall j, l$
  repeat
    for each measurement $i \leftarrow 1$ to N
      calculate auxiliary quantities $\mathbf{Y}_i$, $\mathbf{Z}_i$, $\mathbf{W}_i$, $\mathbf{c}_i$, $\mathbf{P}_i$
    for any scale factor $s_l^j < 0$
      calculate K_l^j as in (26) {Ensuring $s_l^j \leftarrow |s_l^j|$}
    otherwise $K_l^j \leftarrow 0$
    calculate uncertainty $\Sigma_{\boldsymbol{\xi}\boldsymbol{\xi}} \leftarrow \left(\sum_{i=1}^{N} \mathbf{Y}_i^\top \mathbf{W}_i \mathbf{Y}_i\right)^{-1}$
    calculate calibration and scale update $\Delta\boldsymbol{\xi}$ as in (24)
    for each measurement $i \leftarrow 1$ to N
      calculate update $\Delta\mathbf{l}_i$ as in (25)
      update measurement $\mathbf{l}_i \leftarrow \Delta\mathbf{l}_i \oplus\!\!\!\!\!\!\bigcirc\; \mathbf{l}_i$
      update total measurement correction $\boldsymbol{\epsilon}_i \leftarrow \mathbf{l}_i \ominus\!\!\!\!\!\!\bigcirc\; \mathbf{l}_i^{(0)}$
    update calibration and scale $\boldsymbol{\xi} \leftarrow \Delta\boldsymbol{\xi} \oplus\!\!\!\!\!\!\bigcirc\; \boldsymbol{\xi}$
  until $\boldsymbol{\xi}$ has converged

(VO) segments, thereby increasing the number of available datapoints. Lastly, we showcased how the proposed method not only estimates calibration parameters but also effectively corrects input odometry. This correction capability and strict adherence to constraints, positions the method as a coherent strategy for trajectory averaging.

A. Simulated trajectory

In order to conduct experiments on simulated data, we have chosen an achievable 3D simulated trajectory shown in Fig. 3 by a continuous line. We set the GT trajectory of sensor A to match this familiar 3D trajectory previously used in [46] and inspired by the trajectory simulator from [10]. The trajectory is defined by the following parametric equations, with $a = 2\,\mathrm{m}$ denoting the half-width in the x axis direction, $b = 1.5\,\mathrm{m}$ in the y and $c = 1.5\,\mathrm{m}$ in the z axis direction,

$$x(t) = a\frac{\cos(t)}{1+\sin^2(t)}, \tag{38}$$

$$y(t) = b\sin(t)x(t), \tag{39}$$

$$z(t) = c\cos(t)y(t). \tag{40}$$

The trajectories of all other sensors were generated by applying the ground truth rigid transformation and multiplying the translational component of ego-motion with $1/s_{\mathrm{GT}}$, with the sensor system connected in a star-shaped network. The ground truth calibration parameters $\mathbf{X}_{\mathrm{GT}}$ are randomly generated, with rotational parameters drawn from a Gaussian distribution $\mathcal{N}(0, \pi/2)$ and translational parameters drawn from $\mathcal{N}(0, 0.2\,\mathrm{m})$. Finally, scale factors were drawn from

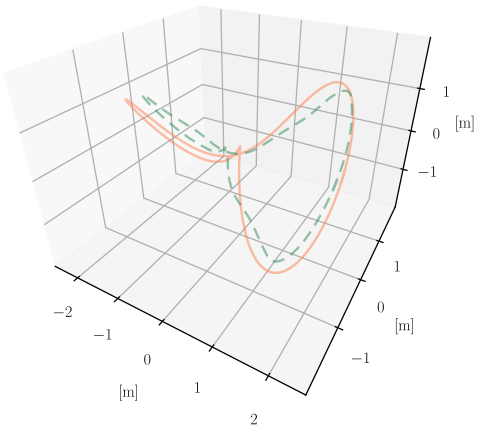

Fig. 3. Example trajectories of two rigidly connected sensors uncorrupted by noise. The trajectory of sensor *A*, described in IV-A, is depicted by an orange continuous line. The trajectory of sensor *B*, depicted by a green dashed line, is generated by applying the ground truth rigid transformation. Relative motions of sensor *A* have the average magnitude of rotation equal to $3.65°$ and the average magnitude of translation of 5.78 cm. To create the simulated input data the *SE(3)* trajectory of each rigidly connected sensor is corrupted by Gaussian noise and the translation estimates of the monocular camera (sensor *B*) are scaled by $s > 0$. This image is best viewed in colour.

a log-uniform distribution spanning from $s_{lower} = 0.01$ to $s_{upper} = 100$. We extracted $N = 300$ equally sampled motion segments per sensor $A_i, B_i, \ldots, M_i$, $i = 1, \ldots, N$ and added noise sampled from a Gaussian distribution. The aforementioned noise instances are drawn from Gaussian distributions centered at 0 with a covariance matrix equaling $\Sigma_{\mathbf{R}} = \bar{\mathbf{r}}\sigma_{\mathbf{R}}\mathbf{I}^3$ or $\Sigma_{\mathbf{t}} = \bar{\mathbf{t}}\sigma_{\mathbf{t}}\mathbf{I}^3$, where $\sigma_{\mathbf{R},\mathbf{t}}$ are expressed as a percentage of the average translational $\bar{\mathbf{t}}$ and rotational $\bar{\mathbf{r}}$ change between poses of the chosen trajectory.

B. Calibration accuracy for a mono-scale system

In order to determine the accuracy and precision of the proposed solution we have put forward three scenarios. In the first one, we assume that uncertainties of all measurements, or more precisely ego-motions, are accurately known. The standard deviation of Gaussian noise for translational parameters varies from approximately $6 \cdot 10^{-4}$ m to $6 \cdot 10^{-3}$ m, and from $6 \cdot 10^{-4}$ rad to $6 \cdot 10^{-3}$ rad for rotational parameters. Results of that experimental scenario are denoted with Σ_{AB} exact. In the second experimental setup the measurement uncertainties are known up to an order of magnitude varying from 10^{-3} m to 10^{-2} m, and from 10^{-3} rad to 10^{-2} rad, with the results denoted by Σ_{AB} ord. mag. Finally, in the third experimental scenario, Σ_{AB} identity the input measurement uncertainties are not known and are initialized as 1. The results of the three experimental scenarios are compared to the performance of the calibration methods presented in [31], [33], [34] and denoted by Wise, Wodtko, Wei, and the dual quaternion and homogeneous matrix solutions of [32], denoted by SchmidtDQ and SchmidtHM, respectively. Results denoted by Wodtko, Wei, SchmidtDQ and SchmidtHM have been achieved using the implementation offered by [34], while the method denoted by Wise was evaluated using the original source code.

TABLE I
ROTATIONAL CALIBRATION ACCURACY OVER 300 TRIALS ON A SIMULATED DATASET.

	$\sigma_{tA} = 1.0$ % $\sigma_{rA} = 5.0$ % $\sigma_{tB} = 5.0$ % $\sigma_{rB} = 1.0$ %	$\sigma_{tA} = 5.0$ % $\sigma_{rA} = 1.0$ % $\sigma_{tB} = 1.0$ % $\sigma_{rB} = 5.0$ %	$\sigma_{tA} = 5.0$ % $\sigma_{rA} = 5.0$ % $\sigma_{tB} = 5.0$ % $\sigma_{rB} = 5.0$ %	$\sigma_{tA} = 5.0$ % $\sigma_{rA} = 10.0$ % $\sigma_{tB} = 10.0$ % $\sigma_{rB} = 5.0$ %	$\sigma_{tA} = 10.0$ % $\sigma_{rA} = 5.0$ % $\sigma_{tB} = 5.0$ % $\sigma_{rB} = 10.0$ %	$\sigma_{tA} = 10.0$ % $\sigma_{rA} = 10.0$ % $\sigma_{tB} = 10.0$ % $\sigma_{rB} = 10.0$ %
	$\mathrm{E}_{\mathbf{R}}$ [°]	$\mathrm{E}_{\mathbf{R}}$ [°]	$\mathrm{E}_{\mathbf{R}}$ [°]	$\mathrm{E}_{\mathbf{R}}$ [°]	$\mathrm{E}_{\mathbf{R}}$ [°]	$\mathrm{E}_{\mathbf{R}}$ [°]
Σ_{AB} exact	$\mathbf{0.2547 \pm 0.1140}$	$\mathbf{0.2684 \pm 0.1147}$	$\mathbf{0.3583 \pm 0.1725}$	$\mathbf{0.5928 \pm 0.2704}$	$\mathbf{0.5626 \pm 0.2665}$	$\mathbf{0.7519 \pm 0.3330}$
Σ_{AB} ord. mag.	$\mathbf{0.2619 \pm 0.1166}$	$\mathbf{0.2692 \pm 0.1147}$	$\mathbf{0.3657 \pm 0.1744}$	$\mathbf{0.6170 \pm 0.2857}$	$\mathbf{0.5657 \pm 0.2660}$	$\mathbf{0.7675 \pm 0.3332}$
Σ_{AB} identity	0.2903 ± 0.1313	0.3046 ± 0.1468	0.4019 ± 0.1959	0.6727 ± 0.3027	0.6204 ± 0.2937	0.8387 ± 0.3996
Wise	0.2706 ± 0.1220	0.2864 ± 0.1300	0.3811 ± 0.1639	$\mathbf{0.5890 \pm 0.2663}$	0.5990 ± 0.2920	0.8220 ± 0.3626
Wodtko	$\mathbf{0.2672 \pm 0.1196}$	$\mathbf{0.2642 \pm 0.1252}$	$\mathbf{0.3657 \pm 0.1634}$	$\mathbf{0.5986 \pm 0.2715}$	$\mathbf{0.5678 \pm 0.2740}$	0.80072 ± 0.3536
Wei	0.3289 ± 0.1557	0.3594 ± 0.1659	0.4763 ± 0.2159	0.7145 ± 0.3378	0.7361 ± 0.3575	0.9975 ± 0.4413
SchmidtDQ	0.6752 ± 0.8722	0.5073 ± 0.5787	0.9317 ± 0.9717	2.0334 ± 2.1563	1.4204 ± 1.5528	2.1521 ± 1.9920
SchmidtHM	0.3419 ± 0.1639	0.2850 ± 0.1313	0.4281 ± 0.1955	0.8951 ± 0.4078	0.6435 ± 0.2747	1.0388 ± 0.4313

TABLE II
TRANSLATIONAL CALIBRATION ACCURACY OVER 300 TRIALS ON A SIMULATED DATASET.

	$\sigma_{tA} = 1.0$ % $\sigma_{rA} = 5.0$ % $\sigma_{tB} = 5.0$ % $\sigma_{rB} = 1.0$ %	$\sigma_{tA} = 5.0$ % $\sigma_{rA} = 1.0$ % $\sigma_{tB} = 1.0$ % $\sigma_{rB} = 5.0$ %	$\sigma_{tA} = 5.0$ % $\sigma_{rA} = 5.0$ % $\sigma_{tB} = 5.0$ % $\sigma_{rB} = 5.0$ %	$\sigma_{tA} = 5.0$ % $\sigma_{rA} = 10.0$ % $\sigma_{tB} = 10.0$ % $\sigma_{rB} = 5.0$ %	$\sigma_{tA} = 10.0$ % $\sigma_{rA} = 5.0$ % $\sigma_{tB} = 5.0$ % $\sigma_{rB} = 10.0$ %	$\sigma_{tA} = 10.0$ % $\sigma_{rA} = 10.0$ % $\sigma_{tB} = 10.0$ % $\sigma_{rB} = 10.0$ %
	$\mathrm{E}_{\mathbf{t}}$ [cm]	$\mathrm{E}_{\mathbf{t}}$ [cm]	$\mathrm{E}_{\mathbf{t}}$ [cm]	$\mathrm{E}_{\mathbf{t}}$[cm]	$\mathrm{E}_{\mathbf{t}}$ [cm]	$\mathrm{E}_{\mathbf{t}}$ [cm]
Σ_{AB} exact	$\mathbf{0.6464 \pm 0.2817}$	0.6629 ± 0.3218	$\mathbf{0.8842 \pm 0.3619}$	$\mathbf{1.4100 \pm 0.6551}$	$\mathbf{1.3581 \pm 0.5921}$	$\mathbf{1.8126 \pm 0.8752}$
Σ_{AB} ord. mag.	$\mathbf{0.6562 \pm 0.2823}$	0.6643 ± 0.3217	$\mathbf{0.9043 \pm 0.3757}$	1.4959 ± 0.6769	1.4665 ± 0.6684	2.1142 ± 0.9406
Σ_{AB} identity	0.7662 ± 0.3359	0.7703 ± 0.3637	1.0399 ± 0.4116	2.0594 ± 1.0128	1.9019 ± 0.8673	2.6477 ± 1.0281
Wise	0.8906 ± 0.3555	0.7987 ± 0.4105	1.2661 ± 0.5581	2.9740 ± 1.1721	2.3917 ± 1.2355	3.9372 ± 1.7983
Wodtko	0.9787 ± 0.4156	0.6592 ± 0.3100	1.2179 ± 0.5215	3.2342 ± 1.3275	1.8910 ± 0.8491	3.7388 ± 1.6718
Wei	0.9992 ± 0.4178	0.7230 ± 0.3406	1.2725 ± 0.5440	3.1770 ± 1.3352	1.9669 ± 0.8794	3.7290 ± 1.6222
SchmidtDQ	1.6735 ± 1.3206	1.2933 ± 1.4360	1.5824 ± 0.7817	3.9805 ± 1.6304	2.4390 ± 1.3009	4.4173 ± 1.9486
SchmidtHM	1.2154 ± 0.5638	$\mathbf{0.6200 \pm 0.2790}$	1.3270 ± 0.6046	4.0323 ± 1.7759	1.7614 ± 0.7077	4.1454 ± 1.8217

We conducted the first set of simulated experiments with randomly generated calibration parameters and scale, as described in Subsec. IV-A and initialized as $\boldsymbol{\xi} = [0., 0., 0., 0., 0., 0., 1.]^{\top}$. At each noise level we executed $N_{exp} = 300$ trials and used rotational, translational and scale errors to define associated metrics

$$\mathrm{E}_{\mathbf{R}} = |\mathbf{e}_{\mathbf{R}}| = |\log(\mathbf{R}_{\mathrm{EST}}\mathbf{R}_{\mathrm{GT}}^{\top})^{\vee}|, \tag{41}$$

$$\mathrm{E}_{\mathbf{t}} = |\mathbf{e}_{\mathbf{t}}| = |\mathbf{t}_{\mathrm{EST}} - \mathbf{t}_{\mathrm{GT}}|, \tag{42}$$

$$\mathrm{E}_{s} = \frac{|e_s|}{s_{\mathrm{GT}}} = \frac{|s_{\mathrm{EST}} - s_{\mathrm{GT}}|}{s_{\mathrm{GT}}}. \tag{43}$$

Additionally, we have defined a failed calibration attempt as a trial in which the algorithm in question fails to converge, $\mathrm{E}_{\mathbf{R}} > 10°$, $\mathrm{E}_{\mathbf{t}} > 10\,\mathrm{cm}$, $\mathrm{E}_{s} > 10\,\%$ or $s < 0$. The solution denoted by Wise had no failed calibration attempts over 300 trials at any noise level, while Wei had a failure rate from 0% to 2.33% and Wodtko from 0% to 6%. The listed failure rates did not depend on measurement noise but on the distance between initialized and ground truth calibration. SchmidtHM failure rate varied from 22.67% to 24.67%, while the dual-quaternion formulation, SchmidtDQ, failure rate linearly rose from 48.33% to 66.67% with increasing noise. Finally, the proposed solution had a varying performance, failing at a rate of 0% to 0.67% regardless of measurement noise. The calibration parameter errors are shown in Tables I, II and III with the lowest average parameter errors and values within 5% displayed in bold. From these results we can see that at the lowest noise levels the proposed approach is rivaled by methods Wodtko, Wei, Wise and SchmidtHM. However, with rising noise the proposed Σ_{AB} exact and Σ_{AB} ord. mag. outperform all other calibration solutions in scale and translation parameter estimation, as well as Wei, SchmidtDQ and SchmidtHM in rotational parameter estimation. The difference in estimation accuracy between experimental scenarios Σ_{AB} exact and Σ_{AB} ord. mag. only starts appearing on very high noise levels for scale factor and translational parameter estimation. The proposed solution in the last calibration scenario Σ_{AB} identity shows poorer performance than with Σ_{AB} exact and Σ_{AB} ord. mag. Although, the proposed solution performs on par or better when initialized with unknown measurement uncertainties. All methods, apart from the proposed solution, show poorer scale, and with that translation, estimation accuracy when the non-metrically scaled sensor is the more noisy of the two. Noise in translational odometry data influences the performance of these algorithms more when present in the non-metrically scaled data, as shown in [34]. When observing calibration accuracy results of asymmetric noise configurations (columns 1 and 2 or columns 4 and 5) we can see that when translational noise is higher in the non-metrically scaled sensor (sensor B),

TABLE III
SCALE ESTIMATION ACCURACY OVER 300 TRIALS ON A SIMULATED DATASET.

	$\sigma_{tA} = 1.0$ % $\sigma_{rA} = 5.0$ % $\sigma_{tB} = 5.0$ % $\sigma_{rB} = 1.0$ %	$\sigma_{tA} = 5.0$ % $\sigma_{rA} = 1.0$ % $\sigma_{tB} = 1.0$ % $\sigma_{rB} = 5.0$ %	$\sigma_{tA} = 5.0$ % $\sigma_{rA} = 5.0$ % $\sigma_{tB} = 5.0$ % $\sigma_{rB} = 5.0$ %	$\sigma_{tA} = 5.0$ % $\sigma_{rA} = 10.0$ % $\sigma_{tB} = 10.0$ % $\sigma_{rB} = 5.0$ %	$\sigma_{tA} = 10.0$ % $\sigma_{rA} = 5.0$ % $\sigma_{tB} = 5.0$ % $\sigma_{rB} = 10.0$ %	$\sigma_{tA} = 10.0$ % $\sigma_{rA} = 10.0$ % $\sigma_{tB} = 10.0$ % $\sigma_{rB} = 10.0$ %
	E_s [%]	E_s [%]	E_s [%]	E_s [%]	E_s [%]	E_s [%]
Σ_{AB} exact	**0.2850 ± 0.2134**	0.2799 ± 0.2147	**0.3852 ± 0.2966**	**0.6207 ± 0.4934**	**0.6070 ± 0.4742**	**0.7793 ± 0.5748**
Σ_{AB} ord. mag.	**0.2936 ± 0.2197**	0.2807 ± 0.2172	0.4756 ± 0.3232	0.8176 ± 0.6858	0.8525 ± 0.6145	1.7940 ± 1.0961
Σ_{AB} identity	0.5936 ± 0.4143	0.5882 ± 0.4290	0.8786 ± 0.4681	2.2248 ± 1.5294	2.2358 ± 1.5003	3.3835 ± 1.3031
Wise	1.0090 ± 0.3462	0.3160 ± 0.2496	1.1301 ± 0.4775	4.0142 ± 0.6895	1.4616 ± 0.8795	4.4118 ± 1.0737
Wodtko	1.0602 ± 0.3602	0.2790 ± 0.2091	1.1127 ± 0.4629	4.1276 ± 0.7094	1.2644 ± 0.7588	4.3247 ± 1.0184
Wei	1.0439 ± 0.3612	0.2817 ± 0.2145	1.0978 ± 0.4694	4.0851 ± 0.7091	1.2541 ± 0.7580	4.2769 ± 1.0159
SchmidtDQ	2.1972 ± 1.9863	1.4231 ± 2.2302	2.3792 ± 2.1626	5.0653 ± 1.2796	2.3184 ± 1.8498	5.4390 ± 1.6065
SchmidtHM	1.1735 ± 0.4193	**0.2611 ± 0.1911**	1.1859 ± 0.4659	4.4287 ± 0.8656	1.2235 ± 0.7207	4.4073 ± 1.0640

the accuracy of translational calibration parameter estimation drops when compared to the mirrored configuration. It is important to note that the proposed GH-based calibration, Σ_{AB} exact, Σ_{AB} ord. mag. and Σ_{AB} identity, is the only method that does not suffer from the asymmetric influence of noise, as shown by almost identical calibration accuracy shown in columns 1 and 2, and columns 4 and 5 of Tables I, II and III. Finally, seeing as the approach denoted by Wodtko performed as well as or better than approaches Wise, Wei, SchmidtDQ and SchmidtHM, we show its distribution of errors along with Σ_{AB} exact and Σ_{AB} ord. mag. in Fig. 4. Again, we have shown that Σ_{AB} exact and Σ_{AB} ord. mag. show superior performance, especially with increasing noise in translational data of the non-metrically scaled sensor. Finally, the difference in performance of Σ_{AB} exact and Σ_{AB} ord. mag. starts appearing only at the highest noise level.

Unfortunately, the Gauss-Helmert model based optimization does not come without its drawbacks. The state is made up of calibration parameters as well as the measurements. It follows that the average runtime of the proposed method in both experimental scenarios ranges from an average of 2.57 s in 10.67 iterations to 3.26 s in 13.24 iterations. k The runtime of remaining methods is roughly equal for all noise level and amounts to Wei ≈ 0.03 s, Wodtko ≈ 0.005 s and Wise ≈ 0.035 s on a PC with a i7-6700HQ CPU at 2.6GHz × 8 and 16GB of 2133MHz DDR4 RAM. The runtime statistic of SchmidtDQ is deceptive because of the amount of failed calibration attempts, with the runtime of successful attempts being roughly equal to 0.7 s. Finally, SchmidtHM has the longest runtime of approximately 3.03 s. The computational complexity of the proposed algorithm with respect to the number of datapoints is $\mathcal{O}(N)$.

C. Calibration accuracy for a multiscale system

Multiscale estimation presents itself as a useful tool when performing odometry-based calibration on short data segments caused by odometry failure or poor temporal synchronization. If the extrinsic calibration of the system remains unchanged, multiscale estimation can allow us to collate multiple shorter segments increasing the volume of available data. To investigate the benefits of multiscale estimation, we constructed a

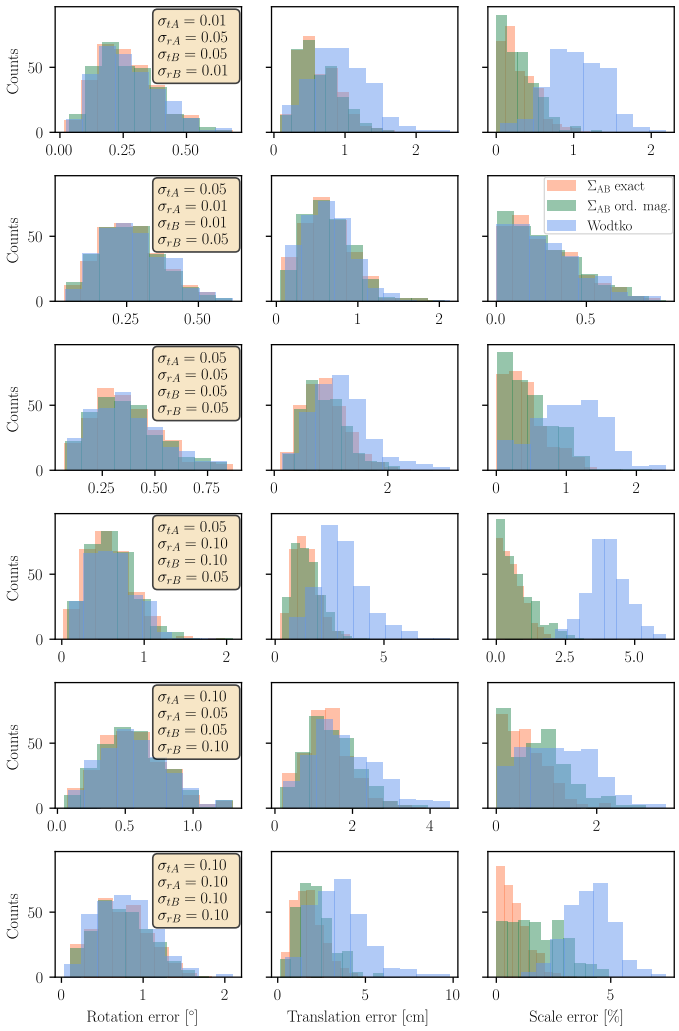

Fig. 4. Histograms of extrinsic calibration and scale error (41), (42), (43), across simulations with varying noise. The per-sensor simulated noise is expressed as a percentage of the average rotational and translational magnitude and asynchronously varies from 1% to 10%. The proposed solution with accurate measurement uncertainty information is denoted in coral, the proposed solution which assumes that the uncertainty of observation is known up to an order of magnitude in green and the competing approach presented in [34] is depicted in blue. The figure is best viewed in color.

TABLE IV
PARAMETER ERRORS WHEN CALIBRATION IS PERFORMED OVER TWO VO SEGMENTS JOINTLY VS. INDIVIDUALLY

	Σ_{AB} ord. mag.					Wodtko				
	$\mathrm{E}_{\mathbf{R}}$ [°]	$\mathrm{E}_{\mathbf{t}}$ [cm]	$\mathrm{E}_{s1}+\mathrm{E}_{s2}$ [%]	E_{s1} [%]	E_{s2} [%]	$\mathrm{E}_{\mathbf{R}}$ [°]	$\mathrm{E}_{\mathbf{t}}$ [cm]	$\mathrm{E}_{s1}+\mathrm{E}_{s2}$ [%]	E_{s1} [%]	E_{s2} [%]
Both VO segments	0.4077	**1.0823**	**1.2860**	×	×	**0.3812**	1.1842	2.1934	×	×
1st segment	0.5311	1.3043	×	0.5577	×	0.5364	1.5068	×	1.0658	×
2nd segment	0.6241	1.3436	×	×	0.9219	0.5406	1.4253	×	×	1.1420

simulated dataset consisting of $N = 300$ data points from the same smooth 3D trajectory (see Fig. 3). Sensor A was considered a reference sensor, whose scale remains unchanged and equal to one throughout the dataset. On the other hand, the relative scale factor of sensor B was modeled to have an abrupt change in the middle of the dataset, at index $i = 150$. The possibility of multiscale estimation allows us to chain these VO segments into one dataset. Then, we have performed $N_{exp} = 300$ trials and to prevent biasing the results we have used trajectories with a symmetric noise configuration, $\sigma_{\mathbf{t}A} = 5\%$, $\sigma_{\mathbf{R}A} = 5\%$, $\sigma_{\mathbf{t}B} = 5\%$ and $\sigma_{\mathbf{R}B} = 5\%$. To create the most realistic conditions possible, we have tested the proposed GH-based method by assuming that we are familiar with data uncertainty up to an order of magnitude. Finally, we have compared the experimental results with the only other method capable of estimating multiple scale factors presented in [34] and denoted with Wodtko.

As seen in Table IV, we have demonstrated that the possibility of extending the dataset by multiscale estimation can improve accuracy of translational and rotational calibration parameter estimation. The sum of scale errors of each segment estimated either jointly or separately remained largely unchanged for both methods, while the estimates of rotational and translational parameters have been shown to improve with joint estimation over both VO segments. The method denoted by Wodtko achieves slightly more accurate rotational parameter estimates but slightly worse performance in translational and significantly worse performance in scale estimation. Unfortunately, the execution time of the proposed approach roughly doubles, while the competing approach retains its fast execution time of 0.008 s.

D. Calibration uncertainty estimation

To test the uncertainty estimation capabilities of the proposed solution, we have conducted a set of $N_{\mathrm{exp}} = 300$ trials per noise level. The calibration parameters and scale have been randomly generated (see Subsec. IV-A). To gauge if the estimated uncertainty provides useful information on the distribution of calibration values, we have to determine the bias of the estimated parameters. The rotational and translation bias is then defined as a sum of differences of estimated and ground truth parameter values, while the relative scale bias is defined as a difference between GT and estimated values divided by the GT scale

$$\mathbf{e}_{\mathbf{R}} = \log\left(\mathbf{R}_{\mathrm{EST}}\mathbf{R}_{\mathrm{GT}}^{\top}\right)^{\vee} \tag{44}$$

$$\mathbf{e}_{\mathbf{t}} = \mathbf{t}_{\mathrm{EST}} - \mathbf{t}_{\mathrm{GT}} \tag{45}$$

$$e_s = \frac{s_{\mathrm{EST}} - s_{\mathrm{GT}}}{s_{GT}} \tag{46}$$

$$\mathbf{e}_{tot} = \begin{bmatrix} \mathbf{e}_{\mathbf{t}} & \mathbf{e}_{\mathbf{R}} & e_s \end{bmatrix}^{\top} \tag{47}$$

$$\longrightarrow \boldsymbol{\mu} = \frac{1}{N_{\mathrm{exp}}}\sum_{n}^{N_{\mathrm{exp}}} \mathbf{e}_{tot,\,n}. \tag{48}$$

From the total parameter difference vector $\mathbf{e}_{tot}$, we can define the associated experimentally achieved covariance matrix over a set of N_{exp} simulated experiments

$$\Sigma_{tot} = \frac{1}{N_{\mathrm{exp}} - 1}\sum_{n}^{N_{\mathrm{exp}}} \mathbf{e}_{tot}\mathbf{e}_{tot}^{\top}. \tag{49}$$

The bias of a set of N_{exp} estimated parameters for $\sigma_{\mathbf{t}A} = 5\%$, $\sigma_{\mathbf{R}A} = 10\%$, $\sigma_{\mathbf{t}B} = 10\%$ and $\sigma_{\mathbf{R}B} = 5\%$ is shown in Table V. Similarly, in Table VI we present the associated standard deviations. The biases of a set of N_{exp} parameters estimated by the proposed GH-based solution are an order of magnitude smaller than the standard deviations for both experimental scenarios. Precisely because the biases of parameters estimated by the proposed solution are significantly smaller than the associated standard deviations, uncertainty information becomes a relevant descriptor of estimated calibration parameters as mentioned in [28].

To visualize the performance of the proposed estimator with respect to parameter covariance matrix estimation, we show the projections of the one-standard-deviation covariance ellipsoids on pairs of axes for rotational, translational and scale blocks of the total covariance matrix in Fig. 5. We can see that both accurate and approximate knowledge of measurement uncertainty empirically yields similar calibration parameter spread. Contrasting that, estimation of calibration parameter uncertainty is significantly affected by inaccuracies in the input measurement uncertainty. Specifically, if the measurement uncertainty has been underestimated, the parameter uncertainty will be underestimated as well. Conversely, an overestimate in measurement uncertainty will result in overestimated parameter uncertainty. By using the metric found in [28] and defined as

$$\varepsilon = \frac{\sqrt{\mathrm{tr}\left((\Sigma_{\mathrm{EST}} - \Sigma_{\mathrm{EXP}})^{\top}(\Sigma_{\mathrm{EST}} - \Sigma_{\mathrm{EXP}})\right)}}{\sqrt{\mathrm{tr}(\Sigma_{\mathrm{EXP}}{}^{\top}\Sigma_{\mathrm{EXP}})}}, \tag{50}$$

we can quantify the average difference between estimated covariance matrices in each experiment and ones calculated

TABLE V
ESTIMATED CALIBRATION PARAMETER BIAS IN THE TWO EXPERIMENTAL SCENARIOS, WITH $s \neq 1$.

	$\mu_{\mathbf{t}x}$ [mm]	$\mu_{\mathbf{t}y}$ [mm]	$\mu_{\mathbf{t}z}$ [mm]	$\mu_{\mathbf{R}x}$ [°]	$\mu_{\mathbf{R}y}$ [°]	$\mu_{\mathbf{R}z}$ [°]	μ_s [1]
Σ_{AB} exact	−0.1138	0.1467	0.1898	−0.0212	−0.0154	0.0104	−0.0003
Σ_{AB} ord. mag.	0.0354	1.2890	0.1159	−0.0234	−0.0275	0.0227	−0.0023

TABLE VI
STANDARD DEVIATION OF ESTIMATED CALIBRATION PARAMETERS IN THE TWO EXPERIMENTAL SCENARIOS, WITH $s \neq 1$.

	$\sigma_{\mathbf{t}x}$ [mm]	$\sigma_{\mathbf{t}y}$ [mm]	$\sigma_{\mathbf{t}z}$ [mm]	$\sigma_{\mathbf{R}x}$ [°]	$\sigma_{\mathbf{R}y}$ [°]	$\sigma_{\mathbf{R}z}$ [°]	σ_s [1]
Σ_{AB} exact	8.7853	7.1861	10.2887	0.3807	0.2366	0.3749	0.0072
Σ_{AB} ord. mag.	8.8188	7.6578	10.3496	0.3892	0.2491	0.4072	0.0091

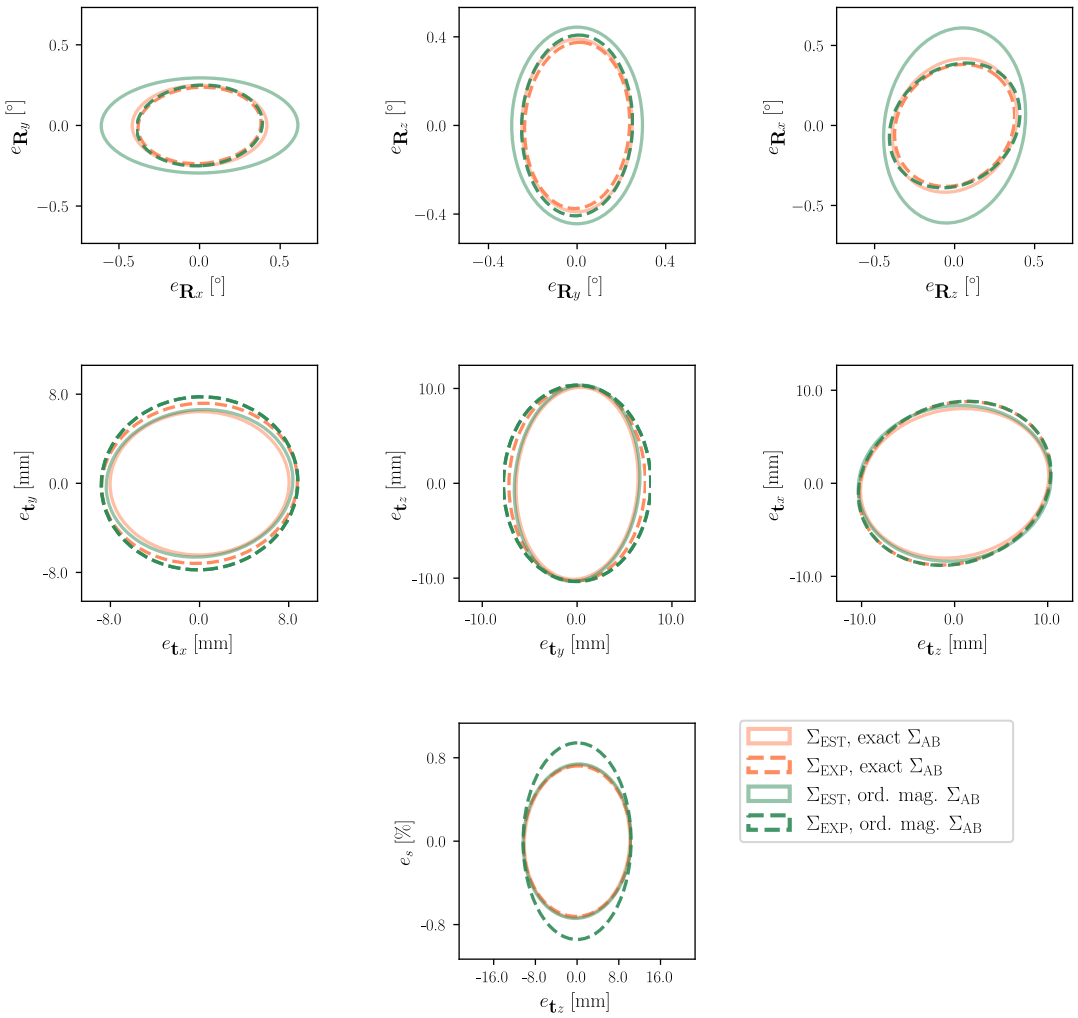

Fig. 5. Projection of the one-standard-deviation covariance ellipsoids on pairs of axes for rotational, translational and scale factor blocks of the total covariance matrix for $\sigma_{\mathbf{t}A} = 5\%$, $\sigma_{\mathbf{R}A} = 10\%$, $\sigma_{\mathbf{t}B} = 10\%$ and $\sigma_{\mathbf{R}B} = 5\%$. The projections of the covariance matrix calculated from the results of N_{exp} experiments with known exact measurement uncertainties are shown with a coral dashed line, while the corresponding projections of the covariance estimated by the proposed solution are shown by a coral continuous. Similarly, experimental and estimated covariance matrix projections for measurement uncertainty known to the order of magnitude are shown in green dashed and continuous lines, respectively. This figure is best viewed in color.

over N_{exp} experiments. The average value of ε per noise level is presented in Fig. 6. Again, we present this measure of discrepancy of estimated to achieved calibration parameter uncertainty for algorithm runs with measurement uncertainty accurately known and known to the order of magnitude. The results presented in Fig. 6 show that the proposed GH-based solution accurately estimates the calibration parameter uncertainty when measurement uncertainty is accurately known over a sizeable range of noise magnitudes. As expected, parameter uncertainty estimated with approximate knowledge

Fig. 6. Discrepancy between rotational, translational and scale factor blocks of the estimated and experimentally achieved parameter covariance matrices when initializing the proposed method with measurement covariance known accurately or to the order of magnitude. Each datapoint represents the average discrepancy over the set of N_{exp} simulated experiments for a given noise level. The figure is best viewed in color.

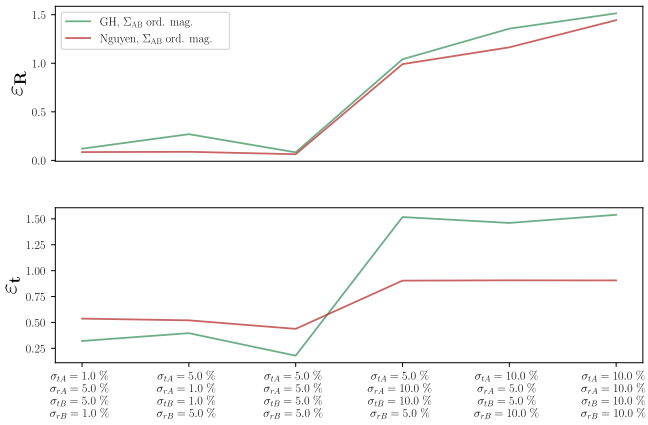

Fig. 7. Discrepancy between rotational and translational blocks of the estimated and experimentally achieved covariance matrices when initializing the proposed method with measurement covariance known accurately or to the order of magnitude and the competing estimator denoted by Nguyen. Each datapoint represents the average discrepancy over the set of N_{exp} simulated experiments for a given noise level. The method denoted by Nguyen does not estimate scale, therefore the scale uncertainty distance is omitted. The figure is best viewed in color.

of measurement uncertainty shows a larger deviation, with the behaviour being governed by how close the approximate measurement uncertainty is to real values.

While the method in [40] also estimates parameter uncertainties and considers unknown scaling, it is restricted to camera calibration. To the best of our knowledge there is no counterpart for the proposed method that exclusively uses odometry data to perform calibration with unknown scale. However, we can construct a simulated dataset, with scale equal to one and compare uncertainty estimation capabilities of the proposed approach with the approach denoted by Nguyen and presented in [28]. The method presented in [28] is capable of estimating calibration parameters and their uncertainty using metrically scaled ego-motion and associated covariance as input. This calibration procedure separates rotational and translational parameter and covariance estimation, increasing the potential for calibration inaccuracies. Finally, we simply remove scale from the calibration parameter estimation of the proposed analytic GH-based solution and use approximate values of measurement uncertainty as input.

The failure rate of Nguyen, when initialized agnostically with the GT parameters set as mentioned in IV-A, ranges from 49% at the lowest noise levels and reaches 100% at $\sigma_{t,\,r} = 5\%$. When initialized through the closed form least-squares method suggested by the authors of [28], the failure rate drops to zero for the three lowest noise configurations and to 2%, 3.67%, and 14.67% for the three highest configurations. The proposed solution retains a 0% failure rate with both initializations. After removing the failed calibration attempts, we present the absolute calibration errors defined by (42) and (41) in Tab. VII. From Tab. VII we can see that the method Nguyen maintains a slightly worse performance for all noise levels. Finally, to fully characterize the performance of the Nguyen method, we have repeated the analysis presented in Fig. 6. Naturally, the discussion of scale was omitted from the analysis. From Fig. 7 we can determine that with an approximate knowledge of measurement covariance, Nguyen slightly outperforms in rotational uncertainty estimation. However, the proposed solution performs slightly better when estimating translational uncertainty at low noise, with the trend reversing at the three highest noise levels. Uncertainty estimation accuracy of the proposed algorithm, as mentioned previously, depends on the quality of approximation, when measurement uncertainty is known only up to the order of magnitude. This behaviour is also present in Nguyen but it is alleviated by correcting the rotational measurements, and the associated uncertainties, during the estimation procedure. Regardless, the proposed method demonstrates more accurate parameter estimation and higher robustness to initialization accuracy, where a relatively poorer calibration parameter estimation accuracy of Nguyen could be attributed to the decoupling of rotational and translational estimation, [28].

E. Multisensor trajectory correction and averaging

One final feature of the proposed calibration method is its ability to correct trajectories of a multisensor system and provide an uncertainty-aware trajectory average. Again we have conducted $N_{exp} = 300$ experiments for a four sensor system out of which only sensor A is metrically scaled. The noise for each experimental run was of magnitude $\sigma_{\mathbf{t}A} = 5\%$, $\sigma_{\mathbf{R}A} = 5\%$, $\sigma_{\mathbf{t}B} = 5\%$, $\sigma_{\mathbf{R}B} = 10\%$, $\sigma_{\mathbf{t}C} = 10\%$, $\sigma_{\mathbf{R}C} = 5\%$, $\sigma_{\mathbf{t}D} = 10\%$ and $\sigma_{\mathbf{R}D} = 10\%$. The calibration parameters, as well as the scale factor, have been kept constant throughout the N_{exp} trials, with the proposed calibration approach being used to estimate the calibration parameters and correct the simulated noisy trajectory. We have quantified the trajectory improvements by comparing the trajectory of each sensor pre- and post- estimation with the ground truth. To make that comparison, we have used the *RPG trajectory evaluation*

TABLE VII
CALIBRATION ACCURACY OVER 300 TRIALS ON A SIMULATED DATASET FOR $s = 1$.

	$\sigma_{tA} = 1.0$ % $\sigma_{rA} = 5.0$ % $\sigma_{tB} = 5.0$ % $\sigma_{rB} = 1.0$ %		$\sigma_{tA} = 5.0$ % $\sigma_{rA} = 1.0$ % $\sigma_{tB} = 1.0$ % $\sigma_{rB} = 5.0$ %		$\sigma_{tA} = 5.0$ % $\sigma_{rA} = 5.0$ % $\sigma_{tB} = 5.0$ % $\sigma_{rB} = 5.0$ %		$\sigma_{tA} = 5.0$ % $\sigma_{rA} = 10.0$ % $\sigma_{tB} = 10.0$ % $\sigma_{rB} = 5.0$ %		$\sigma_{tA} = 10.0$ % $\sigma_{rA} = 5.0$ % $\sigma_{tB} = 5.0$ % $\sigma_{rB} = 10.0$ %		$\sigma_{tA} = 10.0$ % $\sigma_{rA} = 10.0$ % $\sigma_{tB} = 10.0$ % $\sigma_{rB} = 10.0$ %	
	$E_{\mathbf{R}}$ [°]	$E_{\mathbf{t}}$ [cm]	$E_{\mathbf{R}}$ [°]	$E_{\mathbf{t}}$ [cm]	$E_{\mathbf{R}}$ [°]	$E_{\mathbf{t}}$ [cm]	$E_{\mathbf{R}}$ [°]	$E_{\mathbf{t}}$[cm]	$E_{\mathbf{R}}$ [°]	$E_{\mathbf{t}}$ [cm]	$E_{\mathbf{R}}$ [°]	$E_{\mathbf{t}}$ [cm]
Σ_{AB} ord. mag.	**0.2680**	**0.6193**	**0.2437**	**0.5933**	**0.3646**	**0.8528**	**0.6046**	**1.3342**	**0.5600**	**1.3328**	**0.7252**	**1.7835**
Nguyen ord. mag.	0.3539	0.6944	0.3384	0.6763	0.4682	0.9285	0.7886	1.5240	0.7596	1.4960	0.9533	2.0136

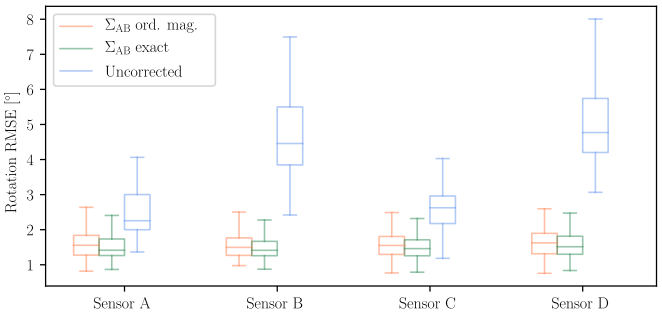

Fig. 8. Average rotational absolute trajectory error of a trajectory with 300 datapoints and a total rotational change of 1095°, as defined by [47]. The ATE of the corrected trajectories in the two experimental scenarios are represented in green and coral and ATE of the uncorrected sensor trajectories is presented in blue. The figure is best viewed in color.

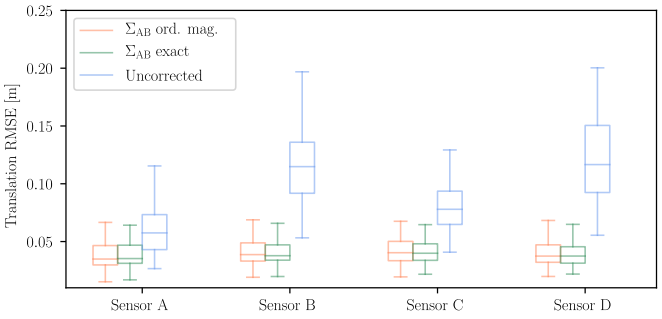

Fig. 9. Average translational absolute trajectory error of a trajectory with 300 datapoints and a total translational change of 17.33 m, as defined by [47]. The ATE of the corrected trajectories in the two experimental scenarios are represented in green and coral and ATE of the uncorrected sensor trajectories is presented in blue.

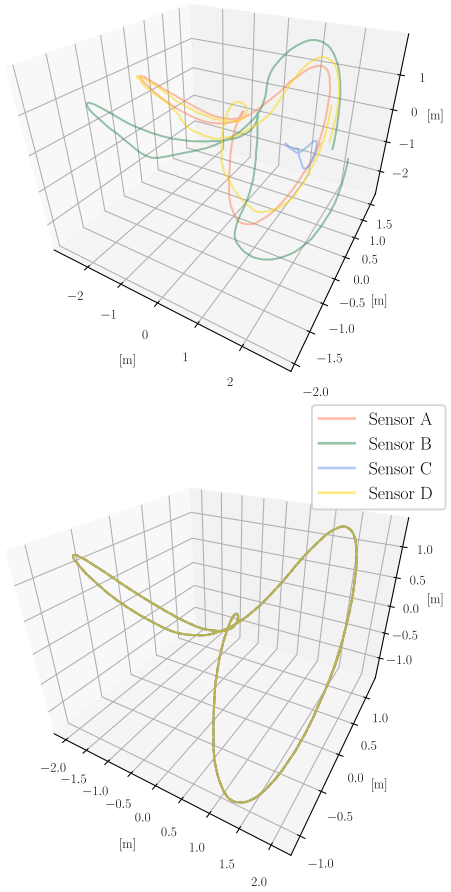

Fig. 10. Example of uncorrected unscaled sensor trajectories (top panel) of a four sensor system and the corrected average trajectory in the ego-centric frame of sensor A (bottom panel). The figure is best viewed in color.

package, [47], to calculate absolute trajectory errors (ATE) for the two proposed experimental scenarios and for the uncorrected sensor trajectories. Rotation and translation Root-Mean-Square-Error (RMSE) per sensor as defined by [47] are depicted in Figures 8 and 9. From the aforementioned ATE results, we can see that the GH-based proposed solutions significantly improve the alignment of trajectories with the ground truth, with the effect being more pronounced at higher odometry noise. In this set of experiments, the most significant improvement was achieved for sensors B and D which both have the same, and highest, noise level in rotational components of ego-motion. It follows that the trajectory correction capabilities of the proposed solution are especially useful when dealing with odometry data with high noise in the rotational components of poses. Additionally, it is worth noting that, with both accurately known measurement uncertainties and with measurement uncertainties known to the order of magnitude, the proposed solution estimates roughly similarly accurate trajectory corrections. Finally, because the Gauss-Helmert model minimizes the corrections to observations but strictly enforces constraints, GH-based calibration solutions can be thought of as trajectory averaging tools, making trajectories of rigidly connected sensors perfectly aligned when transformed using the estimated $\mathbf{X}$. In Fig. 10 we can see the transition from noisy unscaled trajectories to an averaged trajectory, with trajectories of all other sensors having been transformed using the estimated $\mathbf{X}$ into the coordinate frame of sensor A.

TABLE VIII
ESTIMATED CALIBRATION PARAMETER ERROR BETWEEN TWO MONOCULAR CAMERAS FOR SIX Room SEQUENCES.

	Room 1			Room 2			Room 3			Room 4			Room 5			Room 6		
	$E_{\mathbf{R}}$ [°]	$E_{\mathbf{t}}$ [cm]	s [1]	$E_{\mathbf{R}}$ [°]	$E_{\mathbf{t}}$ [cm]	s [1]	$E_{\mathbf{R}}$ [°]	$E_{\mathbf{t}}$ [cm]	s [1]	$E_{\mathbf{R}}$ [°]	$E_{\mathbf{t}}$[cm]	s [1]	$E_{\mathbf{R}}$ [°]	$E_{\mathbf{t}}$ [cm]	s [1]	$E_{\mathbf{R}}$ [°]	$E_{\mathbf{t}}$ [cm]	s [1]
GH, Σ_{DSO}	0.0358	1.4009	1.0876	**0.0280**	**0.3809**	0.3970	0.0333	2.7110	0.9151	0.0185	**1.5807**	0.8910	0.0414	**0.1703**	0.8277	**0.0227**	**0.2929**	0.8456
GH, $\Sigma_{\mathbf{I}}$	0.0200	**1.1400**	1.0690	0.3287	1.3599	0.3811	0.1113	2.1150	0.8998	0.0243	1.8895	0.8819	**0.0271**	**0.1745**	0.8226	0.0673	0.6784	0.8268
Wise	0.0204	**1.1940**	1.0708	0.4313	1.9403	0.3836	0.2077	1.8841	0.8990	0.0297	1.9354	0.8779	0.0310	**0.1760**	0.8220	0.0953	0.8422	0.8262
Wodtko	**0.0144**	**1.1912**	1.0708	0.6176	2.1861	0.3837	0.2558	1.7986	0.8988	0.0451	1.9383	0.8778	0.0321	0.1837	0.8220	0.1156	0.8757	0.8261
Wei	0.0189	**1.1897**	1.0708	0.0495	1.4774	0.3833	**0.0313**	2.3337	0.8987	**0.0127**	1.9281	0.8778	**0.0281**	**0.1741**	0.8220	0.0297	0.7783	0.8262
SchmidtDQ	**0.0146**	**1.1869**	1.0706	0.9579	2.7736	0.3838	0.2710	1.7763	0.8988	0.0530	1.9403	0.8778	0.0338	0.1845	0.8219	0.1237	0.8793	0.8261
SchmidtHM	0.0214	**1.1962**	1.0709	0.8845	2.5369	0.3839	0.3149	**1.6700**	0.8990	0.0679	1.9470	0.8779	0.0480	0.1915	0.8220	0.1498	0.8918	0.8262

V. EXPERIMENTS ON REAL-WORLD DATA

In order to validate the previous simulation results on real-world data, we have performed similar analyses on the TUM VI dataset [48]. The dataset was recorded using two monochrome cameras in a stereo setup and an IMU on a handheld mount, with the GT odometry acquired using a motion capture system and IMU. The calibration values, determined using a target-based procedure, represent the best possible calibration achievable under laboratory conditions. Although not an absolute ground truth, we will henceforth refer to them as such for the purpose of comparison in our experiments. In the chosen six *Room* sequences GT odometry is available throughout the duration of the sequence. In this experimental setup we have viewed the two cameras as monocular, with the trajectory of the two monocular cameras estimated using the open source VO solution DSO, [49]. Additionally, the variances of the estimated camera poses have been used, unaltered, as estimated by DSO. In order to compare the calibration results of the proposed algorithm with other single sensor methods we have estimated the similarity transformation between the cameras **Cam1** and **Cam2**, while using the GT odometry to analyze the influence of the accuracy of odometry variance estimation on calibration accuracy.

In this section we present calibration accuracy results of the proposed method when using uncertainty estimated by DSO, $\mathrm{GH}, \Sigma_{\mathrm{DSO}}$, and with an agnostic uncertainty input, $\mathrm{GH}, \Sigma_{\mathbf{I}}$. Again, the proposed approach is compared to methods Wise, Wodtko, Wei, SchmidtDQ and SchmidtHM on all six *Room* TUM VI dataset sequences, [48]. In all six sequences **Cam1** was chosen as sensor A and all of the algorithms were initialized with calibration parameters $\boldsymbol{\xi} = [0., 0., 0., 0., 0., 0., 1.]^\top$ Finally, because the ground truth scale information cannot exist, we have reported the scale values directly as estimated. The accuracy of performed hand-eye calibration is reported in Tab. VIII, with best results, and values within 5% of it, shown in bold. Additionally, the minimum, average and maximum errors over the six *Room* dataset are shown in Tab. IX. We can see that the proposed method with uncertainty information and the competing approach denoted with Wei both consistently maintain smallest rotational errors, which never exceed $0.1°$. On the other hand, translational estimation performance varies less across methods, with the GH-based solution having the lowest average error. Minimum translational error of all algorithms was achieved on one dataset, while the maximum error of all methods was found on datasets *Room 2* and *Room 3*. Interestingly, the proposed solution performed poorly on dataset *Room 3* but achieved significantly better results than the competition on *Room 2*. Similarly, the proposed solution outperformed all others on *Room 6*. This can conceivably be explained by uncharacteristically accurate DSO uncertainty estimation. Unfortunately, determining the accuracy of pose uncertainty at each timestep is virtually impossible. However, by comparing the DSO estimated poses and GT values, it is clear that the DSO provided measurement variance is usually multiple orders of magnitude smaller than the one suggested by pose to GT errors. This makes calibration parameter uncertainty estimates unreliable at best, forcing us to rely on simulated experiments. While further investigation into methods of pose uncertainty estimation will produce more accurate calibration parameter, and parameter uncertainty, estimates, it lies outside the scope of this paper.

TABLE IX
MINIMUM, AVERAGE AND MAXIMUM ERRORS OVER SIX *Room* DATASETS.

	Minimum		Average		Maximum	
	$E_{\mathbf{R}}$ [°]	$E_{\mathbf{t}}$ [cm]	$E_{\mathbf{R}}$ [°]	$E_{\mathbf{t}}$ [cm]	$E_{\mathbf{R}}$ [°]	$E_{\mathbf{t}}$ [cm]
GH, Σ_{DSO}	0.0185	**0.1703**	0.0300	**1.0895**	**0.0414**	2.7110
GH, $\Sigma_{\mathbf{I}}$	0.0200	**0.1745**	0.0965	1.2262	0.3287	2.1150
Wise	0.0204	**0.1760**	0.1359	1.3287	0.4313	**1.9403**
Wodtko	0.0144	0.1837	0.1801	1.3623	0.6176	2.1861
Wei	**0.0127**	**0.1741**	**0.0284**	1.3136	0.0495	2.3337
SchmidtDQ	0.0146	0.1845	0.2423	1.4568	0.9597	2.7736
SchmidtHM	0.0214	0.1915	0.2478	1.4056	0.8845	2.5369

Finally, in order to show the benefits of multiscale calibration we have performed the same calibration procedure on paired data of consecutive sequences, the results of which are shown in Tab. X. The best results, and values within 5%, are shown in bold. Comparing the results of Tab. VIII and Tab. X we can note that all of the methods scale estimates remain very similar, with GH, Σ_{DSO} estimates seemingly being the most stable. Here we have demonstrated that by jointly calibrating sequences that contain more noisy odometry data with less noisy ones, we can improve calibration results. Again, it is worth noting that the accuracy of the proposed method depends not only on the noise present in the data but also on the accuracy of input measurement uncertainty. The accuracy of joint calibration performed by GH, Σ_{DSO} generally lies between the values for each individual sequence, with some notable improvements being made when one of the sequences allows for very accurate calibration, e.g. when calibrating sequences Room 1 & 2.

TABLE X
ESTIMATED CALIBRATION PARAMETER ERROR BETWEEN TWO CAMERAS ON MULTI-SEQUENCE DATASETS.

	Room 1 & 2				Room 3 & 4				Room 5 & 6			
	$\mathrm{E}_{\mathbf{R}}$ [°]	$\mathrm{E}_{\mathbf{t}}$ [cm]	s_1 [1]	s_2 [1]	$\mathrm{E}_{\mathbf{R}}$ [°]	$\mathrm{E}_{\mathbf{t}}$ [cm]	s [1]	s_2 [1]	$\mathrm{E}_{\mathbf{R}}$ [°]	$\mathrm{E}_{\mathbf{t}}$ [cm]	s [1]	s_2 [1]
GH, Σ_{DSO}	**0.0263**	**0.3912**	1.0876	0.3996	**0.0278**	2.4029	0.9147	0.8910	0.0288	**0.2003**	0.8275	0.8455
GH, $\Sigma_{\mathbf{I}}$	0.0293	1.0390	1.0709	0.3787	0.0765	1.9929	0.9031	0.8800	0.0223	0.3081	0.8212	0.8276
Wodtko	0.2096	1.4768	1.0733	0.3824	0.1414	**1.8713**	0.9021	0.8762	**0.0190**	0.4654	0.8200	0.8262

VI. CONCLUSION

In this paper, we have presented an unscaled hand-eye calibration solution based on the Gauss-Helmert model. The proposed formulation allows the algorithm to not only maintain good performance with highly noisy input data, but also to apply corrections based on measurement uncertainty information. Through testing on simulated odometry data, we have demonstrated that our method can outperform the competing solutions in specific high-noise scenarios, particularly when reliable measurement uncertainty information is available. However, the magnitude of improvement depends significantly on the quality of the uncertainty information provided. Furthermore, we have shown the benefits of extending data sequences by utilizing multiscale estimation. The use of valuable information contained in measurement uncertainty was shown to achieve solid accuracy in calibration and scale parameter uncertainty estimation, being the first exclusively odometry-based estimator to address the estimation of scale uncertainty. Moreover, the evaluation on simulated odometry data has shown that with reasonably accurate measurement uncertainty information, the proposed method achieves significant reduction in absolute trajectory error. Finally, real-world data testing corroborates the results of simulated experiments highlighting the need for more accurate odometry uncertainty estimation methods. While GH-based uncertainty aware hand-eye calibration is still firmly an offline procedure, developing an online GH-based calibration solution presents itself as a prospective avenue for future research. Unlike the current offline procedure, online GH-based calibration holds the potential to enable real-time probabilistic calibration monitoring and on-the-fly recalibration.

APPENDIX
ELEMENTARY JACOBIAN BLOCKS

Here we have provided full derivations of the used partial derivatives of unscaled hand-eye constraints. Firstly, we have calculated the derivative of the rotation action with regard to $\mathbf{r} \in \mathbb{R}^3 \cong$ *TSO(3)*,

$$\begin{aligned}\frac{\partial \mathbf{R}\mathbf{v}}{\partial \mathbf{r}} &= \lim_{\delta\mathbf{r}\to 0} \frac{\mathrm{Exp}(\delta\mathbf{r})\mathbf{R}\mathbf{v} - \mathbf{R}\mathbf{v}}{\delta\mathbf{r}} \\ &\approx \lim_{\delta\mathbf{r}\to 0} \frac{(\mathbf{I}_3 + \delta\mathbf{r}^\wedge)\mathbf{R}\mathbf{v} - \mathbf{R}\mathbf{v}}{\delta\mathbf{r}} \\ &= \lim_{\delta\mathbf{r}\to 0} \frac{-(\mathbf{R}\mathbf{v})^\wedge \delta\mathbf{r}}{\delta\mathbf{r}} = -(\mathbf{R}\mathbf{v})^\wedge. \end{aligned} \tag{51}$$

This partial Jacobian was reached using the expression $\mathbf{x}^\wedge\mathbf{y} = -\mathbf{y}^\wedge\mathbf{x}$ from [41] and left multiplicative incrementation (10). The other partial derivative associated with rotational parameters can be calculated using a similar procedure,

$$\begin{aligned}\frac{\partial \mathbf{r}}{\partial \mathbf{r}} &= \lim_{\delta\mathbf{r}\to 0} \frac{\mathrm{Log}\left(\mathrm{Exp}(\delta\mathbf{r})\mathrm{Exp}(\mathbf{r})\right) - \mathbf{r}}{\delta\mathbf{r}} \\ &\approx \lim_{\delta\mathbf{r}\to 0} \frac{\mathrm{Log}\left(\mathrm{Exp}(\mathbf{J}_l^{-1}(\mathbf{r})\delta\mathbf{r} + \mathbf{r})\right) - \mathbf{r}}{\delta\mathbf{r}} \\ &= \lim_{\delta\mathbf{r}\to 0} \frac{\mathbf{J}_l^{-1}(\mathbf{r})\delta\mathbf{r} + \mathbf{r} - \mathbf{r}}{\delta\mathbf{r}} = \mathbf{J}_l^{-1}(\mathbf{r}). \end{aligned} \tag{52}$$

Lastly,

$$\begin{aligned}\frac{\partial \mathbf{M}\mathbf{r}}{\partial \mathbf{r}} &= \lim_{\delta\mathbf{r}\to 0} \frac{\mathbf{M}\cdot\mathrm{Log}\left(\mathrm{Exp}(\delta\mathbf{r})\mathrm{Exp}(\mathbf{r})\right) - \mathbf{M}\mathbf{r}}{\delta\mathbf{r}} \\ &\approx \lim_{\delta\mathbf{r}\to 0} \frac{\mathbf{M}\cdot\mathrm{Log}\left(\mathrm{Exp}(\mathbf{J}_l^{-1}(\mathbf{r})\delta\mathbf{r} + \mathbf{r})\right) - \mathbf{M}\mathbf{r}}{\delta\mathbf{r}} \\ &= \lim_{\delta\mathbf{r}\to 0} \frac{\mathbf{M}\mathbf{J}_l^{-1}(\mathbf{r})\delta\mathbf{r} + \mathbf{M}\mathbf{r} - \mathbf{M}\mathbf{r}}{\delta\mathbf{r}} = \mathbf{M}\mathbf{J}_l^{-1}(\mathbf{r}), \end{aligned} \tag{53}$$

where $\mathbf{M}$ represents any $\mathbb{R}^3$ matrix not dependent on $\mathbf{r}$. For completion we list the partial derivatives with regard to the translation vector $\mathbf{t}$ and the scale factor s,

$$\frac{\partial \mathbf{R}\mathbf{t}}{\partial \mathbf{t}} = \mathbf{R}, \quad \frac{\partial \mathbf{t}}{\partial \mathbf{t}} = \mathbf{I}_3, \quad \frac{\partial s\mathbf{v}}{\partial s} = \mathbf{v}. \tag{54}$$

Here we present partial Jacobians present in covariance propagation (35). By dividing Jacobians (36) and (37) into four 3×3 blocks

$$\mathcal{J} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{B}_2 \\ \mathbf{B}_3 & \mathbf{B}_4 \end{bmatrix}, \tag{55}$$

we can see that blocks $\mathbf{B}_1$ and $\mathbf{B}_2$ of both Jacobians can be calculated using (54) and (51). Block $\mathbf{B}_3$ is always equal to $\mathbf{0}_{3\times 3}$, while block $\mathbf{B}_4$ of Jacobian (36) can be calculated using the definition of a left group Jacobian [42], and $\mathbf{Ad}_{\mathbf{R}^\top} =$

$\mathbf{Ad}_{\mathbf{R}}^{\top} = \mathbf{R}^{\top}$,

$$
\begin{aligned}
&\frac{\partial \mathrm{Log}(\mathbf{R}^{\top}\mathbf{R}')}{\partial \mathbf{r}} = \left\{ \mathbf{r}'' = \mathrm{Log}(\mathbf{R}^{\top}\mathbf{R}') \right\} \\
&= \lim_{\delta\mathbf{r}\to 0} \frac{\mathrm{Log}((\mathrm{Exp}(\delta\mathbf{r})\mathbf{R})^{\top}\mathbf{R}') - \mathrm{Log}(\mathbf{R}^{\top}\mathbf{R}')}{\delta\mathbf{r}} \\
&= \lim_{\delta\mathbf{r}\to 0} \frac{\mathrm{Log}(\mathbf{R}^{\top}\mathrm{Exp}(-\delta\mathbf{r})\mathbf{R}') - \mathbf{r}''}{\delta\mathbf{r}} \\
&= \lim_{\delta\mathbf{r}\to 0} \frac{\mathrm{Log}(\mathbf{R}^{\top}\mathrm{Exp}(-\delta\mathbf{r})\mathbf{R}\mathbf{R}^{\top}\mathbf{R}') - \mathbf{r}''}{\delta\mathbf{r}} \\
&= \lim_{\delta\mathbf{r}\to 0} \frac{\mathrm{Log}(\mathrm{Exp}(-\mathbf{Ad}_{\mathbf{R}^{\top}}\delta\mathbf{r})\mathbf{R}^{\top}\mathbf{R}') - \mathbf{r}''}{\delta\mathbf{r}} \\
&= \lim_{\delta\mathbf{r}\to 0} \frac{\mathrm{Log}(\mathrm{Exp}(-\mathbf{R}^{\top}\delta\mathbf{r})\mathbf{R}^{\top}\mathbf{R}') - \mathbf{r}''}{\delta\mathbf{r}} \\
&= \lim_{\delta\mathbf{r}\to 0} \frac{\mathrm{Log}(\mathrm{Exp}(-\mathbf{J}_l^{-1}(\mathbf{r}'')\mathbf{R}^{\top}\delta\mathbf{r} + \mathbf{r}'')) - \mathbf{r}''}{\delta\mathbf{r}} \\
&= -\mathbf{J}_l^{-1}(\mathbf{r}'')\mathbf{R}^{\top} = -\mathbf{J}_l^{-1}(\mathrm{Log}(\mathbf{R}^{\top}\mathbf{R}'))\mathbf{R}^{\top}. \qquad (56)
\end{aligned}
$$

Similarly, using $\mathbf{J}_r^{-1}(\mathbf{r}) = \mathbf{J}_l^{-1}(\mathbf{r})\mathbf{R}$ from [41], we can show that the $\mathbf{B}_4$ block of Jacobian (37) equals

$$
\begin{aligned}
&\frac{\partial \mathrm{Log}(\mathbf{R}^{\top}\mathbf{R}')}{\partial \mathbf{r}'} = \left\{ \mathbf{r}'' = \mathrm{Log}(\mathbf{R}^{\top}\mathbf{R}') \right\} \\
&= \lim_{\delta\mathbf{r}'\to 0} \frac{\mathrm{Log}(\mathbf{R}^{\top}\mathbf{R}'\mathbf{R}'^{\top}\mathrm{Exp}(\delta\mathbf{r}')\mathbf{R}') - \mathrm{Log}(\mathbf{R}^{\top}\mathbf{R}')}{\delta\mathbf{r}'} \\
&= \lim_{\delta\mathbf{r}\to 0} \frac{\mathrm{Log}(\mathbf{R}^{\top}\mathbf{R}'\mathrm{Exp}(\mathbf{Ad}_{\mathbf{R}'^{\top}}\delta\mathbf{r})) - \mathbf{r}''}{\delta\mathbf{r}'} \\
&= \lim_{\delta\mathbf{r}\to 0} \frac{\mathrm{Log}(\mathbf{R}^{\top}\mathbf{R}'\mathrm{Exp}(\mathbf{R}'^{\top}\delta\mathbf{r})) - \mathbf{r}''}{\delta\mathbf{r}'} \\
&= \lim_{\delta\mathbf{r}\to 0} \frac{\mathrm{Log}(\mathrm{Exp}(\mathbf{J}_r^{-1}(\mathbf{r}'')\mathbf{R}'^{\top}\delta\mathbf{r}' + \mathbf{r}'')) - \mathbf{r}''}{\delta\mathbf{r}'} \\
&= \mathbf{J}_r^{-1}(\mathbf{r}'')\mathbf{R}'^{\top} = \mathbf{J}_l^{-1}(\mathrm{Log}(\mathbf{R}^{\top}\mathbf{R}'))\mathbf{R}^{\top}\mathbf{R}'\mathbf{R}'^{\top} \\
&= \mathbf{J}_l^{-1}(\mathrm{Log}(\mathbf{R}^{\top}\mathbf{R}'))\mathbf{R}^{\top} = -(56). \qquad (57)
\end{aligned}
$$

REFERENCES

[1] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.

[2] A. Geiger, F. Moosmann, Ö. Car, and B. Schuster, "Automatic camera and range sensor calibration using a single shot," in *2012 IEEE International Conference on Robotics and Automation*, 2012, pp. 3936–3943.

[3] J. Peršić, I. Marković, and I. Petrović, "Extrinsic 6DoF calibration of a radar–LiDAR–camera system enhanced by radar cross section estimates evaluation," *Robotics and Autonomous Systems*, vol. 114, pp. 217–230, 2019.

[4] D. Scaramuzza, A. Harati, and R. Siegwart, "Extrinsic self calibration of a camera and a 3D laser range finder from natural scenes," in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2007, pp. 4164–4169.

[5] J. Levinson and S. Thrun, "Automatic online calibration of cameras and lasers," in *Robotics: Science and Systems*, 06 2013.

[6] J. Lv, J. Xu, K. Hu, Y. Liu, and X. Zuo, "Targetless calibration of LiDAR-IMU system based on continuous-time batch estimation," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 9968–9975.

[7] T. Scott, A. A. Morye, P. Piniés, L. M. Paz, I. Posner, and P. Newman, "Choosing a time and place for calibration of lidar-camera systems," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 2016, pp. 4349–4356.

[8] A. Censi, A. Franchi, L. Marchionni, and G. Oriolo, "Simultaneous calibration of odometry and sensor parameters for mobile robots," *IEEE Transactions on Robotics*, vol. 29, no. 2, pp. 475–492, 2013.

[9] R. Ishikawa, T. Oishi, and K. Ikeuchi, "LiDAR and camera calibration using motions estimated by sensor fusion odometry," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 7342–7349.

[10] B. Della Corte, H. Andreasson, T. Stoyanov, and G. Grisetti, "Unified motion-based calibration of mobile multi-sensor platforms with time delay estimation," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 902–909, 2019.

[11] S. Schneider, T. Luettel, and H.-J. Wuensche, "Odometry-based online extrinsic sensor calibration," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013, pp. 1287–1292.

[12] J. Brookshire and S. Teller, "Extrinsic calibration from per-sensor egomotion," 07 2012.

[13] R. Tsai and R. Lenz, "Real time versatile robotics hand/eye calibration using 3D machine vision," in *Proceedings. 1988 IEEE International Conference on Robotics and Automation*, 1988, pp. 554–561 vol.1.

[14] K. H. Strobl and G. Hirzinger, "Optimal hand-eye calibration," in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006, pp. 4647–4653.

[15] C. Park, P. Moghadam, S. Kim, S. Sridharan, and C. Fookes, "Spatiotemporal camera-LiDAR calibration: A targetless and structureless approach," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1556–1563, 2020.

[16] L. Heng, B. Li, and M. Pollefeys, "Camodocal: Automatic intrinsic and extrinsic calibration of a rig with multiple generic cameras and odometry," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013, pp. 1793–1800.

[17] Y. Shiu and S. Ahmad, "Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form AX=XB," *IEEE Transactions on Robotics and Automation*, vol. 5, no. 1, pp. 16–29, 1989.

[18] F. Park and B. Martin, "Robot sensor calibration: solving AX=XB on the Euclidean group," *IEEE Transactions on Robotics and Automation*, vol. 10, no. 5, pp. 717–721, 1994.

[19] J. C. K. Chou and M. Kamel, "Finding the position and orientation of a sensor on a robot manipulator using quaternions," *The International Journal of Robotics Research*, vol. 10, no. 3, pp. 240–254, 1991.

[20] R. Horaud and F. Dornaika, "Hand-eye calibration," *The International Journal of Robotics Research*, vol. 14, no. 3, pp. 195–210, 1995.

[21] K. Daniilidis, "Hand-eye calibration using dual quaternions," *The International Journal of Robotics Research*, vol. 18, no. 3, pp. 286–298, 1999.

[22] M. Horn, T. Wodtko, M. Buchholz, and K. Dietmayer, "Online extrinsic calibration based on per-sensor ego-motion using dual quaternions," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 982–989, 2021.

[23] M. Giamou, Z. Ma, V. Peretroukhin, and J. Kelly, "Certifiably globally optimal extrinsic calibration from per-sensor egomotion," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 367–374, 2019.

[24] K. Huang and C. Stachniss, "Extrinsic multi-sensor calibration for mobile robots using the Gauss-Helmert model," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, pp. 1490–1496.

[25] M. K. Ackerman and G. S. Chirikjian, "A probabilistic solution to the AX=XB problem: Sensor calibration without correspondence," in *Geometric Science of Information*, F. Nielsen and F. Barbaresco, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 693–701.

[26] M. K. Ackerman, A. Cheng, and G. Chirikjian, "An information-theoretic approach to the correspondence-free AX=XB sensor calibration problem," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, 2014, pp. 4893–4899.

[27] Q. Ma, H. Li, and G. S. Chirikjian, "New probabilistic approaches to the AX = XB hand-eye calibration without correspondence," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 2016, pp. 4365–4371.

[28] H. Nguyen and Q.-C. Pham, "On the covariance of $\boldsymbol{X}$ in $\boldsymbol{AX} = \boldsymbol{XB}$," *IEEE Transactions on Robotics*, vol. 34, no. 6, pp. 1651–1658, 2018.

[29] N. Andreff, R. Horaud, and B. Espiau, "Robot hand-eye calibration using structure-from-motion," *The International Journal of Robotics Research*, vol. 20, no. 3, pp. 228–248, 2001.

[30] J. Heller, M. Havlena, A. Sugimoto, and T. Pajdla, "Structure-from-motion based hand-eye calibration using l_∞ minimization," in *The IEEE/CVF Conference on Computer Vision and Pattern Recognition 2011*, 2011, pp. 3497–3503.

[31] Wei, Li, Naiguang, Lu, Mingli, Dong, and Xiaoping, Lou, "Calibration-free robot-sensor calibration approach based on second-order cone programming," *MATEC Web Conf.*, vol. 173, p. 02005, 2018.

[32] J. Schmidt, F. Vogt, and H. Niemann, "Calibration–free hand–eye calibration: A structure–from–motion approach," in *Pattern Recognition*, W. G. Kropatsch, R. Sablatnig, and A. Hanbury, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 67–74.

[33] E. Wise, M. Giamou, S. Khoubyarian, A. Grover, and J. Kelly, "Certifiably optimal monocular hand-eye calibration," in *2020 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, 2020, pp. 271–278.

[34] T. Wodtko, M. Horn, M. Buchholz, and K. Dietmayer, "Globally optimal multi-scale monocular hand-eye calibration using dual quaternions," in *2021 International Conference on 3D Vision (3DV)*, 2021, pp. 249–257.

[35] B. Wang, Z. Zhao, Y. Chen, and J. Yu, "A novel robust point cloud fitting algorithm based on nonlinear Gauss–Helmert model," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–12, 2023.

[36] S. Vogel, D. Ernst, I. Neumann, and H. Alkhatib, "Recursive Gauss-Helmert model with equality constraints applied to the efficient system calibration of a 3D laser scanner," *Journal of Applied Geodesy*, vol. 16, no. 1, pp. 37–57, 2022.

[37] T. Dang, C. Hoffmann, and C. Stiller, "Continuous stereo self-calibration by camera parameter tracking," *IEEE Transactions on Image Processing*, vol. 18, no. 7, pp. 1536–1550, 2009.

[38] Z. Wang, H. Shen, H. Du, J. Zhou, and X. Jin, "Research on calibration method of extrinsic parameters of lidar and camera carried by UAV," in *2022 4th International Conference on Data-driven Optimization of Complex Systems (DOCS)*, 2022, pp. 1–6.

[39] M. Ulrich and M. Hillemann, "Generic hand–eye calibration of uncertain robots," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 11 060–11 066.

[40] ——, "Uncertainty-aware hand–eye calibration," *IEEE Transactions on Robotics*, pp. 1–19, 2023.

[41] T. D. Barfoot, *State Estimation for Robotics*. Cambridge University Press, 2017.

[42] J. Solà, J. Deray, and D. Atchuthan, "A micro Lie theory for state estimation in robotics," 2021.

[43] A. Ruszczyński, *Nonlinear Optimization*. Princeton University Press, 2006.

[44] B. Wang, Z. Zhao, S. Wang, J. Yu, and Y. Chen, "Robust LS-VCE for the nonlinear gauss–helmert model: Case studies for point cloud fitting and geodetic symmetric transformation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1–13, 2024.

[45] K.-R. Koch, "Outlier detection for the nonlinear gauss helmert model with variance components by the expectation maximization algorithm," *Journal of Applied Geodesy*, vol. 8, no. 3, 2014.

[46] M. Čolaković Bencerić, J. Peršić, I. Marković, and I. Petrović, "On hand-eye calibration via on-manifold Gauss-Newton optimization," in *Intelligent Autonomous Systems 17*, 2023, pp. 378–391.

[47] Z. Zhang and D. Scaramuzza, "A tutorial on quantitative trajectory evaluation for visual(-inertial) odometry," in *IEEE/RSJ Int. Conf. Intell. Robot. Syst. (IROS)*, 2018.

[48] D. Schubert, T. Goll, N. Demmel, V. Usenko, J. Stueckler, and D. Cremers, "The TUM VI benchmark for evaluating visual-inertial odometry," in *International Conference on Intelligent Robots and Systems (IROS)*, October 2018.

[49] J. Engel, V. Koltun, and D. Cremers, "Direct sparse odometry," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 3, pp. 611–625, 2018.