A Certifiably Correct Algorithm for Generalized Robot-World and Hand-Eye Calibration

Emmett Wise[1], Pushyami Kaveti[2], Qilong Cheng[1], Wenhao Wang[1], Hanumant Singh[2], Jonathan Kelly[1], David M. Rosen[2], and Matthew Giamou[3]

Abstract

Automatic extrinsic sensor calibration is a fundamental problem for multi-sensor platforms. Reliable and general-purpose solutions should be computationally efficient, require few assumptions about the structure of the sensing environment, and demand little effort from human operators. In this work, we introduce a fast and certifiably globally optimal algorithm for solving a generalized formulation of the *robot-world and hand-eye calibration* (RWHEC) problem. The formulation of RWHEC presented is "generalized" in that it supports the simultaneous estimation of multiple sensor and target poses, and permits the use of monocular cameras that, alone, are unable to measure the scale of their environments. In addition to demonstrating our method's superior performance over existing solutions, we derive novel identifiability criteria and establish *a priori* guarantees of global optimality for problem instances with bounded measurement errors. We also introduce a complementary Lie-algebraic local solver for RWHEC and compare its performance with our global method and prior art. Finally, we provide a free and open-source implementation of our algorithms and experiments.

Keywords

Hand-eye and robot-world calibration, parameter identification, convex optimization, certifiable estimation

1 Introduction

Calibration is an essential but often painful process when working with common sensors for robot perception. In particular, *extrinsic* calibration refers to the problem of finding the spatial transformations between multiple sensors rigidly mounted to a fixed or mobile sensing platform. Existing approaches vary in terms of the type and number of sensors involved, assumptions about the robot's motion or the geometry of the environment, and operator involvement or expertise. Critically, extrinsic calibration errors can have catastrophic consequences for downstream perception tasks that rely on fusion of data from multiple sensors. For example, an autonomous vehicle using multiple cameras for simultaneous localization and mapping (SLAM) must accurately fuse images from different cameras into a coherent model of the world.

In this work, we focus on the very general and widespread *robot-world and hand-eye calibration* (RWHEC) formulation of extrinsic calibration (Zhuang et al. 1994). The RWHEC problem can be adapted to a wide variety of sensor configurations, and conveniently distills extrinsic calibration down to estimation of two rigid transformations represented as elements $\mathbf{X}$ and $\mathbf{Y}$ of the special Euclidean group $\mathrm{SE}(3)$, that define the kinematic relationship depicted in Figure 1. The matrix equations

$$\mathbf{A}_i \mathbf{X} = \mathbf{Y} \mathbf{B}_i, \ i = 1, \dots, N, \tag{1}$$

are formed from noisy measurements $\mathbf{A}_i, \mathbf{B}_i \in \mathrm{SE}(3)$. RWHEC is named after its original application to a robot manipulator ("hand") holding a camera ("eye"), but it can be applied to any sensor and target[1] configuration forming the kinematic loop shown in Figure 1.

Existing calibration procedures are error prone, especially when used by operators without sufficient expertise. These procedures often require that the operator excite

[1]Space & Terrestrial Autonomous Robotic Systems Laboratory, University of Toronto Institute for Aerospace Studies, Toronto, ON, Canada
[2]Institute for Experiential Robotics, Northeastern University, Boston, MA, USA
[3]Autonomous Robotics and Convex Optimization Laboratory, McMaster University, Hamilton, ON, Canada

Corresponding author:
Matthew Giamou, McMaster University, 1280 Main St W, Hamilton, ON L8S 4L8, Canada.
Email: giamoum@mcmaster.ca

a sensor platform with particular motions, or carefully select initial parameters close enough to the true solution. Without awareness of these idiosyncrasies, the optimizer for a calibration procedure may fail to converge to a critical point, or return a locally optimal solution that is inferior to the global minimizer. Often, the only indicator of inaccurate calibration parameters is the failure of downstream algorithms, which can place nearby people in danger and damage the robot or other infrastructure. To avoid these potentially catastrophic perception failures, end-users without significant expertise need calibration algorithms that automatically *certify* the global optimality of their solution. To this end, we present the following major contributions:

1. the first certifiably globally optimal solver for a general formulation of multi-sensor extrinsic calibration;

2. the first theoretical analysis of parameter identifiability for multi-sensor robot-world and hand-eye calibration; and

3. a free and open-source implementation of our method and experiments.[2]

We begin by surveying the extensive literature on robot-world and hand-eye calibration (RWHEC) in Section 2, followed by a summary of our mathematical notation in Section 3. We develop a detailed description of our generalized formulation of RWHEC in Section 4. Section 5 describes a convex relaxation of RWHEC that leads to a semidefinite program (SDP), while Section 6 presents identifiability criteria in the form of geometric constraints on sensor measurements and motion. Section 7 proves that our SDP relaxations of identifiable problems with sufficiently low noise are guaranteed to be tight, enabling the extraction of the global minimum to the original nonconvex problem. In Section 8 and Section 9, we experimentally demonstrate the superior properties of our RWHEC method as compared with existing solvers on synthetic and real-world data, respectively. Finally, Section 10 discusses remaining challenges and promising avenues for future work.

2 Related Work

We begin our survey of related work by reviewing RWHEC methods in Section 2.1. This involves defining three common categories of algorithms: separate, joint, and probabilistic.[3] The characteristics of all methods discussed are summarized in Table 1, which explicitly highlights the novelty of our solution. We also survey relevant applications of certifiably correct convex optimization methods to estimation problems in robotics and computer vision in Section 2.2.
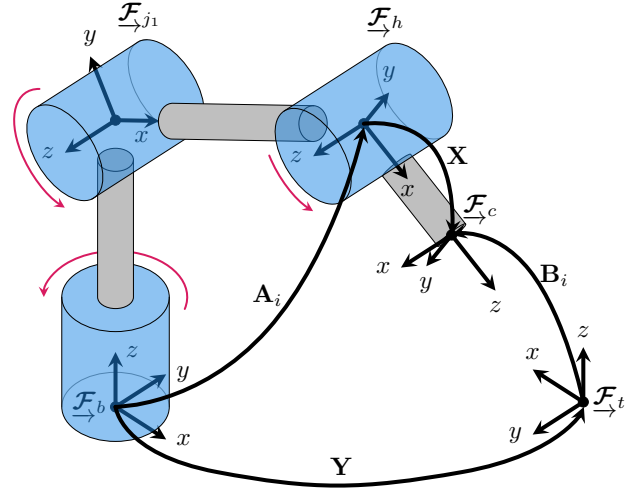

Figure 1. A diagram of the conventional application of RWHEC. In this application, the objective is to estimate $\mathbf{X} \in \mathrm{SE}(3)$, the transformation from the wrist of a robotic manipulator to a camera, and $\mathbf{Y} \in \mathrm{SE}(3)$, the transformation from the manipulator base to a known target. In this diagram, the base, joint 1, hand, camera, and target reference frames are labelled $\underrightarrow{\mathcal{F}}_b$, $\underrightarrow{\mathcal{F}}_{j_1}$, $\underrightarrow{\mathcal{F}}_h$, $\underrightarrow{\mathcal{F}}_c$, and $\underrightarrow{\mathcal{F}}_t$, respectively. The red arrows indicate the axis of joint rotation. At time i, we use the forward kinematics of the manipulator to compute $\mathbf{A}_i$, the transformation from the manipulator base to the wrist, and measure $\mathbf{B}_i$, a noisy estimate of the transformation from the target to the camera.

2.1 Robot-World and Hand-Eye Calibration

A subset of RWHEC methods solve for the rotation and translation components of $\mathbf{X}$ and $\mathbf{Y}$ separately. One simple method is to ignore nonlinear constraints on the rotation components of $\mathbf{X}$ and $\mathbf{Y}$ and solve the resulting overdetermined linear system in Equation (1) with a least squares approach. After projecting the estimated rotation matrices onto the rotation manifold $\mathrm{SO}(3)$, the translation components of $\mathbf{X}$ and $\mathbf{Y}$ are trivial to compute. This approach is used by Li et al. (2010) and extended to multiple $\mathbf{X}$ or $\mathbf{Y}$ variables in the generalized RWHEC formulation of Wang et al. (2022). The identifiability requirements of RWHEC are explored by Shah (2013), but to our knowledge have not been extended to the generalized case. Since these two-stage closed-form solvers ignore the coupling between the rotation and translation cost functions, they can corrupt translation estimates with error from noisy rotation measurements. Furthermore, since they all solve a linear relaxation of the true nonlinear RWHEC problem, these methods are best described as *approximation schemes*.

To decrease this sensitivity of RWHEC to measurement noise, some methods jointly estimate the translation and rotation components of $\mathbf{X}$ and $\mathbf{Y}$. Tabb and Yousef (2017) introduce a variety of joint RWHEC cost functions and parameterizations that are less sensitive to measurement noise than two-stage closed-form RWHEC schemes. Horn

Table 1. A non-exhaustive summary of relevant calibration methods for RWHEC algorithms. The columns track the following algorithmic properties: whether the method jointly estimates rotation and pose, whether a probabilistic problem formulation is employed, whether a post-hoc certificate of global optimality is produced by the algorithm, the algorithm's ability to model multi-frame problems, support for scale-free monocular sensors, and the 3D pose representation used.

Method	Joint	Probabilistic	Certifiable	Multiple $\mathbf{X}$s and/or $\mathbf{Y}$s	Scale-free	Poses
Li et al. (2010)	✗	✗	✗	✗	✗	DQ
Shah (2013)	✗	✗	✗	✗	✗	SE(3)
Wang et al. (2022)	✗	✗	✗	$\mathbf{X}$s or $\mathbf{Y}$s	✗	SE(3)
Tabb and Yousef (2017)	✓	✗	✗	✗	✗	Multiple
Horn et al. (2023)	✓	✗	✓	$\mathbf{X}$s or $\mathbf{Y}$s	✗	DQ
Evangelista et al. (2023)	✓	✗	✗	$\mathbf{X}$s	✗	SE(3)
Dornaika and Horaud (1998)	✓	✓	✗	✗	✗	Multiple
Strobl and Hirzinger (2006)	✓	✓	✗	✗	✗	SE(3)
Heller et al. (2014)	✓	✗	✓	✗	✗	Multiple
Ours	✓	✓	✓	$\mathbf{X}$s and $\mathbf{Y}$s	✓	SE(3)

et al. (2023) use a dual quaternion (DQ) representation for poses to cast the generalized RWHEC problem as a quadratically constrained quadratic program (QCQP) and use a methodology similar to ours to solve a convex relaxation of their problem formulation. However, we demonstrate in Section 8 that the unit DQ representation leads to poorer performance than the homogeneous transformation matrices we use. This is most likely because the set of unit quaternions $\mathbb{H}$ is a *double cover* of $\mathrm{SO}(3)$ in that the quadratic function $f : \mathbb{H} \to \mathrm{SO}(3)$ which converts quaternions to rotation matrices maps $\mathbf{q} \in \mathbb{H}$ to the same rotation matrix as $-\mathbf{q}$. As a result, for each measurement $\mathbf{A}_i$ or $\mathbf{B}_i$ appearing in Equation (1), there are two distinct dual quaternions that represent the same element of $\mathrm{SE}(3)$. Therefore, there are 2^{2N} DQ-based formulations of any RWHEC problem with N pairs of measurements, and their solutions may vary significantly in quality. Most existing work, including Horn et al. (2023), ignores this unfortunate consequence of representing rotations as unit quaternions and does not describe a principled or heuristic approach to lifting elements of $\mathrm{SO}(3)$ to one of two valid unit quaternions.

Evangelista et al. (2023) use a local nonlinear method to solve a variant of generalized RWHEC with an objective function formed from camera reprojection error. This is in contrast to our approach and the classical RWHEC framework, which abstracts out direct sensor measurements to work with poses. One advantage of using reprojection error is that it additionally allows the estimation of intrinsic camera parameters, which is outside the scope of this work.

Although joint estimation methods are more robust to measurement noise than two-stage closed-form solvers, the methods surveyed thus far do not use a rigorous probabilistic error model. As a result, these methods cannot model the effect of measurements with varying accuracy. Probabilistic RWHEC algorithms are introduced

by Dornaika and Horaud (1998) and Strobl and Hirzinger (2006). Dornaika and Horaud (1998) solve a nonlinear probabilistic formulation of RWHEC, but do not use an on-manifold method like the local solver presented in this work. Strobl and Hirzinger (2006) treat the RWHEC problem as an iteratively re-weighted nonlinear optimization problem, where the translation and rotation errors are corrupted by zero-mean Gaussian noise. While both methods account for the probabilistic nature of the problem, neither can provide a certificate of optimality. The probabilistic framework of Ha (2023) considers anisotropic noise models for measurements and estimates the uncertainty of a solution obtained by a local iterative method. Additionally, uncertainty-aware methods leveraging the Gauss-Helmert model have been applied to the hand-eye calibration problem with unknown scale (Ulrich and Hillemann 2024; Čolaković-Bencerić et al. 2025), but this approach has not yet been applied to RWHEC.

The origins of our certifiably optimal approach to calibration lie in the formulation of RWHEC and hand-eye calibration as global polynomial optimization problems by Heller et al. (2014). They solve semidefinite programming (SDP) relaxations of their polynomial programs and explore multiple representations of $\mathrm{SE}(3)$, but they do not use a probabilistic framework and only solve a standard variant of RWHEC (see Table 1 for details).

2.2 Certifiably Correct Estimation

Convex SDP relaxations of quadratically constrained quadratic programs are a powerful tool for finding globally optimal solutions to geometric estimation problems in robotics and computer vision (Carlone et al. 2015; Rosen et al. 2021; Cifuentes et al. 2022). In addition to the pioneering work of Heller et al. (2014), our approach is heavily influenced by the SE-Sync algorithm introduced in

by Rosen et al. (2019). SE-Sync was the first efficient and certifiably optimal algorithm for simultaneous localization and mapping (SLAM), which, like generalized RWHEC, is formulated over many unknown pose variables. Rosen et al. (2019) reveal and exploit the smooth manifold structure of an SDP relaxation of a QCQP formulation of pose-graph SLAM with an MLE objective function. This approach spawned a host of solutions to spatial perception problems including point cloud registration (Yang et al. 2021), multiple variants of localization and mapping (Holmes and Barfoot 2023; Fan et al. 2020; Tian et al. 2021; Papalia et al. 2024; Dumbgen et al. 2023; Yu and Yang 2024), hand-eye calibration (Giamou et al. 2018; Wise et al. 2020; Wodtko et al. 2021), camera pose estimation (Garcia-Salguero et al. 2021; Zhao 2020), as well as research into tools and optimization methods for user-specified problems (Dümbgen et al. 2023; Rosen 2021; Yang and Carlone 2020).

An alternative approach for obtaining global optima for noisy geometric estimation problems is outlined by Wu et al. (2022). In contrast to using convex SDP relaxations, Wu et al. (2022) use a Gröbner basis method to solve polynomial optimization problems over a single pose variable. This symbolic method was also extended to problems with unknown scale, including hand-eye calibration with a monocular camera (Xue et al. 2025). While promising, this approach has not been demonstrated to scale efficiently to problems with many unknown pose variables. Similarly, a branch-and-bound approach is applied to find global optima of hand-eye calibration problems in Heller et al. (2012) and Heller et al. (2016), but these methods do not scale efficiently to problems with many sensors.

In this work, we extend the multi-frame RWHEC formulation introduced by Wang et al. (2022) to the scale-free case involving monocular sensors, and we form a joint objective function within a maximum likelihood estimation (MLE) framework. This leads to an estimation problem over a graph similar in nature to SE-Sync (Rosen et al. 2019), but whose feasible set lacks some of the nice geometric structure required to employ the fast low-dimensional Riemannian optimization algorithms used in this prior work. Finally, we prove fundamental identifiability and global optimality theorems for the multi-frame RWHEC problem first introduced by Wang et al. (2022) and extended by Section 4.5.

3 Notation

In this paper, lowercase Latin and Greek characters represent scalar variables. We reserve lowercase and uppercase boldface characters for vectors and matrices, respectively. For an integer $N > 0$, $[N]$ denotes the index set $\{1, \ldots, N\}$. The space of $n \times n$ symmetric and symmetric positive semidefinite (PSD) matrices are written as $\mathbb{S}^n$ and $\mathbb{S}^n_+$, respectively, and we also use $\mathbf{A} \succeq \mathbf{B}$ ($\mathbf{A} \succ \mathbf{B}$) to indicate that $\mathbf{A} - \mathbf{B}$ is PSD (positive definite). For a general matrix $\mathbf{A}$, we use $\mathbf{A}^\dagger$ to indicate the Moore-Penrose pseudoinverse. The (right) nullspace or kernel of $\mathbf{A}$ is written as $\ker(\mathbf{A})$, and its range is written as $\mathcal{R}(\mathbf{A})$.

For a directed graph $\vec{\mathcal{G}} = (\mathcal{V}, \vec{\mathcal{E}})$, we write each directed edge as $e = (i, j)$ and say that e *leaves* vertex $i \in \mathcal{V}$ and *enters* vertex $j \in \mathcal{V}$. Additionally, $\delta^-(v)$ denotes the set of incident edges leaving vertex $v \in \mathcal{V}$, and $\delta^+(v)$ denotes the set of incident edges entering v.

A right-handed reference frame is written as $\underrightarrow{\mathcal{F}}$ (Barfoot 2024). The translation from $\underrightarrow{\mathcal{F}}_a$ to $\underrightarrow{\mathcal{F}}_b$ described in $\underrightarrow{\mathcal{F}}_a$ is written as $\mathbf{p}_a^{ba} \in \mathbb{R}^3$. The matrix $\mathbf{R}_{ab} \in \mathrm{SO}(3)$ represents the rotation taking coordinates expressed in $\underrightarrow{\mathcal{F}}_b$ to $\underrightarrow{\mathcal{F}}_a$. Likewise, the rigid special Euclidean transformation from $\underrightarrow{\mathcal{F}}_b$ to $\underrightarrow{\mathcal{F}}_a$ is

$$\mathbf{T}_{ab} = \begin{bmatrix} \mathbf{R}_{ab} & \mathbf{p}_a^{ba} \\ \mathbf{0}^\top & 1 \end{bmatrix} \in \mathrm{SE}(3). \tag{2}$$

The skew-symmetric operator $(\cdot)^\wedge$ acts on the vector $\mathbf{p} \in \mathbb{R}^3$ such that

$$\mathbf{p}^\wedge = \begin{bmatrix} 0 & -p_3 & p_2 \\ p_3 & 0 & -p_1 \\ -p_2 & p_1 & 0 \end{bmatrix}, \tag{3}$$

and the operator $(\cdot)^\vee$ denotes its inverse. The standard Kronecker product of $\mathbf{A}$ and $\mathbf{B}$ is written as $\otimes$, and the Kronecker sum $\oplus$ is

$$\mathbf{A} \oplus \mathbf{B} = \mathbf{A} \otimes \mathbf{I} + \mathbf{I} \otimes \mathbf{B}. \tag{4}$$

The function $\mathrm{Diag}\,(\mathbf{A}_1, \ldots, \mathbf{A}_N)$ creates a block-diagonal matrix with its ordered matrix arguments on the diagonal. A vector $\mathbf{x} \sim \mathcal{N}\,(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is Gaussian distributed with a mean of $\boldsymbol{\mu} \in \mathbb{R}^n$ and covariance of $\boldsymbol{\Sigma} \in \mathbb{S}^n_+$. Finally, $\mathrm{Lang}(\mathbf{M}, \kappa)$ denotes a Langevin distribution over $\mathrm{SO}(3)$ with mode $\mathbf{M} \in \mathrm{SO}(3)$ and concentration $\kappa \geq 0$.

4 Generalized Robot-World and Hand-Eye Calibration

In Section 4.1, we review the geometric constraints of the robot-world and hand-eye calibration problem. In Section 4.2, we formulate RWHEC as maximum likelihood estimation. In Section 4.3, we extend our problem formulation to support monocular cameras observing targets of unknown size. In Section 4.4, we convert our calibration problems into QCQPs in standard form. In Section 4.5, we extend our probabilistic formulation to calibrate an arbitrary number of decision variables (i.e., multiple $\mathbf{X}$s and/or $\mathbf{Y}$s). Finally, in Section 4.6, we present a dimensionality reduction strategy for our formulation.

4.1 Geometric Constraints

While the robot-world and hand-eye geometric constraints apply to a large set of calibration problems (e.g., multiple cameras on a mobile manipulator or fixed cameras tracking a known target (Wang et al. 2022)), for convenience we will begin, without loss of generality, with terminology appropriate for calibrating a camera mounted on the 'hand' of a robotic manipulator as shown in Figure 1. Fix reference frames $\mathcal{F}_b, \mathcal{F}_h, \mathcal{F}_t, \mathcal{F}_c$ to the manipulator base, manipulator hand, target, and camera, respectively. Using joint encoder data and kinematic parameters of the robot arm, we are able to estimate the transform $\mathbf{T}_{bh}$. Additionally, a camera observing a target of known scale enables estimation of the camera pose relative to the target, $\mathbf{T}_{tc}$. At each discrete point in time indexed by $i \in [N]$, these two measurements are related by

$$\mathbf{T}_{bh}(i)\mathbf{T}_{hc} = \mathbf{T}_{bt}\mathbf{T}_{tc}(i). \tag{5}$$

In the RWHEC problem, we wish to estimate $\mathbf{T}_{hc}$ and $\mathbf{T}_{bt}$, which we assume are static extrinsic transformation parameters. We introduce simpler symbols

$$\begin{aligned}
\mathbf{A}_i &\triangleq \mathbf{T}_{bh}(i) \\
\mathbf{X} &\triangleq \mathbf{T}_{hc} \\
\mathbf{Y} &\triangleq \mathbf{T}_{bt} \\
\mathbf{B}_i &\triangleq \mathbf{T}_{tc}(i),
\end{aligned} \tag{6}$$

for our matrices, resulting in the following familiar RWHEC constraint:

$$\mathbf{A}_i\mathbf{X} = \mathbf{Y}\mathbf{B}_i. \tag{7}$$

We may separate the rotational and translational components of Equation (7) according to:

$$\mathbf{R}_{\mathbf{A}_i}\mathbf{R}_{\mathbf{X}} = \mathbf{R}_{\mathbf{Y}}\mathbf{R}_{\mathbf{B}_i}, \tag{8a}$$

$$\mathbf{R}_{\mathbf{A}_i}\mathbf{t}_{\mathbf{X}} + \mathbf{t}_{\mathbf{A}_i} = \mathbf{R}_{\mathbf{Y}}\mathbf{t}_{\mathbf{B}_i} + \mathbf{t}_{\mathbf{Y}}, \tag{8b}$$

where $\mathbf{R}_{\mathbf{A}_i}, \mathbf{R}_{\mathbf{X}}, \mathbf{R}_{\mathbf{Y}}, \mathbf{R}_{\mathbf{B}_i}$ and $\mathbf{t}_{\mathbf{A}_i}, \mathbf{t}_{\mathbf{X}}, \mathbf{t}_{\mathbf{Y}}, \mathbf{t}_{\mathbf{B}_i}$ are the rotation and translation components of $\mathbf{A}_i, \mathbf{X}, \mathbf{Y}, \mathbf{B}_i$, respectively.

Note that alternative interpretations of the $\mathbf{A}\mathbf{X} = \mathbf{Y}\mathbf{B}$ equations are possible. For example, consider defining $\mathbf{A}_i$ and $\mathbf{B}_i$ as the inverse of their assignments in Equation (6):

$$\begin{aligned}
\mathbf{A}_i &\triangleq \mathbf{T}_{hb}(i) = \mathbf{T}_{bh}^{-1}(i), \\
\mathbf{B}_i &\triangleq \mathbf{T}_{ct}(i) = \mathbf{T}_{tc}^{-1}(i).
\end{aligned} \tag{9}$$

The kinematic chain in Equation (5) becomes

$$\mathbf{T}_{hb}(i)\mathbf{T}_{bt} = \mathbf{T}_{hc}\mathbf{T}_{ct}(i), \tag{10}$$

which swaps the definitions of $\mathbf{X}$ and $\mathbf{Y}$. Equation (10) corresponds to the setup used for our experiments in

Section 9, which is illustrated in Figure 6. This flexibility affords practitioners some freedom in defining variables for their RWHEC problem, but care must be taken to conform to the noise models discussed in the next section.[4]

4.2 Maximum Likelihood Estimation

Using Equations (8a) and (8b), we can formulate an MLE problem for the unknown states $\mathbf{X}$ and $\mathbf{Y}$. We use $\tilde{\mathbf{x}}$ to indicate a noisy measurement of the true value $\underline{\mathbf{x}}$ of a quantity of interest $\mathbf{x}$. Our probabilistic formulation of RWHEC requires the following assumptions:

1. $\tilde{\mathbf{R}}_{\mathbf{A}_i} = \underline{\mathbf{R}}_{\mathbf{A}_i}$ and $\tilde{\mathbf{t}}_{\mathbf{A}_i} = \underline{\mathbf{t}}_{\mathbf{A}_i} \ \forall \ i = 1, \dots, N$ (i.e., they are noiseless measurements);

2. $\tilde{\mathbf{t}}_{\mathbf{B}_i} = \underline{\mathbf{t}}_{\mathbf{B}_i} + \boldsymbol{\epsilon}_i$, where $\boldsymbol{\epsilon}_i \sim \mathcal{N}\left(0, \sigma_i{}^2\mathbf{I}\right)$ and σ_i is the standard deviation of the translation component of measurements $\mathbf{B}_i \ \forall \ i = 1, \dots, N$; and

3. $\tilde{\mathbf{R}}_{\mathbf{B}_i} = \underline{\mathbf{R}}_{\mathbf{B}_i}\mathbf{R}_{\boldsymbol{\epsilon}_i}$, where $\mathbf{R}_{\boldsymbol{\epsilon}_i} \sim \text{Lang}(\mathbf{I}, \kappa_i)$, where κ_i is the concentration of the rotations of the measurement $\mathbf{B}_i \ \forall \ i = 1, \dots, N$.

These assumptions are consistent with convention and accurately approximate many practical scenarios (Ha 2023, Section IV.C), including the classic "eye-in-hand" setup shown in Figure 1, where measurement $\mathbf{A}_i$ is computed with a calibrated manipulator's forward kinematics and $\mathbf{B}_i$ is a function of a noisy camera measurement of a target. In order to maintain consistency with the majority of robot-world and hand-eye calibration literature, we slightly abuse our notation and use undecorated matrices $\mathbf{A}_i$ and $\mathbf{B}_i$. More precisely, we overload the idealized expressions in Equation (6) and henceforth write

$$\begin{aligned}
\mathbf{A}_i &\triangleq \underline{\mathbf{T}}_{bh}(i) = \begin{bmatrix} \mathbf{R}_{\mathbf{A}_i} & \mathbf{t}_{\mathbf{A}_i} \\ \mathbf{0}^\top & 1 \end{bmatrix}, \\
\mathbf{B}_i &\triangleq \tilde{\mathbf{T}}_{tc}(i) = \begin{bmatrix} \tilde{\mathbf{R}}_{\mathbf{B}_i} & \tilde{\mathbf{t}}_{\mathbf{B}_i} \\ \mathbf{0}^\top & 1 \end{bmatrix}.
\end{aligned} \tag{11}$$

With the introduction of noise into $\mathbf{B}_i$ in Equation (11), the constraints in Equation (7) no longer hold exactly for the true values of $\mathbf{X}$ and $\mathbf{Y}$. Therefore, the best estimate for $\mathbf{X}$ and $\mathbf{Y}$ is the solution to the following maximum likelihood problem:

$$\max_{\mathbf{X},\mathbf{Y} \in \text{SE}(3)} p(\{\mathbf{B}_1, \dots, \mathbf{B}_N\}|\mathbf{X}, \mathbf{Y}). \tag{12}$$

Since the random variables $\tilde{\mathbf{R}}_{\mathbf{B}_i}$ and $\tilde{\mathbf{t}}_{\mathbf{B}_i}$ are assumed to be distributed independently of one another and the random variables in $\mathbf{B}_j$ for all $j \neq i$, taking the negative log-likelihood of the objective in Equation (12) gives us the following equivalent minimization problem:

$$\min_{\mathbf{X},\mathbf{Y} \in \text{SE}(3)} -\sum_{i=1}^{N} (\log(p(\tilde{\mathbf{t}}_{\mathbf{B}_i}|\mathbf{X}, \mathbf{Y})) + \log(p(\tilde{\mathbf{R}}_{\mathbf{B}_i}|\mathbf{X}, \mathbf{Y}))). \tag{13}$$

The conditional log-likelihood of $\tilde{\mathbf{t}}_{\mathbf{B}_i}$ is

$$
\begin{aligned}
\log p(\tilde{\mathbf{t}}_{\mathbf{B}_i}|\mathbf{X},\mathbf{Y}) = \\
-\frac{1}{2\sigma_i^2}\left\|\mathbf{R}_{\mathbf{Y}}^\top(\underline{\mathbf{R}}_{\mathbf{A}_i}\mathbf{t}_{\mathbf{X}} + \underline{\mathbf{t}}_{\mathbf{A}_i} - \mathbf{t}_{\mathbf{Y}}) - \tilde{\mathbf{t}}_{\mathbf{B}_i}\right\|_2^2.
\end{aligned} \quad (14)
$$

Left–multiplying the Euclidean norm's argument with $\mathbf{R}_{\mathbf{Y}}$ makes the log-likelihood quadratic in decision variables $\mathbf{X}$ and $\mathbf{Y}$:

$$
\begin{aligned}
\log p(\tilde{\mathbf{t}}_{\mathbf{B}_i}|\mathbf{X},\mathbf{Y}) = \\
-\frac{1}{2\sigma_i^2}\left\|\underline{\mathbf{R}}_{\mathbf{A}_i}\mathbf{t}_{\mathbf{X}} + \underline{\mathbf{t}}_{\mathbf{A}_i} - \mathbf{t}_{\mathbf{Y}} - \mathbf{R}_{\mathbf{Y}}\tilde{\mathbf{t}}_{\mathbf{B}_i}\right\|_2^2,
\end{aligned} \quad (15)
$$

which we recognize as a weighted error function for the translation constraint in Equation (8b). Similarly, the log-likelihood of $p(\tilde{\mathbf{R}}_{\mathbf{B}_i}|\mathbf{X},\mathbf{Y})$ is (Rosen et al. 2019):

$$
\begin{aligned}
\log p(\tilde{\mathbf{R}}_{\mathbf{B}_i}|\mathbf{X},\mathbf{Y}) = \\
-c(\kappa_i) - \frac{\kappa_i}{2}\left\|\underline{\mathbf{R}}_{\mathbf{A}_i}\mathbf{R}_{\mathbf{X}} - \mathbf{R}_{\mathbf{Y}}\tilde{\mathbf{R}}_{\mathbf{B}_i}\right\|_{\mathrm{F}}^2 + 3,
\end{aligned} \quad (16)
$$

which is a weighted error function for the rotation constraint in Equation (8a). Our MLE problem is the following QCQP:

Problem 1. *Maximum Likelihood Estimation for RWHEC*

$$
\min_{\substack{\mathbf{R}_{\mathbf{X}},\mathbf{R}_{\mathbf{Y}}\in\mathrm{SO}(3) \\ \mathbf{t}_{\mathbf{X}},\mathbf{t}_{\mathbf{Y}}\in\mathbb{R}^3, s^2=1}} J_{\mathbf{t}} + J_{\mathbf{R}}, \quad (17)
$$

where

$$
J_{\mathbf{t}} \triangleq \frac{1}{2}\sum_{i=1}^N \frac{1}{\sigma_i^2}\left\|\underline{\mathbf{R}}_{\mathbf{A}_i}\mathbf{t}_{\mathbf{X}} + s\underline{\mathbf{t}}_{\mathbf{A}_i} - \mathbf{t}_{\mathbf{Y}} - \mathbf{R}_{\mathbf{Y}}\tilde{\mathbf{t}}_{\mathbf{B}_i}\right\|_2^2 \quad (18a)
$$

$$
J_{\mathbf{R}} \triangleq \frac{1}{2}\sum_{i=1}^N \kappa_i\left\|\underline{\mathbf{R}}_{\mathbf{A}_i}\mathbf{R}_{\mathbf{X}} - \mathbf{R}_{\mathbf{Y}}\tilde{\mathbf{R}}_{\mathbf{B}_i}\right\|_{\mathrm{F}}^2. \quad (18b)
$$

Note that to ensure that the terms of the objective are quadratic or constant, we have homogenized Equation (18a) with the quadratically constrained variable $s^2 = 1$. Homogenization simplifies the SDP relaxation employed in Section 5 and the analysis in Section 7. Additionally, we ensure that Problem 1 is a QCQP by using the following constraints for each $\mathrm{SO}(3)$ variable (Tron et al. 2015):

$$
\mathbf{R}\mathbf{R}^\top = \mathbf{I}, \quad (19a)
$$

$$
\mathbf{R}^\top\mathbf{R} = \mathbf{I}, \quad (19b)
$$

$$
\mathbf{R}_{\sigma_i(1)} \times \mathbf{R}_{\sigma_i(2)} = \mathbf{R}_{\sigma_i(3)}, \ \forall i \in \{1,2,3\}, \quad (19c)
$$

where $\mathbf{R}_i$ denotes the ith column of $\mathbf{R}$, and $G \triangleq \{\sigma_i\}$ is the group of cyclic permutations of $\{1,2,3\}$.[5] For generic $\mathbf{R} \in \mathrm{SO}(d)$, the constraints (23) would include $\det(\mathbf{R}) =$

1. However, the determinant is a polynomial of degree d, which for $d = 3$ is cubic. For $\mathrm{SO}(3)$ in particular, we can enforce $\det(\mathbf{R}) = 1$ using the simpler set of quadratic constraints involving the cross-product operation appearing in Equation (19c).

4.3 Monocular Cameras

If the scale of a target observed by a monocular camera is unknown, we can use a scaled pose sensor abstraction to model measurements. Repeating the MLE derivation from Section 4.2, we replace the camera translation measurement model with

$$
\tilde{\mathbf{t}}_{\mathbf{B}_i} = \alpha\mathbf{R}_{\mathbf{Y}}^\top(\underline{\mathbf{R}}_{\mathbf{A}_i}\mathbf{t}_{\mathbf{X}} + \underline{\mathbf{t}}_{\mathbf{A}_i} - \mathbf{t}_{\mathbf{Y}}) + \boldsymbol{\epsilon}_i, \quad (20)
$$

where $\alpha \in \mathbb{R}$ is the unknown scale, $\boldsymbol{\epsilon}_i \sim \mathcal{N}\left(0, \sigma_i^2\mathbf{I}\right)$, and σ_i is the standard deviation of the unscaled translation estimate. Consequently, the monocular MLE problem is the following QCQP:

Problem 2. *Maximum Likelihood Estimation for Monocular RWHEC*

$$
\min_{\substack{\mathbf{R}_{\mathbf{X}},\mathbf{R}_{\mathbf{Y}}\in\mathrm{SO}(3) \\ \mathbf{t}_{\mathbf{X},\alpha},\mathbf{t}_{\mathbf{Y},\alpha}\in\mathbb{R}^3, \alpha\in\mathbb{R}}} J_{\mathbf{t},\alpha} + J_{\mathbf{R}}, \quad (21)
$$

where

$$
J_{\mathbf{t},\alpha} \triangleq \frac{1}{2}\sum_{i=1}^N \frac{1}{\sigma_i^2}\left\|\underline{\mathbf{R}}_{\mathbf{A}_i}\mathbf{t}_{\mathbf{X},\alpha} + \alpha\underline{\mathbf{t}}_{\mathbf{A}_i} - \mathbf{t}_{\mathbf{Y},\alpha} - \mathbf{R}_{\mathbf{Y}}\tilde{\mathbf{t}}_{\mathbf{B}_i}\right\|_2^2. \quad (22)
$$

The variables $\mathbf{t}_{\mathbf{X},\alpha} \triangleq \alpha\mathbf{t}_{\mathbf{X}}$ and $\mathbf{t}_{\mathbf{Y},\alpha} \triangleq \alpha\mathbf{t}_{\mathbf{Y}}$ have "absorbed" the scale parameter α and must be divided by α after solving Problem 2. To maintain our assumption from Section 4.2 that measurements $\mathbf{A}_i$ are noiseless, the measurements $\mathbf{B}_i$ are assumed to come from the monocular camera. Our introduction of the unknown scale parameter α has produced a naturally homogeneous QCQP that has replaced the homogenizing variable s used in Problem 1, which can now be interpreted as the special case of Problem 2 for known scale $\alpha = 1$. Therefore, in Section 4.4 we will deal solely with the monocular case in Problem 2.

4.4 Quadratically Constrained Quadratic Programming

Herein we convert Problem 2 to a standard QCQP form with a vectorized decision variable and constraints defined by real symmetric matrices. The state vector is

$$
\mathbf{x} \triangleq \begin{bmatrix} \mathbf{t}_{\mathbf{X},\alpha}^\top & \mathbf{t}_{\mathbf{Y},\alpha}^\top & \mathbf{r}_{\mathbf{X}}^\top & \mathbf{r}_{\mathbf{Y}}^\top & \alpha \end{bmatrix}^\top, \quad (23)
$$

where $\mathbf{r}_{\mathbf{X}} = \mathrm{vec}\left(\mathbf{R}_{\mathbf{X}}\right)$ and $\mathbf{r}_{\mathbf{Y}} = \mathrm{vec}\left(\mathbf{R}_{\mathbf{Y}}\right)$. Equation (23) allows us to write the rotation part of the objectives of

monocular RWHEC problems as

$$J_{\mathbf{R}} = \frac{1}{2} \sum_{i=1}^{N} \kappa_i \mathbf{x}^\top \mathbf{M}_{\mathbf{R}_i}^\top \mathbf{M}_{\mathbf{R}_i} \mathbf{x}, \qquad (24)$$

where

$$\mathbf{M}_{\mathbf{R}_i} \triangleq \begin{bmatrix} \mathbf{0}_{9\times6} & \mathbf{I} \otimes \underline{\mathbf{R}}_{\mathbf{A}_i} & -\tilde{\mathbf{R}}_{\mathbf{B}_i} \otimes \mathbf{I} & \mathbf{0}_{9\times1} \end{bmatrix}, \quad (25)$$

Equation (25) and many expressions to follow are obtained through a straightforward application of the column-major vectorization identity (Henderson and Searle 1981)

$$\mathrm{vec}\left(\mathbf{A}\mathbf{X}\mathbf{B}\right) = (\mathbf{B}^\top \otimes \mathbf{A})\mathrm{vec}\left(\mathbf{X}\right), \qquad (26)$$

where $\mathbf{A}$, $\mathbf{X}$, and $\mathbf{B}$ are any compatible matrices. The translation components of the monocular RWHEC problem's objective can now be written as

$$J_{\mathbf{t}} = \frac{1}{2} \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \mathbf{x}^\top \mathbf{M}_{\mathbf{t}_i}^\top \mathbf{M}_{\mathbf{t}_i} \mathbf{x}, \qquad (27)$$

where

$$\mathbf{M}_{\mathbf{t}_i} \triangleq \begin{bmatrix} \mathbf{R}_{\mathbf{A}_i} & -\mathbf{I} & \mathbf{0}_{3\times9} & -\tilde{\mathbf{t}}_{\mathbf{B}_i}^\top \otimes \mathbf{I} & \underline{\mathbf{t}}_{\mathbf{A}_i} \end{bmatrix}. \quad (28)$$

The objective function of Problem 2 is now completely described by a quadratic forms with associated symmetric matrix

$$\mathbf{Q} \triangleq \frac{1}{2} \sum_{i=1}^{N} \kappa_i \mathbf{M}_{\mathbf{R}_i}^\top \mathbf{M}_{\mathbf{R}_i} + \frac{1}{2} \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \mathbf{M}_{\mathbf{t}_i}^\top \mathbf{M}_{\mathbf{t}_i}. \quad (29)$$

Consequently, we can rewrite Problem 2 as

$$\min_{\mathbf{x}} \mathbf{x}^\top \mathbf{Q} \mathbf{x}, \\ \text{s.t. } \mathbf{R}_{\mathbf{X}}, \mathbf{R}_{\mathbf{Y}} \in \mathrm{SO}(3). \qquad (30)$$

4.5 Generalizing RWHEC to Multiple Sensors and Targets

We can extend robot-world and hand-eye calibration to robots employing more than one sensor, and calibration procedures involving more than one target (Wang et al. 2022). In this generalized form, we *jointly* estimate a collection of M hand-eye transformations $\mathbf{X}_1, \ldots, \mathbf{X}_M$ and P base-target transformations $\mathbf{Y}_1, \ldots, \mathbf{Y}_P$. As shown in Figure 2, we can model the set of available measurements as a bipartite directed graph $\vec{\mathcal{G}} = (\mathcal{V}, \vec{\mathcal{E}})$, where

$$\mathcal{V} \triangleq \mathcal{V}_{\mathbf{X}} \cup \mathcal{V}_{\mathbf{Y}} = [M + P], \qquad (31)$$

and each edge $e = (j, k) \in \vec{\mathcal{E}} \subseteq \mathcal{V}_{\mathbf{X}} \times \mathcal{V}_{\mathbf{Y}}$ represents a set of observations of target k by camera j. This multi-frame approach was introduced by Wang et al. (2022)

for problems like "multiple eye-in-hand" calibration where either $M = 1$ or $P = 1$.

The notation we use to describe generalized RWHEC is inspired by the elegant graph-theoretic treatment of pose SLAM in Rosen et al. (2019). One notable feature of our formulation is that all observations involving variables $\mathbf{X}_j$ and $\mathbf{Y}_k$ are associated with a *single* directed edge (j, k).[6] Therefore, each edge $e = (j, k) \in \vec{\mathcal{E}}$ is labelled with a *set* $\mathcal{D}_e$ of all $N_e \geq 1$ noisy observations involving unknown variables $\mathbf{X}_j$ and $\mathbf{Y}_k$:

$$\mathcal{D}_e \triangleq \{(\mathbf{A}_{e,i}, \mathbf{B}_{e,i}, \sigma_{e,i}, \kappa_{e,i}) \in \mathrm{SE}(3)^2 \times \mathbb{R}_+^2 \mid i \in [N_e]\}. \quad (32)$$

The problem graph $\vec{\mathcal{G}}$ summarizes the connection between $|\vec{\mathcal{E}}|$ coupled RWHEC subproblems of the form in Problem 1 or Problem 2. This graphical structure enables us to describe a joint RWHEC problem involving all unknown variables $\mathbf{X}_j$ and $\mathbf{Y}_k$ indexed by $\mathcal{V}$:

Problem 3. *Generalized Monocular RWHEC*

$$\min_{\substack{\mathbf{X}_j, \mathbf{Y}_k \in \mathrm{SE}(3), \\ \alpha \in \mathbb{R}}} \frac{1}{2} \sum_{(j,k) \in \vec{\mathcal{E}}} \sum_{i=1}^{N_{(j,k)}} J_{ijk}(\mathbf{X}_j, \mathbf{Y}_k, \alpha),$$

where

$$J_{ijk}(\mathbf{X}, \mathbf{Y}, \alpha) \triangleq \\ \frac{1}{\sigma_{(j,k),i}^2} \left\| \mathbf{R}_{\mathbf{A}_{(j,k),i}} \mathbf{t}_{\mathbf{X}} + \alpha \underline{\mathbf{t}}_{\mathbf{A}_{(j,k),i}} - \mathbf{t}_{\mathbf{Y}} - \mathbf{R}_{\mathbf{Y}} \tilde{\mathbf{t}}_{\mathbf{B}_{(j,k),i}} \right\|_2^2 \\ + \kappa_{(j,k),i} \left\| \mathbf{R}_{\mathbf{A}_{(j,k),i}} \mathbf{R}_{\mathbf{X}} - \mathbf{R}_{\mathbf{Y}} \tilde{\mathbf{R}}_{\mathbf{B}_{(j,k),i}} \right\|_{\mathrm{F}}^2. \quad (33)$$

It is worth noting that Problem 2 is a special case of Problem 3 for a graph $\vec{\mathcal{G}}$ with only two vertices ($M = 1 = P$). Additionally, we will refer to the special case of Problem 3 with known scale ($\alpha = 1$) as the *standard* RWHEC problem:

Problem 4. *Generalized Standard RWHEC*

$$\min_{\mathbf{X}_j, \mathbf{Y}_k \in \mathrm{SE}(3)} \frac{1}{2} \sum_{(j,k) \in \vec{\mathcal{E}}} \sum_{i=1}^{N_{(j,k)}} J_{ijk}(\mathbf{X}_j, \mathbf{Y}_k, 1).$$

For the remainder of this section, we will concern ourselves solely with generalized monocular RWHEC formulation in Problem 3. Our experiments in Sections 8 and 9 will deal with both the standard and monocular cases, but our identifiability analysis in Section 6 only applies to Problem 4.

While our problem formulation admits the use of heterogeneous measurement precisions, to ease notation in the sequel we will assume that all measurements associated with a single edge $e \in \vec{\mathcal{E}}$ have common rotational and
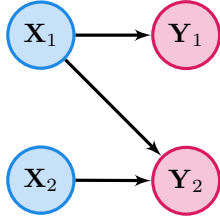
Figure 2. An example of the bipartite graph structure of a simple generalized RWHEC problem with multiple **X**s and **Y**s. Each edge corresponds to a set of measurements $\mathcal{D}_e$ in Equation (32) forming an instance of Problem 1 or Problem 2 involving $\mathbf{X}_j$ and $\mathbf{Y}_k$ for $e = (j, k)$.

translational precisions:[7]

$$\sigma_{e,i} = \sigma_e, \quad \kappa_{e,i} = \kappa_e \; \forall i \in [N_e]. \tag{34}$$

For the monocular case, the parameterization of unknown scale used in Equation (22) limits our multi-sensor extension to cases using either a single target, or targets with the same unknown scale α (e.g., the fiducial markers of equal size deployed in the experiments of Section 9). Our approach can be easily extended to support multiple unknown target scales by introducing target-specific scales α_i or using the conformal special orthogonal group $\mathrm{CSO}(3)$ employed by Yu and Yang (2024).[8]

We can simplify our notation by introducing several matrices structured by blocks mirroring the graph Laplacian (Rosen et al. 2019, Section 4.1). Let $L(\vec{\mathcal{G}}^\tau) \in \mathbb{S}^{3(M+P)}$ be the symmetric (3×3)-block-structured matrix of the form

$$L(\vec{\mathcal{G}}^\tau)_{jk} = \begin{cases} \sum_{e \in \delta^-(j)} \frac{1}{\sigma_e^2} N_e \mathbf{I} & j = k \\ -\frac{1}{\sigma_{(j,k)}^2} \sum_{i=1}^{N_{(j,k)}} \mathbf{R}_{\mathbf{A}_{(j,k),i}}^\top & (j, k) \in \vec{\mathcal{E}} \\ -\frac{1}{\sigma_{(k,j)}^2} \sum_{i=1}^{N_{(k,j)}} \mathbf{R}_{\mathbf{A}_{(k,j),i}} & (k, j) \in \vec{\mathcal{E}} \\ \mathbf{0}_{3 \times 3} & \text{otherwise.} \end{cases} \tag{35}$$

Similarly, let $L(\vec{\mathcal{G}}^\rho) \in \mathbb{S}^{9(M+P)}$ be the symmetric (9×9)-block-structured matrix such that

$$L(\vec{\mathcal{G}}^\rho)_{jk} =$$
$$\begin{cases} \sum_{e \in \delta^-(j)} \kappa_e N_e \mathbf{I} \otimes \mathbf{I} & j = k \\ -\kappa_{(j,k)} \sum_{i=1}^{N_{(j,k)}} \tilde{\mathbf{R}}_{\mathbf{B}_{(j,k),i}} \otimes \mathbf{R}_{\mathbf{A}_{(j,k),i}}^\top & (j, k) \in \vec{\mathcal{E}} \\ -\kappa_{(k,j)} \sum_{i=1}^{N_{(k,j)}} \tilde{\mathbf{R}}_{\mathbf{B}_{(k,j),i}}^\top \otimes \mathbf{R}_{\mathbf{A}_{(k,j),i}} & (k, j) \in \vec{\mathcal{E}} \\ \mathbf{0}_{9 \times 9} & \text{otherwise.} \end{cases} \tag{36}$$

Let $\mathbf{v} \in \mathbb{R}^{3(M+P) \times 1}$ be the (3×1)-block-structured vector with

$$\mathbf{v}_l = \begin{cases} \sum_{e \in \delta^-(l)} \frac{1}{\sigma_e^2} \sum_{i=1}^{N_e} \mathbf{R}_{\mathbf{A}_{e,i}}^\top \mathbf{t}_{\mathbf{A}_{e,i}} & l \leq M \\ -\sum_{e \in \delta^+(l)} \frac{1}{\sigma_e^2} \sum_{i=1}^{N_e} \mathbf{t}_{\mathbf{A}_{e,i}} & \text{otherwise,} \end{cases} \tag{37}$$

and

$$v = \sum_{e \in \vec{\mathcal{E}}} \frac{1}{\sigma_e^2} \sum_{i=1}^{N_e} \mathbf{t}_{\mathbf{A}_{e,i}}^\top \mathbf{t}_{\mathbf{A}_{e,i}}. \tag{38}$$

The matrix $\boldsymbol{\Sigma}$ is the symmetric (9×9)-block-structured matrix determined by

$$\boldsymbol{\Sigma}_{jk} = \begin{cases} \sum_{e \in \delta^+(j)} \frac{1}{\sigma_e^2} \sum_{i=1}^{N_e} (\tilde{\mathbf{t}}_{\mathbf{B}_{e,i}} \tilde{\mathbf{t}}_{\mathbf{B}_{e,i}}^\top) \otimes \mathbf{I} & j = k \\ \mathbf{0}_{9 \times 9} & \text{otherwise.} \end{cases} \tag{39}$$

Let $\mathbf{U} \in \mathbb{R}^{3(M+P) \times 9(M+P)}$ be the (3×9)-block-structured matrix

$$\mathbf{U}_{jk} = \begin{cases} \sum_{e \in \delta^+(j)} \frac{1}{\sigma_e^2} \sum_{i=1}^{N_e} \tilde{\mathbf{t}}_{\mathbf{B}_{e,i}}^\top \otimes \mathbf{I} & j = k \\ -\frac{1}{\sigma_e^2} \sum_{i=1}^{N_{(j,k)}} \tilde{\mathbf{t}}_{\mathbf{B}_{(j,k),i}}^\top \otimes \mathbf{R}_{\mathbf{A}_{(j,k),i}}^\top & (j, k) \in \vec{\mathcal{E}}, \\ \mathbf{0}_{3 \times 9} & \text{otherwise.} \end{cases} \tag{40}$$

Let $\mathbf{u} \in \mathbb{R}^{9(M+P) \times 1}$ be the (9×1)-block-structured vector

$$\mathbf{u}_l = \begin{cases} \mathbf{0} & l \leq M \\ -\sum_{e \in \delta^+(l)} \sum_{i=1}^{N_e} \tilde{\mathbf{t}}_{\mathbf{B}_{e,i}} \otimes \mathbf{t}_{\mathbf{A}_{e,i}} & \text{otherwise.} \end{cases} \tag{41}$$

Using the matrices in Equations (35) to (41), we can define the monocular RWHEC objective function matrix as follows:

$$\mathbf{Q} = \frac{1}{2} \begin{bmatrix} L(\vec{\mathcal{G}}^\tau) & \mathbf{v} & \mathbf{U} \\ \mathbf{v}^\top & v & \mathbf{u}^\top \\ \mathbf{U}^\top & \mathbf{u} & \boldsymbol{\Sigma} + L(\vec{\mathcal{G}}^\rho) \end{bmatrix}. \tag{42}$$

With this new notation, the standard optimization problem for multiple **X**s and **Y**s can be stated as a QCQP whose objective function is a quadratic form:

Problem 5. *Homogeneous QCQP Formulation of Monocular Generalized RWHEC*

$$\min_{\mathbf{x}} \mathbf{x}^\top \mathbf{Q} \mathbf{x}$$
$$\text{s.t. } \mathbf{R}_{\mathbf{X}_j} \in \mathrm{SO}(3) \; \forall \, j \in \mathcal{V}_{\mathbf{X}} \tag{43}$$
$$\mathbf{R}_{\mathbf{Y}_k} \in \mathrm{SO}(3) \; \forall \, k \in \mathcal{V}_{\mathbf{Y}}.$$

The state vector used in Problem 5 is

$$\mathbf{x} \triangleq \begin{bmatrix} \mathbf{t}^\top & \alpha & \mathbf{r}^\top \end{bmatrix}^\top, \tag{44}$$

where

$$\mathbf{t} \triangleq \begin{bmatrix} \mathbf{t}_x^\top & \mathbf{t}_y^\top \end{bmatrix}^\top$$
$$\mathbf{t}_x \triangleq \begin{bmatrix} \mathbf{t}_{\mathbf{X}_1}^\top & \cdots & \mathbf{t}_{\mathbf{X}_M}^\top \end{bmatrix}^\top \tag{45}$$
$$\mathbf{t}_y \triangleq \begin{bmatrix} \mathbf{t}_{\mathbf{Y}_1}^\top & \cdots & \mathbf{t}_{\mathbf{Y}_P}^\top \end{bmatrix}^\top,$$

and

$$\mathbf{r} \triangleq \begin{bmatrix} \mathbf{r}_x^\top & \mathbf{r}_y^\top \end{bmatrix}^\top$$
$$\mathbf{r}_x \triangleq \begin{bmatrix} \mathrm{vec}\left(\mathbf{R}_{\mathbf{X}_1}\right)^\top & \cdots & \mathrm{vec}\left(\mathbf{R}_{\mathbf{X}_M}\right)^\top \end{bmatrix}^\top \tag{46}$$
$$\mathbf{r}_y \triangleq \begin{bmatrix} \mathrm{vec}\left(\mathbf{R}_{\mathbf{Y}_1}\right)^\top & \cdots & \mathrm{vec}\left(\mathbf{R}_{\mathbf{Y}_P}\right)^\top \end{bmatrix}^\top.$$

4.6 Reducing the Dimension of the QCQP

If we know the optimal rotation matrices $\mathbf{R}^\star_{\mathbf{X}_j}, \mathbf{R}^\star_{\mathbf{Y}_k}$ for $j \in \mathcal{V}_{\mathbf{X}}$ and $k \in \mathcal{V}_{\mathbf{Y}}$, then the unconstrained optimal $\mathbf{X}$ translation vectors $\mathbf{t}^\star_{\mathbf{X}_j}$ for $j \in \mathcal{V}_{\mathbf{X}}$, $\mathbf{Y}$ translation vectors $\mathbf{t}^\star_{\mathbf{Y}_k}$ for $k \in \mathcal{V}_{\mathbf{Y}}$, and scale $\alpha^\star$ can be recovered by solving the following linear system:

$$\begin{bmatrix} \mathbf{t}^{\star \top}_\alpha & \alpha^\star \end{bmatrix}^\top = - \begin{bmatrix} L(\vec{\mathcal{G}}^\tau) & \mathbf{v} \\ \mathbf{v}^\top & v \end{bmatrix}^\dagger \begin{bmatrix} \mathbf{U} \\ \mathbf{u}^\top \end{bmatrix} \mathbf{r}^\star. \quad (47)$$

Using the generalized Schur complement (Gallier 2010), we *reduce* $\mathbf{Q}$ to

$$\mathbf{Q}' = \frac{1}{2} \left(\mathbf{\Sigma} + L(\vec{\mathcal{G}}^\rho) - \begin{bmatrix} \mathbf{U}^\top & \mathbf{u} \end{bmatrix} \begin{bmatrix} L(\vec{\mathcal{G}}^\tau) & \mathbf{v} \\ \mathbf{v}^\top & v \end{bmatrix}^\dagger \begin{bmatrix} \mathbf{U} \\ \mathbf{u}^\top \end{bmatrix} \right). \quad (48)$$

The result is a reduced form of Problem 5 that only depends on the rotation variables:

Problem 6. *Reduced QCQP Formulation of Generalized Monocular RWHEC*

$$\begin{aligned} \min_{\mathbf{r} \in \mathbb{R}^{9(M+P)}} \quad & \mathbf{r}^\top \mathbf{Q}' \mathbf{r}, \\ \text{s.t.} \quad & \mathbf{R}_{\mathbf{X}_j} \in \mathrm{SO}(3) \; \forall \, j \in \mathcal{V}_{\mathbf{X}}, \\ & \mathbf{R}_{\mathbf{Y}_k} \in \mathrm{SO}(3) \; \forall \, k \in \mathcal{V}_{\mathbf{Y}}. \end{aligned} \quad (49)$$

The computational effort required to solve Problem 6 with the solution method described in Section 5 depends not only on the total number of matrix variables $M + P$, but also on the structure of $\vec{\mathcal{G}}$. For example, consider the two directed graphs shown in Figure 3, either of which is capable of representing eight cameras observing an arm with a target of unknown scale: Figure 3a corresponds to the interpretation in Equation (5), whereas Figure 3b corresponds to Equation (10). The fill patterns after taking the Schur complement of $\mathbf{Q}$ for these two graphs are displayed in Figure 4. In contrast to the multiple $\mathbf{Y}$s case in Figure 3a, the multiple $\mathbf{X}$s formulation in Figure 3b has a sparse block arrowhead pattern, which decreases the time required to solve the semidefinite relaxation of the QCQP. Many RWHEC problems are more complex than the cases in Figure 3 (i.e., involve multiple $\mathbf{X}$s and $\mathbf{Y}$s that are densely interconnected), but these two limiting cases are common in practice. In all experiments in Section 8 and Section 9, our method uses the conic operator splitting method COSMO (Garstka et al. 2021), which is able to exploit sparsity in $\mathbf{Q}'$. Therefore, whenever possible, practitioners using our method are encouraged to formulate their particular generalized RWHEC problem in a manner that promotes sparsity in the reduced objective matrix. We recognize that this leeway is not always available, as the sensors used to produce measurements $\mathbf{B}_k$ may only align well with the isotropic noise models of Section 4.2 for a single formulation. The interested reader can learn more about the impact of different interpretations of RWHEC from Ha (2023), where this topic is explored alongside a detailed discussion of anisotropic noise distributions beyond the scope of this work.
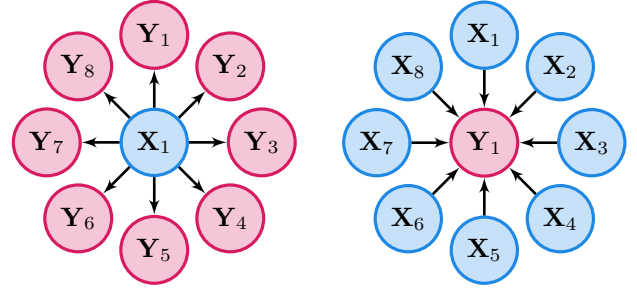
(a) A graph with multiple $\mathbf{Y}$s.   **(b)** A graph with multiple $\mathbf{X}$s.

Figure 3. Two directed graphs representing generalized RWHEC problems. Depending on the interpretation of variables discussed in Section 4.2, it is possible to use either graph for the same RWHEC scenario. Problems with the structure in Figure 3b lead to an objective function with greater sparsity (see Figure 4).
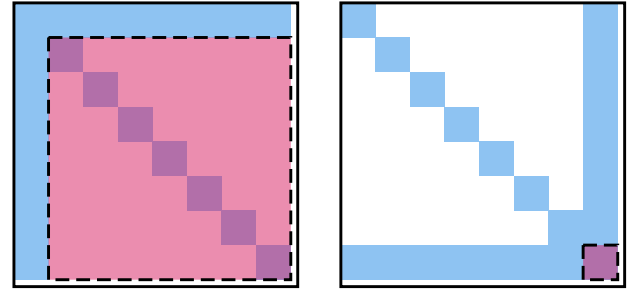
(a) Sparsity pattern of $\mathbf{Q}'$ for a formulation with multiple $\mathbf{Y}$s.   **(b)** Sparsity pattern of $\mathbf{Q}'$ for a formulation with multiple $\mathbf{X}$s.

Figure 4. Sparse matrix patterns for the multiple $\mathbf{X}$s and $\mathbf{Y}$s formulations in Figure 3. In these diagrams, the blue areas with no borders are the default fill patterns of the problems, while the red areas with dashed borders are the matrix areas that are filled by the Schur complement dimensionality reduction. The blue fill areas represent the general fill pattern of the sparse matrix and may have internal sparsity patterns. The final empty row and column represent the terms in the objective function involving the homogenization term, which only exists in the constraints in the unscaled case.

5 Certifiably Globally Optimal Extrinsic Calibration

In this section, we present a convex SDP relaxation of our calibration problem. Deriving the standard *Shor relaxation* of Problem 6 and its dual requires the homogenization of

the quadratic $SO(3)$ constraints in Equation (19) (Cifuentes et al. 2022, Section 1). Specifically, we introduce quadratically constrained variable $s^2 = 1$ and homogenize the linear and constant parts of Equation (19):

$$\mathbf{R}\mathbf{R}^\top = s^2\mathbf{I}, \tag{50a}$$

$$\mathbf{R}^\top\mathbf{R} = s^2\mathbf{I}, \tag{50b}$$

$$\mathbf{R}_{\sigma_i(1)} \times \mathbf{R}_{\sigma_i(2)} = s\mathbf{R}_{\sigma_i(3)} \ \forall i \in \{1, 2, 3\}. \tag{50c}$$

Treating the special case with known scale ($\alpha = 1$) in Problem 4 also requires that we replace α in Problem 6 with s, which ensures that the objective function is homogeneous.

Applying the theory of duality for generalized inequalities to the homogenized constraints in Problem 6 yields the following expression for its *Lagrangian* (Boyd and Vandenberghe 2004, Section 5.9):

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = \lambda_s + \mathbf{x}^\top\mathbf{Z}(\boldsymbol{\lambda})\mathbf{x}, \tag{51}$$

where

$$\begin{aligned}
\mathbf{Z}(\boldsymbol{\lambda}) &\triangleq \mathbf{Q}' + \mathbf{P}_1(\boldsymbol{\lambda}) + \mathbf{P}_2(\boldsymbol{\lambda}), \\
\mathbf{P}_1(\boldsymbol{\Lambda}_1, \dots, \boldsymbol{\Lambda}_{2(M+P)}) &\triangleq \\
&\mathrm{Diag}(-(\boldsymbol{\Lambda}_1 \oplus \boldsymbol{\Lambda}_2), \\
&\qquad \vdots \\
&\quad -(\boldsymbol{\Lambda}_{2(M+P)-1} \oplus \boldsymbol{\Lambda}_{2(M+P)}), \\
&\quad \sum_{i=1}^{2(M+P)} \mathrm{tr}(\boldsymbol{\Lambda}_i)),
\end{aligned} \tag{52}$$

and

$$\begin{aligned}
\mathbf{P}_2 &\triangleq \begin{bmatrix} \mathbf{P}_D & \mathbf{p} \\ \mathbf{p}^\top & -\lambda_s \end{bmatrix}, \\
\mathbf{P}_D &\triangleq \mathrm{Diag}\left(\mathbf{P}_h(\boldsymbol{\lambda}_1), \dots, \mathbf{P}_h(\boldsymbol{\lambda}_{(M+P)})\right), \\
\mathbf{p} &\triangleq -\begin{bmatrix} \mathbf{P}_o\boldsymbol{\lambda}_1 \\ \vdots \\ \mathbf{P}_o\boldsymbol{\lambda}_{(M+P)} \end{bmatrix}, \\
\boldsymbol{\Lambda}_1, &\dots, \boldsymbol{\Lambda}_{2(M+P)} \in \mathbb{S}^3, \\
\boldsymbol{\lambda}_1, &\dots, \boldsymbol{\lambda}_{(M+P)} \in \mathbb{R}^9.
\end{aligned} \tag{53}$$

We can subdivide $\boldsymbol{\lambda}$ into components

$$\boldsymbol{\lambda}_l \triangleq \begin{bmatrix} \boldsymbol{\lambda}_{l,ijk}^T & \boldsymbol{\lambda}_{l,jki}^T & \boldsymbol{\lambda}_{l,kij}^T \end{bmatrix}^\top \ \forall l = 1, \dots, M+P, \tag{54}$$

where $\boldsymbol{\lambda}_{l,ijk}, \boldsymbol{\lambda}_{l,jki}, \boldsymbol{\lambda}_{l,kij} \in \mathbb{R}^3$. Finally, the matrices $\mathbf{P}_o$ and $\mathbf{P}_h$ are

$$\begin{aligned}
\mathbf{P}_o &\triangleq \begin{bmatrix} \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \\ \mathbf{I} & \mathbf{0} & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{9\times9}, \\
\mathbf{P}_h(\boldsymbol{\lambda}_l) &\triangleq \begin{bmatrix} \mathbf{0}_{3\times3} & -\boldsymbol{\lambda}_{l,ijk}^\wedge & \boldsymbol{\lambda}_{l,kij}^\wedge \\ \boldsymbol{\lambda}_{l,ijk}^\wedge & \mathbf{0}_{3\times3} & -\boldsymbol{\lambda}_{l,jki}^\wedge \\ -\boldsymbol{\lambda}_{l,kij}^\wedge & \boldsymbol{\lambda}_{l,jki}^\wedge & \mathbf{0}_{3\times3} \end{bmatrix} \in \mathbb{S}^9.
\end{aligned} \tag{55}$$

Algorithm 1 The Certifiable RWHEC Algorithm

Input: Problem graph $\vec{\mathcal{G}} = (\mathcal{V}, \vec{\mathcal{E}})$ and data $\mathcal{D}_e \ \forall e \in \vec{\mathcal{E}}$.
Output: Estimates $\hat{\alpha}, \hat{\mathbf{X}}_j, \hat{\mathbf{Y}}_k \ \forall j, k \in \mathcal{V}$ and the suboptimality bound $\hat{\rho}$.

1: **function** CERTIFIABLE-RWHEC($\vec{\mathcal{G}}, \{\mathcal{D}_e\}_{e \in \vec{\mathcal{E}}}$)
2:      Form $\mathbf{Q}'$ from the inputs
3:      Solve Problem 7 for $\boldsymbol{\lambda}^\star$
4:      Set $d^\star \leftarrow \lambda_s^\star$
5:      Compute $\mathbf{r}' \in \ker(\mathbf{Z}(\boldsymbol{\lambda}^\star))$
6:      Set $\hat{\mathbf{r}} \leftarrow$ NORMALIZE($\mathbf{r}'$)
7:      Recover $\hat{\mathbf{t}}$ and $\hat{\alpha}$ via Equation (47)
8:      Reshape $\hat{\mathbf{r}}$ and $\hat{\mathbf{t}}$ into $\hat{\mathbf{X}}_j, \hat{\mathbf{Y}}_k \ \forall j, k \in \mathcal{V}$
9:      Set $p \leftarrow (\hat{\mathbf{r}})^\top\mathbf{Q}'\hat{\mathbf{r}}$
10:      Set $\hat{\rho} \leftarrow (p - d^\star)/d^\star$
11:      **return** $\left\{\hat{\alpha}, \hat{\mathbf{X}}_j, \hat{\mathbf{Y}}_k, \hat{\rho}\right\}$
12: **end function**

The minimum of the Lagrangian function is only defined if $\mathbf{Z}(\boldsymbol{\lambda}) \succeq 0$. As a result, the Lagrangian dual problem is the following SDP:

Problem 7. *Dual of Monocular RWHEC*

$$\begin{aligned}
\max_{\boldsymbol{\lambda}} \ & \lambda_s, \\
\text{s.t. } & \mathbf{Z}(\boldsymbol{\lambda}) \succeq 0,
\end{aligned} \tag{56}$$

where $\boldsymbol{\lambda} \in \mathbb{R}^{22(M+P)+1}$ and λ_s is the component of $\boldsymbol{\lambda}$ corresponding to the homogenizing constraint $s^2 = 1$ in the primal problem.

Our complete certifiable RWHEC algorithm is presented in Algorithm 1. Given a solution $\boldsymbol{\lambda}^\star$ to dual Problem 7, the KKT conditions for SDPs tell us that solution $\mathbf{r}^\star$ to the corresponding primal problem is a null vector of $\mathbf{Z}(\boldsymbol{\lambda}^\star)$ (Cifuentes et al. 2022). Given a null vector of $\mathbf{Z}(\boldsymbol{\lambda}^\star)$, we can scale it by the unique factor that sets the determinant of $\mathbf{R}^\star$ to 1 in order to recover $\mathbf{r}^\star$. Then, we can determine the optimal $\alpha^\star$ and scaled translation $\mathbf{t}^\star$ with Equation (47) and correct $\mathbf{t}^\star$ by a factor of $1/\alpha^\star$.

In addition to providing the computational benefits of solving a convex problem, our SDP relaxation allows us to numerically *certify* the global optimality of a primal solution $\mathbf{x}^\star$: by weak Lagrangian duality (Boyd and Vandenberghe 2004), $\lambda_s^\star$ is a lower bound on the optimal objective value of the primal Problem 6. Therefore, when this *duality gap* is small with respect to machine precision, we obtain a *post hoc* certificate of the optimality of our primal solution. We provide an example of this certification procedure on real experimental data in Section 9.3.

6 Uniqueness of Solutions

In this section, we derive conditions on measurement data which ensure that the standard generalized RWHEC

formulation of Problem 4 has a unique solution in the absence of noise (i.e., sufficient conditions for a problem instance to be an *identifiable* model).[9] In addition to precisely characterizing which robot and sensor motions lead to a well-posed calibration problem with identifiable extrinsic parameters, the results in this section are used in Section 7 to prove that SDP relaxations of our QCQP are tight, even when noisy measurements are used. A similar analysis is conducted for hand-eye calibration of a single monocular camera in Andreff et al. (2001) and the standard RWHEC problem in Shah (2013), but to our knowledge the multi-sensor and multi-target case has not been addressed until now.

6.1 The Rotation-Only Case

We begin by characterizing the uniqueness of solutions to the RWHEC problem with $\mathrm{SO}(3)$-valued data:

Problem 8. *Rotation-Only Generalized RWHEC*

$$
\min_{\mathbf{x}} \frac{1}{2} \sum_{(j,k)\in\vec{\mathcal{E}}} \sum_{i=1}^{N_{(j,k)}} \left\| \mathbf{R}_{\mathbf{A}_{(j,k),i}} \mathbf{R}_{\mathbf{X}_j} - \mathbf{R}_{\mathbf{Y}_k} \mathbf{R}_{\mathbf{B}_{(j,k),i}} \right\|_{\mathrm{F}}^2
$$
$$
\text{s.t. } \mathbf{R}_{\mathbf{X}_j} \in \mathrm{SO}(3) \ \forall j \in [M]
$$
$$
\mathbf{R}_{\mathbf{Y}_k} \in \mathrm{SO}(3) \ \forall k \in [P].
$$
(57)

Problem 8 with a single pair of variables $\mathbf{R}_{\mathbf{X}}$ and $\mathbf{R}_{\mathbf{Y}}$ is sometimes called *conjugate rotation averaging* in the computer vision literature (Hartley et al. 2013). It is essentially Problem 4 without translation variables, and with $\kappa_{e,i} = 1 \ \forall e \in \vec{\mathcal{E}}$ for notational convenience.

Theorem 1. *Consider an instance of Problem 8 induced by a weakly connected bipartite directed graph $\vec{\mathcal{G}} = (\mathcal{V}, \vec{\mathcal{E}})$ with* exact *rotation measurements $\underline{\mathbf{R}}_{\mathbf{A}_{e,i}}, \underline{\mathbf{R}}_{\mathbf{B}_{e,i}} \in \mathrm{SO}(3)$ for $i \in [N_e]$ associated with each edge $e \in \vec{\mathcal{E}}$. If a single rotation $\mathbf{R}_{\mathbf{X}_v}$ or $\mathbf{R}_{\mathbf{Y}_v}$ for some $v \in \mathcal{V}$ can be uniquely determined with the problem data, then this instance of Problem 8 has a unique solution.*

Proof. Exact measurements imply that each residual in Problem 8 is zero:

$$
\underline{\mathbf{R}}_{\mathbf{A}_{e,i}} \mathbf{R}_{\mathbf{X}_j} \underline{\mathbf{R}}_{\mathbf{B}_{e,i}}^\top = \mathbf{R}_{\mathbf{Y}_k} \ \forall e \in \vec{\mathcal{E}}, \ i \in [N_e]. \quad (58)
$$

Equation (58) gives us an explicit expression for $\mathbf{R}_{\mathbf{Y}_k}$ in terms of $\mathbf{R}_{\mathbf{X}_j}$ for any $(j,k) \in \vec{\mathcal{E}}$. Since $\underline{\mathbf{R}}_{\mathbf{A}_{e,i}}$ and $\underline{\mathbf{R}}_{\mathbf{B}_{e,i}}$ are rotation matrices and therefore invertible, we can also solve Equation (58) for $\mathbf{R}_{\mathbf{X}_j}$ in terms of $\mathbf{R}_{\mathbf{Y}_k}$:

$$
\underline{\mathbf{R}}_{\mathbf{A}_{e,i}}^\top \mathbf{R}_{\mathbf{Y}_k} \underline{\mathbf{R}}_{\mathbf{B}_{e,i}} = \mathbf{R}_{\mathbf{X}_j} \ \forall e \in \vec{\mathcal{E}}, \ i \in [N_e]. \quad (59)
$$

Now suppose (by our hypothesis) that there is some $v \in \mathcal{V}$ such that we can uniquely identify the associated rotation

$\mathbf{R}_{\mathbf{X}_v}$ or $\mathbf{R}_{\mathbf{Y}_k}$. Since $\vec{\mathcal{G}}$ is weakly connected, there is a semi-path connecting v and any other $w \in \mathcal{V}$. By repeatedly applying the identities in Equations (58) and (59), we can uniquely identify the rotations associated with each vertex along the semi-path joining v to w. Since every vertex in $\vec{\mathcal{G}}$ can be joined to v by such a path, this shows that the rotations associated with *all* vertices in $\vec{\mathcal{G}}$ are uniquely identifiable.

Theorem 1 tells us that in the idealized noise-free version of Problem 8, a single rotation with a unique solution implies that the entire problem has a unique solution. The following corollary is a direct consequence of our result and Theorem 2.3 in Shah (2013):

Corollary 2. *Consider an instance of Problem 8 with weakly connected graph $\vec{\mathcal{G}} = (\mathcal{V}, \vec{\mathcal{E}})$ and exact measurements. If there exists an edge $e = (j,k) \in \vec{\mathcal{E}}$ with $i_1, i_2, i_3 \in [N_e]$ such that $\mathbf{R}_{\mathbf{A}_{e,i_2}}^\top \mathbf{R}_{\mathbf{A}_{e,i_1}}$ and $\mathbf{R}_{\mathbf{A}_{e,i_3}}^\top \mathbf{R}_{\mathbf{A}_{e,i_1}}$ have distinct principal axes, then there is a unique solution. Furthermore, this solution is an element of a 1-dimensional vector space containing all* unconstrained *minimizers of the homogeneous objective function of Problem 8.*

Proof. Theorem 2.3 in Shah (2013) tells us that both $\mathbf{R}_{\mathbf{X}_j}$ and $\mathbf{R}_{\mathbf{Y}_k}$ have a unique solution, satisfying the hypothesis of Theorem 1 and guaranteeing the existence of a unique solution. Furthermore, if we use Equation (26) to vectorize the inner sum over i in the homogeneous objective of Problem 8 for the edge $e = (j,k)$ satisfying our hypotheses we get

$$
\sum_{i=1}^{N_{(j,k)}} \left\| \mathbf{R}_{\mathbf{A}_{(j,k),i}} \mathbf{R}_{\mathbf{X}_j} - \mathbf{R}_{\mathbf{Y}_k} \mathbf{R}_{\mathbf{B}_{(j,k),i}} \right\|_{\mathrm{F}}^2 = \tag{60}
$$
$$
\left\| \mathbf{Q}_{(j,k)} \mathbf{r}_{(j,k)} \right\|_2^2,
$$

where

$$
\mathbf{r}_{(j,k)} \triangleq \begin{bmatrix} \mathrm{vec}\left(\mathbf{R}_{\mathbf{X}_j}\right) \\ \mathrm{vec}\left(\mathbf{R}_{\mathbf{Y}_k}\right) \end{bmatrix}. \tag{61}
$$

Furthermore, $\mathbf{Q}_{(j,k)}$ has a 1-dimensional kernel in our case of exact measurements (Andreff et al. 2001, Lemma 1). Since the objective of Problem 8 is nonnegative and attains a value of zero for our unique solution, any *unconstrained* minimizer must also attain a value of zero and therefore satisfy Equation (58). Each element $\eta \mathbf{r}_{(j,k)} \in \ker(\mathbf{Q}_{(j,k)})$ is identified with $\eta \mathbf{R}_{\mathbf{X}_v}$ and $\eta \mathbf{R}_{\mathbf{Y}_v}$ for all $v \in \mathcal{V}$ by the same argument from weak connectedness used in the proof of Theorem 1. Therefore, we have established the existence of a 1-dimensional vector space of unconstrained minimizers of Problem 8 induced by the informative edge e and parameterized by $\eta \in \mathbb{R}$.

Corollary 2 provides us with a geometrically interpretable sufficient condition for uniqueness: if there is a

single sensor-target pair in the problem graph that gathered measurements from orientations related by rotations about two distinct axes, then the entire multi-frame rotational calibration problem has a unique solution. However, Theorem 1 in its full generality suggests that there exist cases where the data from more than one edge is required to verify that there is a unique solution.[10] Finally, the fact that the unique solution guaranteed by Corollary 2 is also a unique (up to scale) unconstrained minimizer of the objective function of Problem 8 will be used in our theorem on global optimality in Section 7.

6.2   The Full $\mathrm{SE}(3)$ *Case*

We can use Theorem 1 to prove a similar result for the inhomogeneous formulation of standard RWHEC in Problem 4 with exact measurements:

Theorem 3. *Consider an instance of Problem 4 induced by a weakly connected bipartite directed graph $\vec{\mathcal{G}} = (\mathcal{V}, \vec{\mathcal{E}})$ with exact pose measurements $\mathbf{A}_{e,i}, \mathbf{B}_{e,i} \in \mathrm{SE}(3)$ for $i \in [N_e]$ associated with each edge $e \in \vec{\mathcal{E}}$. If a single transformation $\mathbf{X}_v$ or $\mathbf{Y}_v$ for some $v \in \mathcal{V}$ can be uniquely determined with the problem data, then this instance of Problem 4 has a unique solution.*

Proof. Since the rotation residual terms in the objective of Problem 4 are independent of translations, Theorem 1 gives us a unique solution for the rotation component of each $\mathbf{X}_j$ and $\mathbf{Y}_k$. Exact measurements imply that the translational residuals are also zero:

$$\mathbf{t}_{\mathbf{X}_j} = \mathbf{R}_{\mathbf{A}_{e,i}}^\top (\mathbf{R}_{\mathbf{Y}_k} \underline{\mathbf{t}}_{\mathbf{B}_{e,i}} + \mathbf{t}_{\mathbf{Y}_k} - \underline{\mathbf{t}}_{\mathbf{A}_{e,i}}) \, \forall e \in \vec{\mathcal{E}}. \quad (62)$$

Equation (62) relates each pair of translation components via an affine equation. Since each rotation matrix $\mathbf{R}_{\mathbf{A}_{e,i}}^\top$ is full rank, any translation $\mathbf{t}_{\mathbf{X}_j}$ or $\mathbf{t}_{\mathbf{Y}_k}$ adjacent to a uniquely determined translation is itself uniquely determined via Equation (62). Therefore, the fact that $\vec{\mathcal{G}}$ is weakly connected allows us to uniquely specify each translation through a procedure analogous to the one used to prove Theorem 1. $\blacksquare$

Once again, the results of Shah (2013) provide us with a geometrically interpretable sufficient condition that is analogous to the rotation-only case in Corollary 2:

Corollary 4. *Consider an instance of Problem 4 with graph $\vec{\mathcal{G}} = (\mathcal{V}, \vec{\mathcal{E}})$ and exact measurements. If there exists an edge $e = (j, k) \in \vec{\mathcal{E}}$ with $i_1, i_2, i_3 \in [N_e]$ such that $\mathbf{R}_{\mathbf{A}_{e,i_2}}^\top \mathbf{R}_{\mathbf{A}_{e,i_1}}$ and $\mathbf{R}_{\mathbf{A}_{e,i_3}}^\top \mathbf{R}_{\mathbf{A}_{e,i_1}}$ have distinct principal axes, then the problem has a unique solution. Furthermore, if*

$$\mathbf{t}_e \triangleq \begin{bmatrix} \mathbf{R}_{\mathbf{Y}_k} \underline{\mathbf{t}}_{\mathbf{B}_{e,1}} \\ \vdots \\ \mathbf{R}_{\mathbf{Y}_k} \underline{\mathbf{t}}_{\mathbf{B}_{e,N_e}} \end{bmatrix} \notin \mathcal{R}(\mathbf{M}_e) \quad (63)$$

where

$$\mathbf{M}_e \triangleq \begin{bmatrix} -\mathbf{R}_{\mathbf{A}_{e,1}} & \mathbf{I} \\ \vdots & \\ -\mathbf{R}_{\mathbf{A}_{e,N_e}} & \mathbf{I} \end{bmatrix} \in \mathbb{R}^{3N_e \times 6}, \quad (64)$$

then this solution is also the unique unconstrained minimizer of the homogeneous objective function of Problem 4.

Proof. Corollary 2 gives us the unique rotation solution (up to scale) to Equation (58) for all edges of $\vec{\mathcal{G}}$. Theorem 3.2 in Shah (2013) tells us that the translation components of $\mathbf{X}_j$ and $\mathbf{Y}_k$ both have a unique solution, satisfying the hypotheses of Theorem 3.

To establish that this solution is also the unique unconstrained minimizer of the objective function, we can parameterize the 1-dimensional vector space of unconstrained minimizers of the rotational cost from Corollary 2 as $\eta \mathbf{R}_{\mathbf{X}_v}$ or $\eta \mathbf{R}_{\mathbf{Y}_v}$ for all $v \in \mathcal{V}$ and a single $\eta \in \mathbb{R}$. Since we have determined with Theorem 3 that there is a *constrained* minimizer corresponding to $\eta = 1$ that attains a value of zero, it suffices to show that this is the only *unconstrained* minimizer that sets the inhomogeneous translation residual associated with our informative edge e to zero:

$$\mathbf{R}_{\mathbf{A}_{e,i}} \mathbf{t}_{\mathbf{X}_j} + \underline{\mathbf{t}}_{\mathbf{A}_{e,i}} - \mathbf{t}_{\mathbf{Y}_k} - \eta \mathbf{R}_{\mathbf{Y}_k} \underline{\mathbf{t}}_{\mathbf{B}_{e,i}} = \mathbf{0}, \, i \in [N_e]. \quad (65)$$

This system of equations can be written as the matrix equation

$$\bar{\mathbf{M}}_e \begin{bmatrix} \mathbf{t}_{\mathbf{X}_j} \\ \mathbf{t}_{\mathbf{Y}_k} \\ \eta \end{bmatrix} = \begin{bmatrix} \underline{\mathbf{t}}_{\mathbf{A}_{e,1}} \\ \vdots \\ \underline{\mathbf{t}}_{\mathbf{A}_{e,N_e}} \end{bmatrix}, \quad (66)$$

where

$$\bar{\mathbf{M}}_e \triangleq \begin{bmatrix} \mathbf{M}_e & \mathbf{t}_e \end{bmatrix} \in \mathbb{R}^{3N_e \times 7}. \quad (67)$$

Since $\mathbf{M}_e$ is full rank (Shah 2013, Theorem 3.2) and $\mathbf{t}_e \notin \mathcal{R}(\mathbf{M}_e)$, the augmented $\bar{\mathbf{M}}_e$ is also full rank. Furthermore, since $\bar{\mathbf{M}}_e$ has more rows than columns ($N_e \geq 3$ by hypothesis), it is injective (Axler 2024, Theorems 3.15 and 3.21) and any solution to Equation (66) is therefore unique. The first part of this proof has furnished us with a solution for $\eta = 1$, and we have demonstrated that this is the unique unconstrained minimizer of the objective in Problem 4 as desired. $\blacksquare$

The requirement that measurements are made from at least three poses that differ by rotations about two distinct axes was first derived with unit quaternions by Zhuang et al. (1994). Theorem 3 reveals that satisfying this condition for a single $\mathbf{X}_j$-$\mathbf{Y}_k$ pair is sufficient to ensure generalized RWHEC has a unique solution in the idealized case without noise. Additionally, note that since $\mathbf{M}_e$ is full rank and $N_e \geq 3$ by hypothesis, $\mathcal{R}(\mathbf{M}_e)$ is a 6-dimensional proper

subspace of $\mathbb{R}^{3N_e}$. Consequently, Equation (63) holds for almost all values of $\mathbf{t}_e \in \mathbb{R}^{3N_e}$, ensuring that in general a unique solution is also an unconstrained minimizer, which is pertinent to the *a priori* optimality guarantees in Section 7. These results indicate that a practitioner calibrating a multi-sensor rig can ensure all parameters are identifiable by exciting their platform about two axes when taking measurements of a single target-sensor pair, so long as the directed graph of measurements $\vec{\mathcal{G}}$ is weakly connected.

7 Global Optimality Guarantees

A solution to the convex SDP relaxation described in Section 5 can provide a *post hoc* upper bound on global suboptimality via the duality gap. In this section, we demonstrate that the SDP relaxation also has *a priori* global optimality guarantees when measurement noise is below a problem-dependent threshold. This is achieved by applying the following theorem (Cifuentes et al. 2022, Theorem 3.9):

Theorem 5. *Consider the family of parametric QCQPs of the form*

$$\min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{x}^\top \mathbf{F}(\boldsymbol{\theta})\mathbf{x} + \mathbf{f}(\boldsymbol{\theta})^\top \mathbf{x} + c(\boldsymbol{\theta}) \\ \text{s.t. } g(\mathbf{x}) = \mathbf{0}, \tag{68}$$

where $\mathbf{F} : \Theta \to \mathbb{S}^n$, $\mathbf{f} : \Theta \to \mathbb{R}^n$, *and* $c : \Theta \to \mathbb{R}$ *are continuous functions of parameter* $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^d$, *and the multivariate constraint function* $g : \mathbb{R}^n \to \mathbb{R}^m$ *is quadratic. Let* $\bar{\boldsymbol{\theta}}$ *be such that the objective function of Equation (68) is strictly convex, and its unique unconstrained minimizer* $\bar{\mathbf{x}}$ *is also the minimizer of the constrained problem. If the Abadie constraint qualification (ACQ) holds at the solution* $\bar{\mathbf{x}}$, *then there is a neighbourhood of* $\bar{\boldsymbol{\theta}}$ *in which the primal and dual SDP relaxations of Equation (68) are tight.*

The ACQ is a weak regularity condition precisely stated in Definition 3.1 of Cifuentes et al. (2022), and it guarantees the existence of Lagrange multipliers at $\bar{\mathbf{x}}$. Informally, if the feasible set of Equation (68) describes a smooth manifold, then the ACQ holds at $\mathbf{x}$ if the rank of the Jacobian $\nabla g(\mathbf{x})$ of the system of constraint equations at $\mathbf{x}$ describing the feasible set is equal to the codimension of the smooth manifold.

By examining Equation (42), Problem 4 can be written in the form of the objective of Equation (68):

$$\mathbf{F}(\boldsymbol{\theta}) = \frac{1}{2} \begin{bmatrix} L(\vec{\mathcal{G}}^\tau) & \mathbf{U}, \\ \mathbf{U}^\top & \boldsymbol{\Sigma} + L(\vec{\mathcal{G}}^\rho) \end{bmatrix}, \\ \mathbf{f}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{v} \\ \mathbf{u} \end{bmatrix}, \\ c(\boldsymbol{\theta}) = \frac{v}{2}, \tag{69}$$

where $\boldsymbol{\theta} \triangleq \mathrm{vec}\left(\{\mathbf{A}_{e,i}, \mathbf{B}_{e,i}\}_{e \in \vec{\mathcal{E}}, i \in [N_e]} \right)$ is a vectorization of the measurement data from which the elements on the right hand side of Equation (69) are formed. However, in order to apply Theorem 5 to generalized RWHEC, we need to demonstrate that the ACQ holds for our redundant implementation of SO(3) constraints in Equation (19):

Theorem 6. *Let* $g : \mathbb{R}^9 \to \mathbb{R}^{21}$ *be a vectorization of the quadratic constraints encoding* SO(3) *in Equation (19). Then the ACQ holds for all* $\mathbf{x}$ *satisfying* $g(\mathbf{x}) = \mathbf{0}$, *and by extension for a "stacked" vectorization of constraints encoding* $SO(3)^n$.

Proof. We will apply Theorem 8 from Appendix B to show that the ACQ holds for g on SO(3). We begin by defining the decomposition $g = (g_1, g_2)$ as in Equation (79). Let $g_1 : \mathbb{R}^9 \to \mathbb{R}^6$ encode a vectorization of the six scalar equations in the symmetric row orthogonality constraint in Equation (19a), and let $g_2 : \mathbb{R}^9 \to \mathbb{R}^{15}$ encode Equations (19b) and (19c). The linear independence constraint qualification (LICQ) holds for g_1 at all vectorized inputs in O(3) $\subset \mathbb{R}^{3 \times 3}$ (Boumal 2023, Sec. 7.3), satisfying the hypothesis of part (a) in Theorem 8. It remains to demonstrate that the hypothesis of part (b) holds. The first block of g_2 (corresponding to Equation (19b)) is an algebraic rearrangement of g_1 and therefore locally constant on SO(3). The second block (Equation (19c)) is equivalent to $\det(\mathbf{R}) = 1$ when combined with g_1 (Tron et al. 2015). Since O(3) \ SO(3) is disconnected from SO(3), this block is also locally constant on SO(3). We have satisfied the hypotheses of Theorem 8, demonstrating that ACQ holds at $\mathbf{x} = \mathrm{vec}(\mathbf{R})$ for all $\mathbf{R} \in SO(3)$. Since the constraint set for each SO(3)-valued variable is independent of the others, the ACQ holds for any point in $SO(3)^N$ as well, as the rank of a block diagonal matrix is the sum of the rank of its blocks.

Theorem 6 establishes that redundant polynomial parameterizations of SO(3) constraints do not seriously interfere with the regularity properties of optimization problems. This result is important because the addition of redundant rotation constraints in a primal QCQP can improve the tightness of its SDP relaxations (Tron et al. 2015; Dümbgen et al. 2023). Additionally, it is worth noting that Theorem 6 is not limited to the various forms of generalized RWHEC studied in this work: it applies to *any* QCQP whose only constrained variables are in SO(3) or SE(3). We are now prepared to prove our main result, which is a straightforward application of Theorem 5:

Proposition 7. *Let* $\bar{\boldsymbol{\theta}} \triangleq \mathrm{vec}\left(\{\mathbf{A}_{e,i}, \mathbf{B}_{e,i}\}_{e \in \vec{\mathcal{E}}, i \in [N_e]} \right)$ *consist of exact measurements parameterizing an instance of Problem 4 as per Equation (68) and Equation (69). If there is a unique solution* $\bar{\mathbf{x}}$ *that is also an unconstrained*

minimizer of the objective function, then there is a neighbourhood of $\bar{\boldsymbol{\theta}}$ in which the SDP relaxation of Problem 4 is tight.

Proof. We must show that $\bar{\boldsymbol{\theta}}$ defines a QCQP satisfying the hypotheses of Theorem 5. We proved that the ACQ holds in Theorem 6, and the objective's dependence on $\boldsymbol{\theta}$ is polynomial and therefore continuous. It remains to demonstrate that the objective is strictly convex, which is equivalent to demonstrating that $\mathbf{F}(\bar{\boldsymbol{\theta}})$ is positive definite (Boyd and Vandenberghe 2004, Example 3.2). Since $\bar{\mathbf{x}}$ is the unique unconstrained minimizer by hypothesis and our objective is convex, $\mathbf{F}(\bar{\boldsymbol{\theta}}) \succ \mathbf{0}$ (Boyd and Vandenberghe 2004, Example 4.5) and our claim follows.

Proposition 7 tells us that a noisy instance of the RWHEC problem has a tight SDP relaxation so long as its measurements are sufficiently close to noise-free measurements describing a problem with a unique solution. The uniqueness result in Theorem 3 therefore gives us simple geometric criteria by which to determine whether an instance of RWHEC is well-posed and globally solvable via SDP relaxation under moderate noise. Note that the linear independence hypothesis $\mathbf{t}_e \notin \mathcal{R}(\mathbf{M}_e)$ in Corollary 4 which ensures that the unique minimizer is also the unconstrained minimizer plays a key role in our proof of Proposition 7. One useful interpretation of our results in this section is that under extremely mild assumptions and moderate noise, identifiable instances of the standard generalized RWHEC problem are precisely those instances which can be efficiently computed via convex relaxation. To our knowledge, this *a priori* global optimality guarantee is the first of its kind for a multi-sensor calibration problem.

8 Simulation Experiments

In this section, we use synthetic data from two simulated robotic systems to compare the accuracy and robustness of our algorithm with a variety of other RWHEC methods. The first system consists of a robotic manipulator with a hand-mounted camera observing a visual fiducial target. Using the simulated manipulator hand poses and camera-target measurements, this system forms a RWHEC problem with one $\mathbf{X}$ and one $\mathbf{Y}$ variable. The second system generates data for a RWHEC problem with four $\mathbf{X}$s and one $\mathbf{Y}$ by simulating a robotic manipulator with a hand-mounted target observed by four stationary cameras. For each system, we generate data to study both standard and monocular RWHEC.

To generate measurements for the robotic manipulator with a hand-mounted camera, we simulate a camera trajectory relative to the target and fix groundtruth values for $\mathbf{X}$ and $\mathbf{Y}$. At each point in time indexed by i, we use the

camera pose relative to the target to determine the camera-target transformation $\mathbf{B}_i$. By combining the ground truth values of $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{B}_i$, we calculate the ground truth values for each $\mathbf{A}_i$. Similarly, for the cameras observing a robotic manipulator with a hand-mounted target, the pose of camera j and the target pose at time i are used to calculate $\mathbf{B}_{ij} \triangleq \mathbf{B}_{e_j,i}$, where e_j is the unique edge in the problem graph associated with camera j. Once again, the ground truth values for $\mathbf{X}_j$, $\mathbf{Y}$, and $\mathbf{B}_{ij}$, are used to compute the ground truth measurements for each $\mathbf{A}_{ij} \triangleq \mathbf{A}_{e_j,i}$. We add samples from a Gaussian distribution $\mathcal{N}\left(\mathbf{0}, \sigma^2 \mathbf{I}\right)$ to the ground truth translation measurements $\bar{\mathbf{t}}_{\mathbf{B}_i}$ and $\bar{\mathbf{t}}_{\mathbf{B}_{ij}}$. Furthermore, we right-perturb the noiseless rotation measurements $\bar{\mathbf{R}}_{\mathbf{B}_i}$ and $\bar{\mathbf{R}}_{\mathbf{B}_{ij}}$ with samples from an isotropic Langevin distribution $\mathbf{R}_n \sim \text{Lang}\left(\mathbf{I}, \kappa\right)$. In each individual experimental trial, the noise parameters are fixed for all measurements: we consider translation noise with standard deviations of $\sigma = 1$ cm or $\sigma = 5$ cm, and rotation noise concentrations of $\kappa = 125$ or $\kappa = 12$. In the monocular RWHEC studies, we scale the translation component of the noisy $\mathbf{B}_i$ measurements by $\alpha = 0.5$. Finally, we generate 100 random runs for each experiment.

Using the data from our simulation studies, we compare the estimated parameter accuracy of our standard and monocular RWHEC methods to those of five standard RWHEC methods and one monocular RWHEC method. The five benchmark RWHEC solvers are the two-stage closed-form methods in Shah (2013) and Wang et al. (2022), the certifiable method in Horn et al. (2023), the probabilistic method in Dornaika and Horaud (1998), and a local on-manifold method (LOM) of our own design (see Appendix A for details of our implementation). For the monocular RWHEC scenarios, we compare the accuracy of our algorithm to LOM. When a scenario has more than one $\mathbf{X}$ to solve for, we compare our methods with Wang et al. (2022), Horn et al. (2023), and LOM.

For the solvers that require initialization, we provide random initial values, the solution from Wang et al. (2022), or parameters close to the ground truth. For each experiment, our method and the method in Horn et al. (2023) are randomly initialized. We attempted randomly initializing LOM, but this scheme resulted in convergence to local minima. Consequently, we also initialize the solver in Dornaika and Horaud (1998) and LOM using the solution from the method in Wang et al. (2022). However, the method in Wang et al. (2022) cannot solve monocular RWHEC problems. In the monocular RWHEC studies, we initialize LOM with calibration parameters that are within 10 cm and 10° of the ground truth values. Additionally, we set the initial scale estimate to 1.

Our globally optimal solver was implemented in Julia with the JuMP modelling language (Dunning et al. 2017),

which provides convenient access to several general-purpose optimization methods. For all experiments in this section and Section 9, our algorithm used the conic operator splitting method (COSMO) solver with default parameters except for residual norm convergence tolerances of $\epsilon_{abs} = \epsilon_{rel} = 5 \times 10^{-11}$, initial step size $\rho = 10^{-4}$, and a maximum of one million iterations to ensure convergence (Garstka et al. 2021). LOM was implemented in Ceres (Agarwal et al. 2023), and we used default settings with the exception of a maximum of 1000 iterations, a function tolerance of 10^{-15}, and a gradient tolerance of 10^{-19} for all experiments.

Finally, it is important to note that the problem formulation solved by LOM uses a standard multiplicative exponentiated Gaussian noise model (Barfoot 2024), and therefore does not minimize the same MLE objective function derived in Section 4.2. Therefore, we cannot expect the global minima of LOM's objective to exactly match those of the problem formulation solved by our convex method. We partially mitigate the experimental effects of this discrepancy by roughly matching the intensity of the isotropic variance σ_R^2 described in Appendix A to the Langevin concentration parameter κ using the asymptotic Gaussian approximation of the von Mises distribution described in Appendix A of Rosen et al. (2019).

8.1 Robot Arm Poses on a Sphere

In this pair of standard and monocular RWHEC experiments, we consider the problem of extrinsically calibrating a hand-mounted camera and a target, relative to a robotic manipulator. A diagram of the type of system we simulate is shown in Figure 1. In particular, we estimate $\mathbf{X} = \mathbf{T}_{hc}$ and $\mathbf{Y} = \mathbf{T}_{bt}$, where $\underrightarrow{\mathcal{F}}_b$, $\underrightarrow{\mathcal{F}}_t$, $\underrightarrow{\mathcal{F}}_h$, $\underrightarrow{\mathcal{F}}_c$, are the manipulator base, target, hand, and camera reference frames, respectively. For each problem instance, the camera collects measurements from 100 distinct poses along a trajectory on the surface of a sphere while its optical axis $\hat{\mathbf{z}}$ is pointed towards the target, $\hat{\mathbf{y}}$ points towards the south pole of the sphere, and $\hat{\mathbf{x}}$ completes the orthonormal frame. Although they contain rotations about two distinct axes and therefore satisfy the identifiability conditions of Corollary 4, trajectories generated in this manner for were empirically observed to not have a unique solution for the monocular RWHEC problem. We found that identifiable monocular RWHEC problem instances could be created by collecting measurements from the surface of two spheres: one with a unit radius, and another with a radius of 0.3 m.

The mean and standard deviation of the estimated translation ($t_{x,err}$, $t_{y,err}$) and rotation ($r_{x,err}$, $r_{y,err}$) errors are shown in Table 2. Generally, the least accurate methods are the two-stage closed-form solvers from Shah (2013) and Wang et al. (2022), while our method and LOM are the most

accurate. In spite of the differences in their formulation, our method and LOM return extremely similar solutions, which outperform the other methods by up to 50 mm and 3 degrees. Additionally, these two methods are more sensitive to translation measurement noise than rotation measurement noise. Finally, all algorithms have roughly the same accuracy when $\kappa = 125$ and $\sigma = 5$ cm. In this case, the rotation measurement noise is low, so the two-stage closed-form solvers are more likely to return accurate rotation estimates.

Table 3 contains the mean and standard deviation of the estimated translation ($t_{x,err}$, $t_{y,err}$), rotation ($r_{x,err}$, $r_{y,err}$), and scale error (α_{err}) for the monocular RWHEC study. As is the case for the standard RWHEC experiment, our method and LOM have very similar accuracy. However, our method does not require an initial estimate of $\mathbf{X}$, $\mathbf{Y}$, or α. Even though monocular RWHEC is ostensibly more challenging than RWHEC, both monocular RWHEC solvers return more accurate estimates than their corresponding RWHEC methods. We posit that this discrepancy is due to constraining the camera trajectory to a single sphere in the RWHEC study. Collecting data on the surface of two spheres may have resulted in more "identifiable" data, improving the accuracy of the estimated parameters (see, e.g., Grebe (2022) for more insight into the effect of sensor trajectory on calibration accuracy).

8.2 Many Cameras Observing a Hand-Mounted Target

In this study, we compare the estimated parameter accuracy of standard and monocular RWHEC algorithms for systems with multiple $\mathbf{X}$s and one $\mathbf{Y}$. The simulated system consists of a robotic manipulator with a hand-mounted target viewed by four fixed cameras. Figure 5 shows a diagram of the base-camera and hand-target transformations for one camera. Each variable $\mathbf{X}_i = \mathbf{T}_{bc_i}$ relates the manipulator base frame $\underrightarrow{\mathcal{F}}_b$ and the ith camera frame $\underrightarrow{\mathcal{F}}_{c_i}$. The unknown transformation from the robot hand to the target is represented by the variable $\mathbf{Y} = \mathbf{T}_{ht}$, where $\underrightarrow{\mathcal{F}}_h$ and $\underrightarrow{\mathcal{F}}_t$ are the manipulator hand and target reference frames, respectively. Each simulated trajectory consists of 108 manipulator poses, ensures that the target is always visible to every fixed camera, and has a unique solution by satisfying the requirements of Corollary 4.

The results of standard and monocular RWHEC studies are shown in Tables 4 and 5, which contain the mean and standard deviation of the estimated translation ($t_{x,err}$, $t_{y,err}$) and rotation ($r_{x,err}$, $r_{y,err}$) error. The errors associated with $\mathbf{X}$ (i.e., $t_{x,err}$ and $r_{x,err}$) are the mean and standard deviation of all four estimated $\mathbf{X}$ transformations. Table 5 also displays the mean and standard deviation of the estimated scale error (α_{err}). As with the previous experiments, our method and LOM produced the most

Table 2. Calibration results for the experiments in Section 8.1 with known scale. The values in each row are estimated by a different algorithm. The mean error magnitude and standard deviation are given in each cell. Dornaika and LOM are initialized with the method in Wang et al. (2022). Initializing LOM with random poses results in convergence to suboptimal local minima with errors orders of magnitude greater than the global solutions and have not been included.

Noise Level	Method (Init.)	$t_{x,err}$ [mm]	$r_{x,err}$ [deg]	$t_{y,err}$ [mm]	$r_{y,err}$ [deg]
$\kappa = 125,$ $\sigma = 1$cm	Shah	20.6 ± 10.5	1.37 ± 0.55	9.9 ± 4.8	1.34 ± 0.57
	Wang	20.6 ± 10.5	1.37 ± 0.55	9.9 ± 4.8	1.34 ± 0.57
	Dornaika (Wang)	19.9 ± 10.2	1.25 ± 0.53	8.8 ± 4.3	1.20 ± 0.54
	Horn	18.5 ± 9.4	1.16 ± 0.51	8.4 ± 4.2	1.10 ± 0.51
	LOM (Wang)	11.1 ± 5.82	0.78 ± 0.37	$\mathbf{3.69 \pm 2.12}$	0.63 ± 0.38
	Ours	$\mathbf{10.9 \pm 5.6}$	$\mathbf{0.77 \pm 0.36}$	3.71 ± 2.13	$\mathbf{0.62 \pm 0.36}$
$\kappa = 125,$ $\sigma = 5$cm	Shah	31.2 ± 15.0	1.61 ± 0.64	21.2 ± 10.0	1.57 ± 0.66
	Wang	31.2 ± 15.0	1.61 ± 0.64	21.2 ± 10.0	1.57 ± 0.66
	Dornaika (Wang)	30.6 ± 14.5	1.49 ± 0.62	20.4 ± 9.7	1.44 ± 0.63
	Horn	29.1 ± 13.6	$\mathbf{1.42 \pm 0.60}$	19.9 ± 9.2	$\mathbf{1.36 \pm 0.59}$
	LOM (Wang)	28.5 ± 13.4	1.45 ± 0.62	$\mathbf{18.5 \pm 8.8}$	1.39 ± 0.62
	Ours	$\mathbf{28.4 \pm 13.0}$	1.42 ± 0.63	$\mathbf{18.5 \pm 8.7}$	1.36 ± 0.61
$\kappa = 12,$ $\sigma = 1$cm	Shah	65.5 ± 38.3	4.34 ± 1.95	31.8 ± 17.2	4.41 ± 1.98
	Wang	65.5 ± 38.3	4.34 ± 1.95	31.8 ± 17.2	4.41 ± 1.98
	Dornaika (Wang)	63.4 ± 37.3	3.92 ± 1.87	27.4 ± 16.2	3.92 ± 1.90
	Horn	45.0 ± 24.8	3.30 ± 1.69	25.8 ± 14.1	3.19 ± 1.55
	LOM (Wang)	15.2 ± 10.0	1.84 ± 0.84	$\mathbf{3.4 \pm 1.8}$	0.88 ± 0.59
	Ours	$\mathbf{15.1 \pm 10.0}$	$\mathbf{1.81 \pm 0.83}$	$\mathbf{3.4 \pm 1.8}$	$\mathbf{0.87 \pm 0.59}$
$\kappa = 12,$ $\sigma = 5$cm	Shah	71.9 ± 44.7	4.74 ± 2.13	37.8 ± 19.3	4.63 ± 2.21
	Wang	71.9 ± 44.7	4.74 ± 2.13	37.8 ± 19.3	4.63 ± 2.21
	Dornaika (Wang)	69.9 ± 43.1	4.33 ± 2.08	33.9 ± 17.9	4.13 ± 2.14
	Horn	50.6 ± 28.4	3.69 ± 1.72	31.0 ± 16.2	3.40 ± 1.72
	LOM (Wang)	48.3 ± 27.3	3.18 ± 1.67	$\mathbf{18.7 \pm 9.04}$	2.72 ± 1.64
	Ours	$\mathbf{47.7 \pm 27.0}$	$\mathbf{3.12 \pm 1.67}$	18.8 ± 9.02	$\mathbf{2.68 \pm 1.65}$

Table 3. Calibration results for the experiments in Section 8.1 with unknown scale. The values in each row are estimated by a different algorithm. The mean error magnitude and standard deviation are given in each cell. LOM is initialized with parameters $10°$ and 10 cm from the ground truth values and a scale of 1. As in Table 2, LOM performed poorly without an accurate initialization and those results have been omitted.

Noise Level	Method (Init.)	$t_{x,err}$ [mm]	$r_{x,err}$ [deg]	$t_{y,err}$ [mm]	$r_{y,err}$ [deg]	α_{err} [%]
$\kappa = 125,$ $\sigma = 1$cm	LOM (close)	$\mathbf{4.61 \pm 2.39}$	$\mathbf{0.51 \pm 0.21}$	$\mathbf{3.82 \pm 1.93}$	$\mathbf{0.22 \pm 0.12}$	$1.78\text{e-}3 \pm 1.54\text{e-}3$
	Ours	4.71 ± 2.43	$\mathbf{0.51 \pm 0.22}$	3.82 ± 1.91	0.250 ± 0.113	$\mathbf{1.76\text{e-}3 \pm 1.54\text{e-}3}$
$\kappa = 125,$ $\sigma = 5$cm	LOM (close)	$\mathbf{23.0 \pm 11.2}$	$\mathbf{0.94 \pm 0.44}$	19.1 ± 11.4	$\mathbf{0.85 \pm 0.40}$	$8.45\text{e-}3 \pm 7.03\text{e-}3$
	Ours	23.8 ± 11.3	1.11 ± 0.53	$\mathbf{19.0 \pm 11.4}$	1.06 ± 0.52	$\mathbf{8.41\text{e-}3 \pm 7.02\text{e-}3}$
$\kappa = 12,$ $\sigma = 1$cm	LOM (close)	4.52 ± 1.97	1.57 ± 0.59	3.78 ± 1.89	$\mathbf{0.20 \pm 0.10}$	$1.88\text{e-}3 \pm 1.37\text{e-}3$
	Ours	$\mathbf{4.50 \pm 1.99}$	$\mathbf{1.54 \pm 0.57}$	$\mathbf{3.77 \pm 1.88}$	0.206 ± 0.098	$\mathbf{1.87\text{e-}3 \pm 1.37\text{e-}3}$
$\kappa = 12,$ $\sigma = 5$cm	LOM (close)	$\mathbf{23.2 \pm 14.1}$	$\mathbf{1.90 \pm 0.81}$	20.0 ± 11.0	$\mathbf{1.07 \pm 0.54}$	$9.24\text{e-}3 \pm 7.69\text{e-}3$
	Ours	24.4 ± 14.3	1.99 ± 0.81	$\mathbf{19.6 \pm 11.0}$	1.25 ± 0.58	$\mathbf{9.16\text{e-}3 \pm 7.69\text{e-}3}$

accurate parameter estimates. The method of Wang et al. (Wang et al. 2022) returns approximate closed-form solutions using a linearized system and the singular value decomposition, yet our experiments demonstrate that it is more accurate than the globally optimal method of Horn et al. (Horn et al. 2023). Recalling that there is a double cover from the unit quaternions to SO(3), we hypothesize that the use of dual quaternions by the method in Horn et al. (2023) leads to its deteriorated performance in the multi-**X** case, as each measurement (i.e., **A** or **B**) must be assigned one of
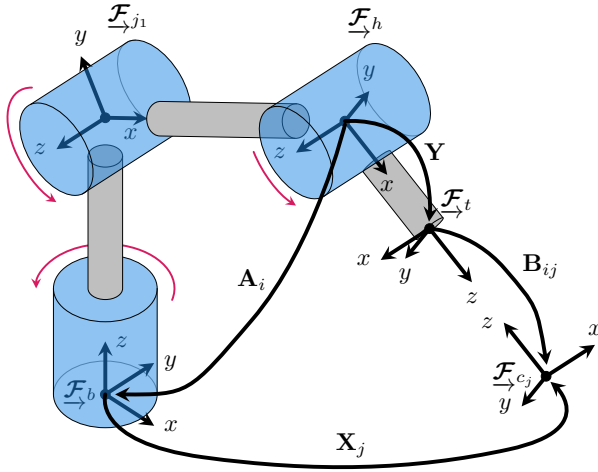
Figure 5. A diagram of a robot arm with a hand-mounted target that is observed by a stationary camera. In this diagram, the base, joint 1, hand, camera j, and target reference frames are labelled $\underrightarrow{\mathcal{F}}_b$, $\underrightarrow{\mathcal{F}}_{j_1}$, $\underrightarrow{\mathcal{F}}_h$, $\underrightarrow{\mathcal{F}}_{c_j}$, and $\underrightarrow{\mathcal{F}}_t$, respectively. The red arrows indicate the axis of joint rotation. At time i, we use the forward kinematics of the manipulator to estimate the transformation from the manipulator base to the hand, $\mathbf{A}_i$. Further, we can measure the transformation from the target to camera j, $\mathbf{B}_{ij}$.

two equivalent unit DQ representations of SE(3). Finally, as in our previous monocular RWHEC study, our algorithm and LOM initialized to within 10 cm and 10° of the true solution achieved similar accuracy.

9 Real-World Experiment

In this section, we discuss our real-world data collection system, data preprocessing procedure, and calibration results.

9.1 Data Collection and Data Preprocessing

In our real world experiment, we use infrared motion capture markers placed on a mobile sensor system to produce measurements $\mathbf{A}_{(j,k),i}$ and fiducial markers in the environment with unknown poses $\mathbf{X}_j$ to form the RWHEC problem shown in Figure 6. Figures 7 and 8 show our mobile system and two images of our indoor experimental environment, respectively. Our mobile system is equipped with eight Point Grey Blackfly S USB cameras, OptiTrack markers, and a VectorNav VN-100 inertial measurement unit (IMU). We place a sufficient number of OptiTrack markers on the system to enable estimation of the relative transformation between the OptiTrack reference frame and a frame fixed to the mobile system. To validate our estimated camera calibration parameters, we approximate ground truth extrinsics $\mathbf{Y}_k$ with an estimate computed using the Kalibr toolbox (Rehder et al. 2016), which requires measurements from the IMU in addition to images of a checkerboard target taken by each camera. Kalibr uses a

local method and therefore requires an initialization which we manually computed, therefore it should only be treated as a rough proxy for the true solution.

A total of sixteen AprilTags (Olson 2011) are mounted in our experimental environment for use as fiducial markers. To evaluate the accuracy of estimated AprilTag poses $\mathbf{X}_j$, we place OptiTrack markers on a single AprilTag to establish ground truth measurements.

Figure 9 shows the trajectory of the mobile system in our experiment. In the first half of the data collection run, the system experiences purely planar motion, and each target is observed at least once. After the planar motion, the system rotates about all three axes and translates perpendicular to the plane of motion from the first half of the trajectory, ensuring sufficient excitation for the RWHEC problem to have a unique solution. Figure 10 is a grid where row k corresponds to camera k, and column j corresponds to AprilTag j. If the square in column j and row k is blue, then the data collected for that camera-target pair enables an identifiable RWHEC subproblem (i.e., an instance of RWHEC with a unique solution). A red square indicates that the data collected for that camera-target pair did not contain sufficient excitation, while a white square indicates that target j was not observed by camera j. Figure 10 indicates that each camera observed at least one target, and that the overall generalized RWHEC problem is described by a weakly connected bipartite directed graph.

Our data preprocessing pipeline for this experiment consists of four steps designed to align measurements in time and remove outliers.[11] First, we rectify the images using the camera intrinsic parameters estimated with Kalibr (see Rehder et al. (2016)). Second, we measure the camera-to-AprilTag transformations $\mathbf{B}_{(j,k),i}$ using an AprilTag detector. Third, we synchronize the camera and OptiTrack measurements by interpolating the OptiTrack system measurements to the camera measurement timestamps with the Lie-algebraic method in Chapter 8 of Barfoot (2024). Finally, we use a RANSAC (Fischler and Bolles 1981) scheme to reject gross outliers. For each ground truth estimate of $\mathbf{Y}_k$ computed with Kalibr, the RWHEC geometric constraint at each measurement time labelled by $i \in [N_{(j,k)}]$ becomes

$$\mathbf{X}_j = \mathbf{A}_{(j,k),i}^{-1}\mathbf{Y}_k\mathbf{B}_{(j,k),i}. \tag{70}$$

Consequently, we can use RANSAC on the model in Equation (70) to determine the $\mathbf{A}_{(j,k),i}$-$\mathbf{B}_{(j,k),i}$ pairs that result in a consistent $\mathbf{X}_j$ transformation. For this RANSAC procedure, our minimum inlier set size is one third of the number of $\mathbf{A}_{(j,k),i}$-$\mathbf{B}_{(j,k),i}$ pairs for camera k and target j. Data from camera-target pairs with only one $\mathbf{A}_{(j,k),i}$-$\mathbf{B}_{(j,k),i}$ measurement are rejected because they cannot be validated using this scheme. An $\mathbf{A}_{(j,k),i}$-$\mathbf{B}_{(j,k),i}$ pair is considered an inlier if it is within 0.6 m and 60° of the

Table 4. Calibration results for the multi-sensor experiments in Section 8.2 with known scale. The values in each row are estimated by a different algorithm. The mean error magnitude and standard deviation are given in each cell. LOM is initialized with the results from Wang. Note we have no a priori estimates of $t_{\mathbf{X}_i}$. Once again, we have not included results from experiments initializing LOM with random poses, as they converged to local minima.

Noise Level	Method (Init.)	$t_{x,err}$ [mm]	$r_{x,err}$ [deg]	$t_{y,err}$ [mm]	$r_{y,err}$ [deg]
$\kappa = 125$, $\sigma = 1$cm	Wang	2.32 ± 1.56	0.469 ± 0.200	2.71 ± 1.37	0.273 ± 0.120
	Horn	12.11 ± 6.49	0.455 ± 0.191	5.43 ± 2.04	0.165 ± 0.067
	LOM (Wang)	0.994 ± 0.416	$\mathbf{0.454 \pm 0.190}$	0.572 ± 0.296	0.022 ± 0.010
	Ours	$\mathbf{0.992 \pm 0.416}$	$\mathbf{0.454 \pm 0.189}$	$\mathbf{0.570 \pm 0.294}$	$\mathbf{0.021 \pm 0.010}$
$\kappa = 125$, $\sigma = 5$cm	Wang	4.95 ± 2.20	0.470 ± 0.196	3.52 ± 1.57	0.260 ± 0.115
	Horn	12.4 ± 6.4	4.55 ± 1.91	5.42 ± 2.33	1.76 ± 0.69
	LOM (Wang)	4.57 ± 1.96	0.460 ± 0.191	2.74 ± 1.34	0.120 ± 0.056
	Ours	$\mathbf{4.56 \pm 1.96}$	$\mathbf{0.459 \pm 0.191}$	$\mathbf{2.69 \pm 1.32}$	$\mathbf{0.111 \pm 0.051}$
$\kappa = 12$, $\sigma = 1$cm	Wang	6.05 ± 4.84	1.54 ± 0.63	7.63 ± 4.04	0.812 ± 0.364
	Horn	69 ± 121	24.6 ± 94.4	48.5 ± 82.5	13.0 ± 76.8
	LOM (Wang)	$\mathbf{0.933 \pm 0.370}$	1.52 ± 0.64	$\mathbf{0.537 \pm 0.262}$	$\mathbf{0.021 \pm 0.010}$
	Ours	$\mathbf{0.933 \pm 0.370}$	$\mathbf{1.50 \pm 0.62}$	$\mathbf{0.537 \pm 0.262}$	$\mathbf{0.021 \pm 0.010}$
$\kappa = 12$, $\sigma = 5$cm	Wang	7.85 ± 4.34	1.52 ± 0.70	9.58 ± 4.46	0.868 ± 0.340
	Horn	77 ± 146	26 ± 106	57 ± 110	14.8 ± 83.9
	LOM (Wang)	$\mathbf{4.57 \pm 1.97}$	1.51 ± 0.682	2.93 ± 1.34	0.104 ± 0.047
	Ours	$\mathbf{4.57 \pm 1.97}$	$\mathbf{1.48 \pm 0.67}$	$\mathbf{2.92 \pm 1.32}$	$\mathbf{0.103 \pm 0.046}$

Table 5. Calibration results for the multi-sensor experiments in Section 8.2 with unknown scale. The values in each row are estimated by a different algorithm. The mean error magnitude and standard deviation are given in each cell. LOM is initialized with parameters $10°$ and 10 cm from the ground truth values and a scale of 1. We have once again omitted experiments initializing LOM with random poses.

Noise Level	Method (Init.)	$t_{x,err}$ [mm]	$r_{x,err}$ [deg]	$t_{y,err}$ [mm]	$r_{y,err}$ [deg]	α_{err} [%]
$\kappa = 125$, $\sigma = 1$cm	LOM (close)	$\mathbf{1.05 \pm 0.45}$	0.469 ± 0.185	0.631 ± 0.249	0.030 ± 0.013	$\mathbf{2.71e\text{-}4 \pm 1.98e\text{-}4}$
	Ours	$\mathbf{1.05 \pm 0.45}$	$\mathbf{0.468 \pm 0.185}$	$\mathbf{0.597 \pm 0.256}$	$\mathbf{0.026 \pm 0.012}$	$\mathbf{2.71e\text{-}4 \pm 1.98e\text{-}4}$
$\kappa = 125$, $\sigma = 5$cm	LOM (close)	5.25 ± 2.20	0.457 ± 0.195	3.15 ± 1.26	0.179 ± 0.079	$\mathbf{1.37e\text{-}3 \pm 1.03e\text{-}3}$
	Ours	$\mathbf{5.21 \pm 2.18}$	$\mathbf{0.454 \pm 0.194}$	$\mathbf{3.03 \pm 1.24}$	$\mathbf{0.159 \pm 0.071}$	$\mathbf{1.37e\text{-}3 \pm 1.03e\text{-}3}$
$\kappa = 12$, $\sigma = 1$cm	LOM (close)	$\mathbf{1.02 \pm 0.44}$	1.50 ± 0.62	$\mathbf{0.530 \pm 0.284}$	0.021 ± 0.010	$2.58e\text{-}4 \pm 2.02e\text{-}4$
	Ours	$\mathbf{1.02 \pm 0.44}$	$\mathbf{1.46 \pm 0.61}$	$\mathbf{0.530 \pm 0.284}$	0.021 ± 0.010	$\mathbf{2.57e\text{-}4 \pm 2.02e\text{-}4}$
$\kappa = 12$, $\sigma = 5$cm	LOM (close)	5.20 ± 2.27	1.48 ± 0.64	3.13 ± 1.56	0.156 ± 0.075	$\mathbf{1.18e\text{-}3 \pm 8.63e\text{-}4}$
	Ours	$\mathbf{5.18 \pm 2.26}$	$\mathbf{1.44 \pm 0.626}$	$\mathbf{3.02 \pm 1.51}$	$\mathbf{0.144 \pm 0.069}$	$\mathbf{1.18e\text{-}3 \pm 8.62e\text{-}4}$

estimated pose $\mathbf{X}_j$. The inlier $\mathbf{A}_{(j,k),i}$-$\mathbf{B}_{(j,k),i}$ pairs are then used for our RWHEC problem. To extend this outlier rejection scheme to monocular RWHEC, we assume that the AprilTag size is known within 10% of the actual value, which we empirically found to be sufficient for reliable outlier rejection.

9.2 Experimental Results

Using the preprocessed data, we obtain the calibration results shown in Table 6. We assume that our hand-measured AprilTag sizes are approximately correct, so the local standard and monocular RWHEC solvers are initialized with the parameters estimated using the method

in Wang et al. (2022). Our estimated AprilTag translation is within 12 cm (or 8% of the ground-truth distance) and $6°$ of the ground truth transformation measured with OptiTrack. The estimated camera calibration parameters are, on average, within about 3 cm and $1°$ of the parameters estimated by Kalibr. We did not expect our algorithm to return the same values as Kalibr because collecting a dedicated calibration dataset for each camera should result in more accurate calibration parameters. As expected, the inexpensive approximation technique in Wang et al. (2022) returned the least accurate rotation estimates. As in our simulation studies, our method and LOM compute parameters with similar accuracy. Interestingly,
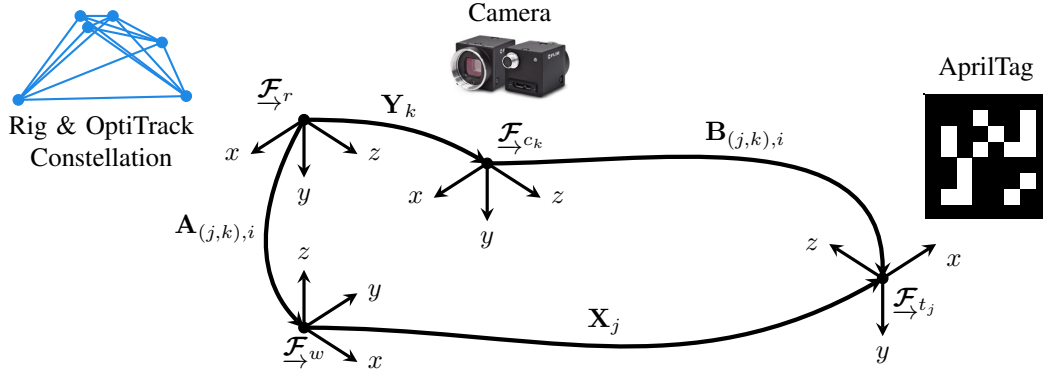
Figure 6. A diagram of the measurements for camera k and target j at time i. The reference frames for the cameras, targets, OptiTrack world, and rig reference frames are $\underset{\rightarrow}{\mathcal{F}}_{c_k}$, $\underset{\rightarrow}{\mathcal{F}}_{t_j}$, $\underset{\rightarrow}{\mathcal{F}}_w$, and $\underset{\rightarrow}{\mathcal{F}}_r$, respectively. The OptiTrack rig constellation enables estimation of the transformation $\mathbf{A}_{(j,k),i}$. The monocular camera observing the AprilTag enables estimation of the transformation $\mathbf{B}_{(j,k),i}$.

Table 6. Calibration results for our real world dataset with known scale. The values in each row are estimated by a different algorithm. The mean error magnitude and standard deviation for $\mathbf{t}_{c_0}^{c_i c_0}$ and $\mathbf{R}_{c_0 c_i}$ are given in each cell. The error in the estimated transformation between AprilTag 20 and the OptiTrack reference frame is also provided. LOM is initialized with the solution from Wang et al. (2022).

Method (Init.)	$\mathbf{t}_{c_0}^{c_i c_0}$ Error [cm]	$\mathbf{R}_{c_0 c_i}$ Error [deg]	$\mathbf{t}_w^{t_{20} w}$ Error [cm]	$\mathbf{R}_{w t_{20}}$ Error [deg]
Wang	3.46 ± 1.69	1.38 ± 0.81	11.7	5.30
LOM (Wang)	3.34 ± 2.44	0.88 ± 0.49	11.4	**4.50**
Ours	3.20 ± 1.88	0.99 ± 0.52	11.5	4.73
LOM with Unknown Scale (Wang)	3.28 ± 2.42	$\mathbf{0.86 \pm 0.47}$	**7.92**	**4.50**
Ours with Unknown Scale	$\mathbf{2.87 \pm 1.94}$	0.99 ± 0.52	7.94	4.72

estimating the scale of the AprilTags improves accuracy. The estimated AprilTag scale is 2.5% smaller than the hand-measured value. From our simulation studies, the
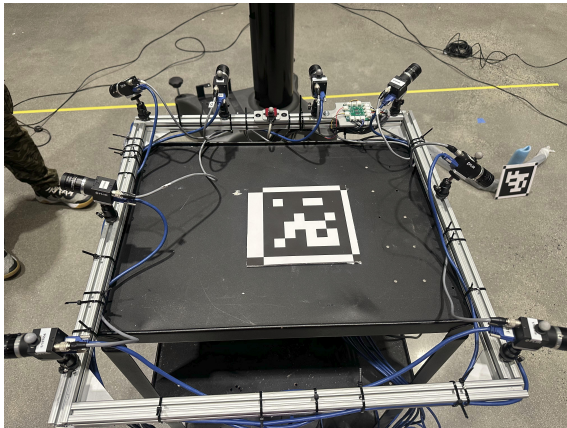

Figure 8. Images from the real-world experiment. These images are from camera 0 and show a subset of the AprilTags in the environment. The bottom left AprilTag in the image on the right has OptiTrack markers, so we can determine the ground truth pose of the AprilTag frame relative to the OptiTrack world frame.

Figure 7. Image of the real-world data collection rig. The data collection rig consists of eight hardware synchronized cameras facing a variety of different directions. Further, the data collection rig includes an IMU and opti track markers. The OptiTrack markers enable us to estimate the rig pose relative to the OptiTrack reference frame.
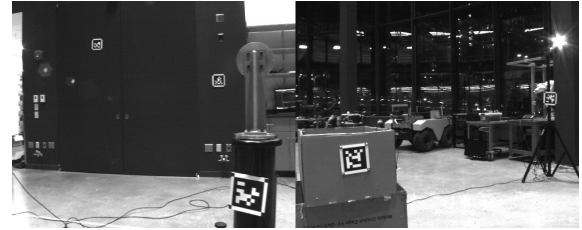
estimated scale error is often within 0.01% of ground truth value, so a scale error this large is unexpected. This scaling error suggests that our manual measurements of the AprilTags is incorrect by approximately 3 mm. This correction improved our estimated camera calibration parameters and AprilTag translations by approximately 0.5 cm and 4 cm, respectively. Therefore, our monocular method can significantly improve estimation accuracy in standard RWHEC problems with noisy fiducial marker size
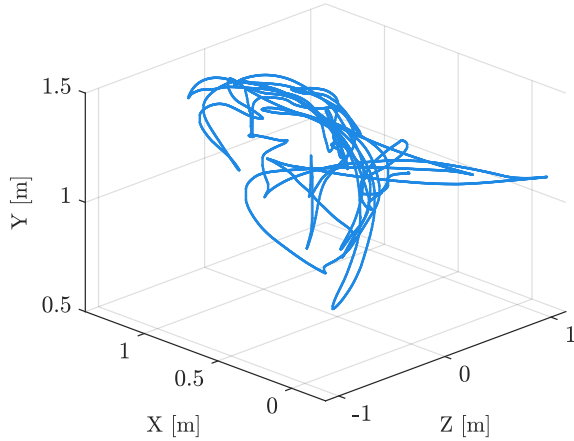
Figure 9. Trajectory of the rig in the real-world experiment. Initially, the platform rotates about the y-axis and moves in the xz-plane. Following the planar motion, the system follows an unconstrained trajectory, which allows for rotation about the x and z axes.
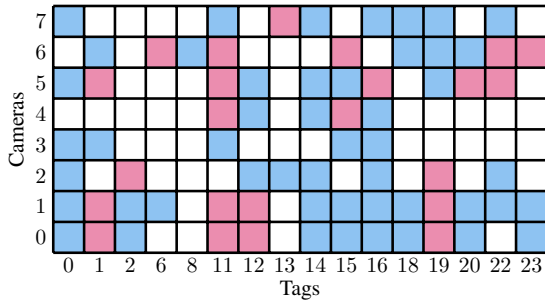

Figure 10. A grid describing the observability of each connection in the bipartite graph generated by our real-world experiment. Blue squares indicate that measurements between tag j and camera k are sufficient for an identifiable RWHEC problem. Red squares indicate that there is a connection between tag j and camera k, but the measurements are insufficient for the problem to be identifiable on its own. A white square indicates that there is no observation of tag j by camera k.

estimates, and future work can investigate a maximum a posteriori (MAP) estimation scheme that leverages prior estimates of parameters in a principled fashion.

9.3 Certification

As described in Section 5, we can numerically certify that our convex method achieves the global minimum. Weak duality tells us that

$$d \le d^\star \le p^\star \le p, \tag{71}$$

where d and p are the dual and primal objective function values attained by candidate solutions $\boldsymbol{\nu}$ and $\mathbf{x}$, and $d^\star$ and $p^\star$ are the optima of the dual and primal objective function values attained by the optimal solutions $\boldsymbol{\nu}^\star$ and $\mathbf{x}^\star$. The

relative suboptimality of our solution $\mathbf{x}$ is

$$\rho \triangleq \frac{p - p^\star}{p^\star}, \tag{72}$$

but we cannot directly compute ρ because do not have access to $p^\star$. Instead, we note that Equation (71) gives us the following upper bound on the relative suboptimality:

$$\hat{\rho} \triangleq \frac{p - d}{d} \ge \rho. \tag{73}$$

For the standard RWHEC problem with known scale, $\hat{\rho} = -6.41\text{e-}9$, and for the monocular case, $\hat{\rho} = 8.55\text{e-}9$. In both cases, the relative duality gap indicates that the primal solution returned by our method attains an objective value that is less than a millionth of a percent off of the lower bound provided by the dual solution. Finally, while weak duality means that $\hat{\rho} < 0$ is not possible in theory, our solutions were computed with floating point arithmetic which approximates $\mathbb{R}$ discretely and inevitably introduces roundoff errors. Therefore, the negative duality gap for the case with known scale is not large enough to cause concern.

9.4 Runtime

Since calibration is typically an offline procedure, we did not expend effort tuning solver parameters to reduce algorithm runtime. All experiments were run on a laptop with an Intel Core i7-8750H CPU and 16 GB of memory. The parameters used by COSMO and Ceres were selected to ensure accurate convergence, and the longest runtime observed was approximately five minutes for our global solver in the monocular real world experiment reported in Table 6. COSMO solved the synthetic problems with the sparsity pattern in Figure 3b substantially faster, taking at most seven seconds. Unsurprisingly, the Ceres implementation of LOM took at most one second.

10 Conclusion

We have presented an efficient and certifiably globally optimal solution to a generalized formulation of multi-sensor extrinsic calibration. Our formulation builds on previous robot-world and hand-eye calibration methods by incorporating monocular cameras and arbitrary multi-sensor and target configurations. Additionally, we have presented a novel local on-manifold solver which matches the accuracy of our global method when provided with a moderately accurate initialization. We have also characterized the subset of multi-sensor RWHEC problem instances which have a unique solution, and used this to prove that RWHEC admits a tight SDP relaxation when measurement noise is not too large. Our experiments verify that global optimality is an important consideration for RWHEC, and that our MLE problem formulation using

rotation matrices is superior to dual quaternion-based methods.

We see our contributions as critical steps towards truly "power-on-and-go" sensor calibration algorithms for multi-sensor robotic systems. Realizing this vision will require extending our RWHEC solver so that it does not depend on specialized calibration targets and is therefore able to handle outliers caused by errors in data association. One potential option for mitigating the effect of outliers on generalized RWHEC is to use a robust truncated least squares objective function, which has been applied to other estimation problems while preserving their QCQP structure (Yang et al. 2021). Additionally, our current MLE formulation assumes that translation and rotation noise is isotropic. Future work can extend our problem formulation with guidance from the study of anisotropic SLAM in Holmes et al. (2024). Finally, truly autonomous sensor calibration will require an active perception strategy which can design trajectories that are information-theoretically optimal (Grebe et al. 2021) or seek out measurements that render parameters identifiable (Yang et al. 2023).

Appendices

A A Local On-Manifold RWHEC Solver

In this appendix, we describe LOM, our local on-manifold solver for the monocular RWHEC problem. The standard or known-scale version is not explicitly included, but it can be easily derived by setting $\alpha = 1$ in the expressions to follow. Given a bipartite graph with nodes $\{\mathbf{X}_1, \ldots, \mathbf{X}_M, \mathbf{Y}_1, \ldots, \mathbf{Y}_P\}$, recall that $\mathcal{D}_{(j,k)}$ is the data describing measurements involving $\mathbf{X}_j$ and $\mathbf{Y}_k$:

$$\mathcal{D}_{(j,k)} = \{(\mathbf{A}_{(j,k),i}, \mathbf{B}_{(j,k),i}) \ \forall \ k = 1, \ldots, N_{(j,k)}\}. \quad (74)$$

The noisy measurement models for a given $(\mathbf{A}_{(j,k),i}, \mathbf{B}_{(j,k),i})$ pair are

$$
\begin{aligned}
\mathbf{R}_{\mathbf{B}_{(j,k),i}} =& \mathbf{R}_{\mathbf{Y}_k}^\top \mathbf{R}_{\mathbf{A}_{(j,k),i}} \mathbf{R}_{\mathbf{X}_j} \exp\left(\epsilon_{R_i}^\wedge\right), \\
\epsilon_{R_i} \sim& \mathcal{N}\left(0, \sigma_R{}^2 \mathbf{I}\right), \\
\mathbf{t}_{\mathbf{B}_{(j,k),i}} =& \mathbf{R}_{\mathbf{Y}_k}^\top \left(\mathbf{R}_{\mathbf{A}_{(j,k),i}} \mathbf{t}_{\mathbf{X}_j} + \alpha \mathbf{t}_{\mathbf{A}_{(j,k),i}} - \mathbf{t}_{\mathbf{Y}_k}\right) + \epsilon_{t_i}, \\
\epsilon_{t_i} \sim& \mathcal{N}\left(0, \sigma_t{}^2 \mathbf{I}\right).
\end{aligned}
\quad (75)
$$

This model leverages the right-perturbation noise framework of Barfoot (2024) to ensure that the model is an on-manifold analogue of the MLE model in Section 4.2. In the experiments of Sections 8 and 9, the σ_t used in Equation (75) is identical to the parameter with that name in Section 4.2, whereas we use the approximation in Appendix A of Rosen et al. (2019) to compute $\sigma_R = 1/\sqrt{\kappa}$, where κ is the concentration parameter of a Langevin distribution.

The associated error distributions are

$$
\begin{aligned}
\mathbf{e}_{R_{(j,k),i}} =& \log\left(\mathbf{R}_{\mathbf{X}_j}^\top \mathbf{R}_{\mathbf{A}_{(j,k),i}}^\top \mathbf{R}_{\mathbf{Y}_k} \mathbf{R}_{\mathbf{B}_{(j,k),i}}\right)^\vee \\
& \sim \mathcal{N}\left(0, \sigma_R{}^2 \mathbf{I}\right), \\
\mathbf{e}_{t_{(j,k),i}} =& \mathbf{R}_{\mathbf{A}_{(j,k),i}} \mathbf{t}_{\mathbf{X}_j} + \alpha \mathbf{t}_{\mathbf{A}_{(j,k),i}} - \mathbf{R}_{\mathbf{Y}_k} \mathbf{t}_{\mathbf{B}_{(j,k),i}} - \mathbf{t}_{\mathbf{Y}_k} \\
& \sim \mathcal{N}\left(0, \sigma_t{}^2 \mathbf{I}\right).
\end{aligned}
\quad (76)
$$

The associated error Jacobians are

$$
\begin{aligned}
\frac{\partial \mathbf{e}_{R_{(j,k),i}}}{\partial \psi_{\mathbf{X}_j}} =& -\mathbf{R}_{\mathbf{B}_{(j,k),i}}^\top \mathbf{R}_{\mathbf{Y}_k}^\top \mathbf{R}_{\mathbf{A}_{(j,k),i}}, \\
\frac{\partial \mathbf{e}_{R_{(j,k),i}}}{\partial \psi_{\mathbf{Y}_k}} =& \mathbf{R}_{\mathbf{B}_{(j,k),i}}^\top \mathbf{R}_{\mathbf{Y}_k}^\top, \\
\frac{\partial \mathbf{e}_{t_{(j,k),i}}}{\partial \mathbf{t}_{\mathbf{X}_j}} =& \mathbf{R}_{\mathbf{A}_{(j,k),i}}, \\
\frac{\partial \mathbf{e}_{t_{(j,k),i}}}{\partial \psi_{\mathbf{Y}_k}} =& \left(\mathbf{R}_{\mathbf{Y}_k} \mathbf{t}_{\mathbf{B}_{(j,k),i}}\right)^\wedge, \\
\frac{\partial \mathbf{e}_{t_{(j,k),i}}}{\partial \mathbf{t}_{\mathbf{Y}_k}} =& -\mathbf{I}, \\
\frac{\partial \mathbf{e}_{t_{(j,k),i}}}{\partial \alpha} =& \mathbf{t}_{\mathbf{A}_{(j,k),i}},
\end{aligned}
\quad (77)
$$

where $\psi_{\mathbf{X}_j}, \psi_{\mathbf{Y}_k} \in \mathbb{R}^3$ are linear perturbations that can be mapped to the Lie algebra $\mathfrak{so}(3)$ via the wedge operator $(\cdot)^\wedge$. LOM uses these Jacobians within the Ceres optimization library (Agarwal et al. 2023) to solve the following nonlinear least squares program:

Problem 9. *Local RWHEC Optimization Problem*

$$\min_{\substack{\mathbf{X}_1, \ldots, \mathbf{X}_M, \\ \mathbf{Y}_1, \ldots, \mathbf{Y}_P, \\ \alpha}} \sum_{(j,k),i} \frac{1}{\sigma_R^2} \mathbf{e}_{R_{(j,k),i}}^\top \mathbf{e}_{R_{(j,k),i}} + \frac{1}{\sigma_t^2} \mathbf{e}_{t_{(j,k),i}}^\top \mathbf{e}_{t_{(j,k),i}}. \quad (78)$$

B Constraint Qualification

Theorem 8. A constraint qualification for locally redundant constraints. *Let $c_1 \colon \mathbb{R}^n \to \mathbb{R}^{m_1}$ and $c_2 \colon \mathbb{R}^n \to \mathbb{R}^{m_2}$ be continuously-differentiable functions, and*

$$
\begin{aligned}
c \colon& \mathbb{R}^n \to \mathbb{R}^{m_1 + m_2} \\
c(\mathbf{x}) =& \left(c_1(\mathbf{x}), \ c_2(\mathbf{x})\right).
\end{aligned}
\quad (79)
$$

Fix $\bar{\mathbf{x}} \in \mathbb{R}^n$, and let $\bar{\mathbf{y}} \triangleq c(\bar{\mathbf{x}}) = (\bar{\mathbf{y}}_1, \bar{\mathbf{y}}_2) \in \mathbb{R}^{m_1 + m_2}$.

(a) *Suppose that the linear independence constraint qualification holds for c_1 at $\bar{\mathbf{x}}$ (i.e. that $\operatorname{rank} \nabla c_1(\bar{\mathbf{x}}) = m_1$), and define:*

$$\mathcal{X} \triangleq \{\mathbf{x} \in \mathbb{R}^n \mid c_1(\mathbf{x}) = \bar{\mathbf{y}}_1\}. \quad (80)$$

Then $\mathcal{X}$ is locally a smooth embedded submanifold of $\mathbb{R}^n$ about $\bar{\mathbf{x}}$; that is, there exists an open set $U \subseteq \mathbb{R}^n$ containing $\bar{\mathbf{x}}$ such that $U \cap \mathcal{X}$ is a smooth embedded submanifold of $\mathbb{R}^n$ of dimension $n - m_1$, and its tangent space at $\bar{\mathbf{x}}$ is given by:

$$T_{\bar{\mathbf{x}}}(\mathcal{X}) = \ker \nabla c_1(\bar{\mathbf{x}}). \tag{81}$$

(b) *Suppose additionally that the second constraint function c_2 is* locally constant *on $\mathcal{X}$ about $\bar{\mathbf{x}}$; that is, there exists an open set $V \subseteq U$ containing $\bar{\mathbf{x}}$ such that $c_2(\mathbf{x}) \equiv \bar{\mathbf{y}}_2$ for all $\mathbf{x} \in \mathcal{X} \cap V$. Then the feasible set*

$$\mathcal{M} \triangleq \{\mathbf{x} \in \mathbb{R}^n \mid c(\mathbf{x}) = \bar{\mathbf{y}}\} \tag{82}$$

determined by all *of the constraints $c(\mathbf{x})$ locally coincides with $\mathcal{X}$:*

$$\mathcal{M} \cap V = \mathcal{X} \cap V. \tag{83}$$

In particular, $\mathcal{M}$ is locally a smooth embedded submanifold of $\mathbb{R}^n$ about $\bar{\mathbf{x}}$ of dimension $n - m_1$, and its tangent space at $\bar{\mathbf{x}}$ is given by:

$$T_{\bar{\mathbf{x}}}(\mathcal{M}) = \ker \nabla c(\bar{\mathbf{x}}). \tag{84}$$

Proof. In the language of differential topology, the statement that the LICQ holds at $\bar{\mathbf{x}}$ (i.e. that $\operatorname{rank} \nabla c_1(\bar{\mathbf{x}}) = m_1$) is equivalent to the statement that the constraint function c_1 is a *submersion* at $\bar{\mathbf{x}}$. This property is *locally stable*; that is, if c_1 is a submersion at $\bar{\mathbf{x}}$, then in fact there exists an open set $U \subseteq \mathbb{R}^n$ containing $\bar{\mathbf{x}}$ such that $c_1|_U : U \to \mathbb{R}^{m_1}$ is a (global) submersion on U (Lee 2013, Prop. 4.1). It follows that the preimage of $\bar{\mathbf{y}}_1$ under $c_1|_U$:

$$(c_1|_U)^{-1}(\bar{\mathbf{y}}_1) \triangleq \{\mathbf{x} \in U \mid c_1(\mathbf{x}) = \bar{\mathbf{y}}_1\} = U \cap \mathcal{X} \tag{85}$$

is a smooth embedded sub manifold of $\mathbb{R}^n$ of dimension $n - m_1$ (Lee 2013, Cor. 5.13), and its tangent space at $\bar{\mathbf{x}}$ is given by (81) (Guillemin and Pollack 1974, Sec. 1.4):

$$T_{\bar{\mathbf{x}}}(\mathcal{X}) = \ker \nabla c_1(\bar{\mathbf{x}}). \tag{86}$$

To prove part (b), note that the hypotheses that $V \subseteq U$ and $c_2(\mathbf{x}) \equiv \bar{\mathbf{y}}_2$ for all $\mathbf{x} \in V$ immediately imply

$$c(\mathbf{x}) = \bar{\mathbf{y}} \iff c_1(\mathbf{x}) = \bar{\mathbf{y}}_1 \quad \forall \mathbf{x} \in V, \tag{87}$$

which is equivalent to (83) [cf. (80) and (82)]. Thus it remains only to show (84). To do so, we will prove that each of $\ker \nabla c(\bar{\mathbf{x}})$ and $T_{\bar{\mathbf{x}}}(\mathcal{M})$ is contained in the other, making use of the fact that:

$$T_{\bar{\mathbf{x}}}(\mathcal{X}) = T_{\bar{\mathbf{x}}}(\mathcal{M}) \tag{88}$$

since $\mathcal{X}$ and $\mathcal{M}$ locally coincide on the neighborhood V of $\bar{\mathbf{x}}$. Suppose first that $\mathbf{v} \in \ker \nabla c(\bar{\mathbf{x}})$. Then:

$$0 = \nabla c(\bar{\mathbf{x}})\mathbf{v} = \begin{pmatrix} \nabla c_1(\bar{\mathbf{x}})\mathbf{v} \\ \nabla c_2(\bar{\mathbf{x}})\mathbf{v} \end{pmatrix} \tag{89}$$

which implies in particular that $\mathbf{v} \in \ker \nabla c_1(\bar{\mathbf{x}}) = T_{\bar{\mathbf{x}}}(\mathcal{X}) = T_{\bar{\mathbf{x}}}(\mathcal{M})$ by (81) and (88); this shows that $\ker \nabla c(\bar{\mathbf{x}}) \subseteq T_{\bar{\mathbf{x}}}(\mathcal{M})$. Conversely, suppose that $\mathbf{v} \in T_{\bar{\mathbf{x}}}(\mathcal{M})$. Then $\nabla c_1(\mathbf{x})\mathbf{v} = 0$ [by (88) and (81)], and we must have $\nabla c_2(\bar{\mathbf{x}})\mathbf{v} = 0$ (since $\mathbf{v} \in T_{\bar{\mathbf{x}}}(\mathcal{M})$ and c_2 is locally constant on $\mathcal{M}$ about $\bar{\mathbf{x}}$ by hypothesis); together these prove that $\mathbf{v} \in \ker \nabla c(\bar{\mathbf{x}})$, and thus $T_{\bar{\mathbf{x}}}(\mathcal{M}) \subseteq \ker \nabla c(\bar{\mathbf{x}})$. Altogether this establishes (84), as claimed.

Notes

1. A calibration target is any specialized structure or fiducial marker whose pose, shape, or appearance is known to a robot or its operator.

2. Our code is available at `https://github.com/utiasSTARS/certifiable-rwhe-calibration`.

3. It is worth noting that only separate and joint methods form mutually exclusive categories.

4. The curious reader is directed to Ha (2023), which provides probabilistic interpretations of different *configurations* of reference frames appearing in prior RWHEC methods.

5. The astute reader may recognize that Equations (19a) and (19b) are redundant, in that either on its own is sufficient for ensuring $\mathbf{R} \in \mathrm{O}(3)$. However, the inclusion of both in the QCQP we formulate is essential as a form of "duality strengthening" (Briales et al. 2017) for the SDP relaxation derived in this section. While redundant in their original quadratic form, these constraints manifest as independent linear constraints in the corresponding relaxed and lifted SDP of Section 5, ultimately increasing the number of noisy problem instances that can be exactly solved by our approach (Wise et al. 2020; Dümbgen et al. 2023).

6. Our approach ensures that $\vec{\mathcal{G}}$ is *simple*, whereas an alternative formulation could use a directed *multigraph* whose edges correspond to individual observations between $\mathbf{X}_j$ and $\mathbf{Y}_k$.

7. The Julia implementation of our method used in the experiments of Sections 8 and 9 supports the use of inhomogeneous measurement precisions within edge data $\mathcal{D}_e$.

8. The applicability of these alternatives depends on whether measurements $\mathbf{A}_{e,i}$ or $\mathbf{B}_{e,i}$ correspond to camera-target measurements. In this work's MLE formulation, we assume the latter is true.

9. We leave the derivation of similar results for the monocular case of Problem 3 for future work.

10. We were indeed able to construct many such cases for a small ($M = P = 2$) synthetic problem graph. However, their complete characterization does not appear to be as mathematically elegant as the well-studied special case treated in Corollary 2.

11. None of the methods presented in this paper are robust, and the same cleaned dataset is used for all comparisons. Addressing challenges with data association is left for future work.

References

Agarwal S, Mierle K and Team TCS (2023) Ceres Solver. URL https://github.com/ceres-solver/ceres-solver.

Andreff N, Horaud R and Espiau B (2001) Robot Hand-Eye Calibration Using Structure-from-Motion. *The International Journal of Robotics Research* 20(3): 228–248.

Axler S (2024) *Linear Algebra Done Right*. Undergraduate Texts in Mathematics. Cham: Springer International Publishing. ISBN 978-3-031-41025-3 978-3-031-41026-0. DOI:10.1007/978-3-031-41026-0.

Barfoot TD (2024) *State Estimation for Robotics*. 2 edition. Cambridge University Press. ISBN 978-1-107-15939-6 978-1-316-67152-8. DOI:10.1017/9781009299909.

Boumal N (2023) *An Introduction to Optimization on Smooth Manifolds*. 1 edition. Cambridge University Press. ISBN 978-1-00-916616-4 978-1-00-916617-1 978-1-00-916615-7. DOI:10.1017/9781009166164.

Boyd SP and Vandenberghe L (2004) *Convex Optimization*. Cambridge, UK ; New York: Cambridge University Press. ISBN 978-0-521-83378-3.

Briales J, Gonzalez-Jimenez J et al. (2017) Convex Global 3D Registration with Lagrangian Duality. In: *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Carlone L, Rosen DM, Calafiore G, Leonard JJ and Dellaert F (2015) Lagrangian duality in 3D SLAM: Verification techniques and optimal solutions. In: *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Hamburg, Germany: IEEE, pp. 125–132. DOI:10.1109/iros.2015.7353364.

Cifuentes D, Agarwal S, Parrilo PA and Thomas RR (2022) On the local stability of semidefinite relaxations. *Mathematical Programming* 193(2): 629–663. DOI:10.1007/s10107-021-01696-1.

Čolaković-Bencerić M, Peršić J, Marković I and Petrović I (2025) Multiscale and Uncertainty-Aware Targetless Hand-Eye Calibration via the Gauss–Helmert Model. *IEEE Transactions on Robotics* 41: 2340–2357. DOI:10.1109/TRO.2025.3548538.

Dornaika F and Horaud R (1998) Simultaneous robot-world and hand-eye calibration. *IEEE transactions on Robotics and Automation* 14(4): 617–622.

Dümbgen F, Holmes C, Agro B and Barfoot TD (2023) Toward Globally Optimal State Estimation Using Automatically Tightened Semidefinite Relaxations.

Dumbgen F, Holmes C and Barfoot TD (2023) Safe and Smooth: Certified Continuous-Time Range-Only Localization. *IEEE Robotics and Automation Letters* 8(2): 1117–1124. DOI:10.1109/LRA.2022.3233232.

Dunning I, Huchette J and Lubin M (2017) JuMP: A Modeling Language for Mathematical Optimization. *SIAM Review* 59(2): 295–320. DOI:10.1137/15M1020575.

Evangelista D, Olivastri E, Allegro D, Menegatti E and Pretto A (2023) A graph-based optimization framework for hand-eye calibration for multi-camera setups. In: *2023 IEEE International Conference on Robotics and Automation (ICRA)*. pp. 11474–11480. DOI:10.1109/ICRA48891.2023.10160758.

Fan T, Wang H, Rubenstein M and Murphey T (2020) CPL-SLAM: Efficient and Certifiably Correct Planar Graph-Based SLAM Using the Complex Number Representation. *IEEE Transactions on Robotics* 36(6): 1719–1737. DOI:10.1109/TRO.2020.3006717.

Fischler MA and Bolles RC (1981) Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* 24(6): 381–395.

Gallier J (2010) The Schur Complement and Symmetric Positive Semidefinite (and Definite) Matrices .

Garcia-Salguero M, Briales J and Gonzalez-Jimenez J (2021) Certifiable relative pose estimation. *Image and Vision Computing* 109: 104142. DOI:10.1016/j.imavis.2021.104142.

Garstka M, Cannon M and Goulart P (2021) COSMO: A Conic Operator Splitting Method for Convex Conic Problems. *Journal of Optimization Theory and Applications* 190(3): 779–810. DOI:10.1007/s10957-021-01896-x.

Giamou M, Ma Z, Peretroukhin V and Kelly J (2018) Certifiably Globally Optimal Extrinsic Calibration from Per-Sensor Egomotion. *arXiv preprint arXiv:1809.03554* .

Grebe C (2022) *A Study of Observability-Aware Trajectory Optimization*. Master's Thesis, University of Toronto, Toronto, Ontario, Canada.

Grebe C, Wise E and Kelly J (2021) Observability-Aware Trajectory Optimization: Theory, Viability, and State of the Art. In: *2021 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*. Karlsruhe, Germany: IEEE. ISBN 978-1-66544-521-4, pp. 1–8. DOI:10.1109/MFI52462.2021.9591177.

Guillemin V and Pollack A (1974) *Differential Topology*. Prentice-Hall.

Ha J (2023) Probabilistic Framework for Hand–Eye and Robot–World Calibration AX = YB. *IEEE Transactions on Robotics* 39(2): 1196–1211. DOI:10.1109/TRO.2022.3214350.

Hartley R, Trumpf J, Dai Y and Li H (2013) Rotation Averaging. *International Journal of Computer Vision* 103(3): 267–305. DOI:10.1007/s11263-012-0601-0.

Heller J, Havlena M and Pajdla T (2012) A branch-and-bound Algorithm for Globally Optimal Hand-eye Calibration. In: *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, pp. 1608–1615.

Heller J, Havlena M and Pajdla T (2016) Globally Optimal Hand-Eye Calibration Using Branch-and-Bound. *IEEE*

Transactions on Pattern Analysis and Machine Intelligence 38(5): 1027–1033. DOI:10.1109/TPAMI.2015.2469299.

Heller J, Henrion D and Pajdla T (2014) Hand-eye and robot-world calibration by global polynomial optimization. In: *2014 IEEE International Conference on Robotics and Automation (ICRA)*. pp. 3157–3164. DOI:10.1109/ICRA.2014.6907313.

Henderson HV and Searle SR (1981) The vec-permutation matrix, the vec operator and Kronecker products: A review. *Linear and Multilinear Algebra* 9(4): 271–288. DOI:10.1080/03081088108817379.

Holmes C and Barfoot TD (2023) An Efficient Global Optimality Certificate for Landmark-Based SLAM. *IEEE Robotics and Automation Letters* 8(3): 1539–1546. DOI:10.1109/LRA.2023.3238173.

Holmes C, Dümbgen F and Barfoot T (2024) On Semidefinite Relaxations for Matrix-Weighted State-Estimation Problems in Robotics. *IEEE Transactions on Robotics* 40: 4805–4824. DOI:10.1109/TRO.2024.3475220.

Horn M, Wodtko T, Buchholz M and Dietmayer K (2023) Extrinsic Infrastructure Calibration Using the Hand-Eye Robot-World Formulation.

Lee J (2013) *Introduction to Smooth Manifolds*. 2nd edition. New York.

Li AG, Wang L and Wu D (2010) Simultaneous robot-world and hand-eye calibration using dual-quaternions and Kronecker product. *International Journal of Physical Sciences* 5: 1530–1536.

Olson E (2011) AprilTag: A robust and flexible visual fiducial system. In: *2011 IEEE International Conference on Robotics and Automation*. Shanghai, China: IEEE. ISBN 978-1-61284-386-5, pp. 3400–3407. DOI:10.1109/ICRA.2011.5979561.

Papalia A, Fishberg A, O'Neill BW, How JP, Rosen DM and Leonard JJ (2024) Certifiably Correct Range-Aided SLAM. *IEEE Transactions on Robotics* 40: 4265–4283. DOI:10.1109/TRO.2024.3454430.

Rehder J, Nikolic J, Schneider T, Hinzmann T and Siegwart R (2016) Extending Kalibr: Calibrating the extrinsics of multiple IMUs and of individual axes. In: *IEEE Intl. Conf. Robotics and Automation (ICRA)*. Stockholm, Sweden, pp. 4304–4311.

Rosen DM (2021) Scalable Low-Rank Semidefinite Programming for Certifiably Correct Machine Perception. In: LaValle SM, Lin M, Ojala T, Shell D and Yu J (eds.) *Algorithmic Foundations of Robotics XIV*, volume 17. Cham: Springer International Publishing. ISBN 978-3-030-66722-1 978-3-030-66723-8, pp. 551–566. DOI:10.1007/978-3-030-66723-8_33.

Rosen DM, Carlone L, Bandeira AS and Leonard JJ (2019) SE-Sync: A certifiably correct algorithm for synchronization over the special Euclidean group. *The International Journal of Robotics Research* 38(2-3): 95–125. DOI:10.1177/0278364918784361.

Rosen DM, Doherty KJ, Terán Espinoza A and Leonard JJ (2021) Advances in Inference and Representation for Simultaneous Localization and Mapping. *Annual Review of Control, Robotics, and Autonomous Systems* 4(1): 215–242. DOI:10.1146/annurev-control-072720-082553.

Shah M (2013) Solving the Robot-World/Hand-Eye Calibration Problem Using the Kronecker Product. *Journal of Mechanisms and Robotics* 5(3): 031007.

Strobl KH and Hirzinger G (2006) Optimal hand-eye calibration. In: *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*. pp. 4647–4653. DOI:10.1109/IROS.2006.282250.

Tabb A and Yousef KMA (2017) Solving the robot-world hand-eye(s) calibration problem with iterative methods. *Machine Vision and Applications* 28: 569–590.

Tian Y, Khosoussi K, Rosen DM and How JP (2021) Distributed Certifiably Correct Pose-Graph Optimization. *IEEE Transactions on Robotics* 37(6): 2137–2156. DOI:10.1109/TRO.2021.3072346.

Tron R, Rosen DM and Carlone L (2015) On the Inclusion of Determinant Constraints in Lagrangian Duality for 3D SLAM. In: *Robotics: Science and Systems (RSS), Workshop "The Problem of Mobile Sensors: Setting Future Goals and Indicators of Progress for SLAM"*, volume 4. RSS Foundation.

Ulrich M and Hillemann M (2024) Uncertainty-Aware Hand–Eye Calibration. *IEEE Transactions on Robotics* 40: 573–591. DOI:10.1109/TRO.2023.3330609.

Wang Y, Jiang W, Huang K, Schwertfeger S and Kneip L (2022) Accurate calibration of multi-perspective cameras from a generalization of the hand-eye constraint. In: *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 1244–1250.

Wise E, Giamou M, Khoubyarian S, Grover A and Kelly J (2020) Certifiably Optimal Monocular Hand-Eye Calibration. In: *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*. IEEE, pp. 271–278.

Wodtko T, Horn M, Buchholz M and Dietmayer K (2021) Globally Optimal Multi-Scale Monocular Hand-Eye Calibration Using Dual Quaternions. In: *2021 International Conference on 3D Vision (3DV)*. IEEE, pp. 249–257.

Wu J, Zheng Y, Gao Z, Jiang Y, Hu X, Zhu Y, Jiao J and Liu M (2022) Quadratic Pose Estimation Problems: Globally Optimal Solutions, Solvability/Observability Analysis, and Uncertainty Description. *IEEE Transactions on Robotics* 38(5): 3314–3335. DOI:10.1109/TRO.2022.3155880.

Xue B, Zhu Y, Liu T, Wu J, Jiao J, Jiang Y, Zhang C, Jiang X and He Z (2025) *S* QPEP: Global Optimal Solutions to Scaled Quadratic Pose Estimation Problems. *IEEE Transactions on Instrumentation and Measurement* : 1–1DOI:10.1109/TIM.2025.3540135.

Yang H and Carlone L (2020) One ring to rule them all: Certifiably robust geometric perception with outliers. *Advances in neural information processing systems* 33: 18846–18859.

Yang H, Shi J and Carlone L (2021) TEASER: Fast and Certifiable Point Cloud Registration. *IEEE Transactions on Robotics* 37(2): 314–333. DOI:10.1109/TRO.2020.3033695.

Yang J, Rebello J and Waslander SL (2023) Next-best-view selection for robot eye-in-hand calibration. In: *2023 20th Conference on Robots and Vision (CRV)*. IEEE, pp. 161–168.

Yu X and Yang H (2024) SIM-Sync: From Certifiably Optimal Synchronization Over the 3D Similarity Group to Scene Reconstruction With Learned Depth. *IEEE Robotics and Automation Letters* 9(5): 4471–4478. DOI:10.1109/LRA.2024.3377006.

Zhao J (2020) An Efficient Solution to Non-Minimal Case Essential Matrix Estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* : 1–1DOI:10.1109/TPAMI.2020.3030161.

Zhuang H, Roth Z and Sudhakar R (1994) Simultaneous robot/world and tool/flange calibration by solving homogeneous transformation equations of the form AX=YB. *IEEE Transactions on Robotics and Automation* 10(4): 549–554. DOI:10.1109/70.313105.