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

Author Names Omitted for Anonymous Review

Abstract—**Extrinsic sensor calibration has been a longstanding problem when working with with multi-sensor platforms. Since the engineering effort required to obtain accurate calibration parameters increases with the number of sensors deployed, robotics researchers have pursued an efficient and broadly applicable approach to multi-sensor batch extrinsic calibration. In this work, we present a fast and certifiably globally optimal solution to a generalized formulation of the robot world and hand-eye calibration (RWHEC) problem. The formulation of RWHEC used is "generalized" in that it supports the estimation of multiple sensor and target poses, and permits the use of monocular cameras that are unable to measure the scale of their environments. In addition to demonstrating superior performance over existing solutions to generalized RWHEC, we provide the first 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.**

Index Terms—**Calibration and Identification, Optimization and Optimal Control, Sensor Fusion**

I. INTRODUCTION

Calibration is an essential but often painful process when working with common sensors for robot perception. In particular, the *extrinsic* calibration problem of finding the spatial transformations between multiple sensors rigidly mounted to a fixed or mobile sensing platform poses a significant challenge. Existing approaches vary in terms of the type and number of sensors involved, requirements placed on the robot's motion or environment geometry, and operator involvement or expertise.

Errors in a robot's extrinsic calibration can have catastrophic consequences for downstream perception tasks which 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 [1]. The RWHEC problem can be adapted to a wide variety of sensor configurations, and conveniently distills extrinsic calibration down to estimation of two rigid transformation variables, $\mathbf{X}$ and $\mathbf{Y}$, 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 $\mathbf{A}_i$ and $\mathbf{B}_i$. 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 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 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;
3) a novel manifold optimization-based solution to our problem;
4) an analysis of scenarios where local minima can degrade the solution of nonlinear solvers;
5) an extensive set of synthetic experiments demonstrating the superior speed and accuracy of our approach compared with local or approximate methods;
6) a real-world experiment demonstrating the performance of our algorithm on a robot arm-camera system; and
7) 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 II, followed by a summary of our mathematical notation in Section III. We develop a detailed description of our generalized formulation of RWHEC in Section IV. Section V describes a convex relaxation of RWHEC that leads to a semidefinite program (SDP), and Section VI presents identifiability criteria in the form of geometric constraints on sensor measurements and motion. Section VII proves that our SDP relaxations of identifiable problems with sufficiently low noise are guaranteed to be tight, enabling the extraction of the global minimum

---

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

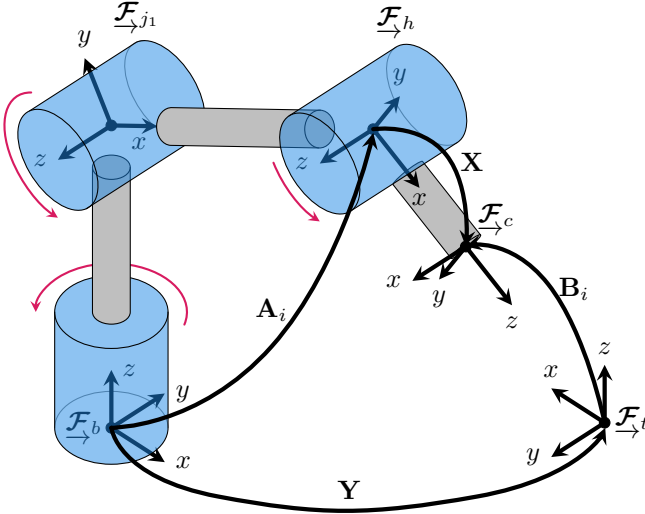[2]Our code will be released at the end of the anonymous review period.

Fig. 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 t_i, we use the forward kinematics of the manipulator to estimate the transform from the manipulator base to the wrist, $\mathbf{A}_i$. Further, we can measure the transform from the target to the camera, $\mathbf{B}_i$.

to the original nonconvex problem. In Section VIII and Section IX, we experimentally demonstrate the superior properties of our RWHEC method as compared with existing solution procedures in the literature on synthetic and real-world data respectively. Finally, Section X discusses remaining challenges and promising avenues for future work.

II. RELATED WORK

In this section, we begin by defining and reviewing three common categories of RWHEC methods: separate, joint, and probabilistic. The characteristics of all methods discussed are summarized in Table I, 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 II-B.

A. Hand-Eye and Robot-World 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 in [2] and extended to multiple $\mathbf{X}$ or $\mathbf{Y}$ variables in the generalized RWHEC formulation of [4]. The identifiability requirements of RWHEC are explored in [3], but to our knowledge these 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.

To decrease this sensitivity of RWHEC to measurement noise, some methods jointly estimate the translation and rotation components of $\mathbf{X}$ and $\mathbf{Y}$. In [5], the authors introduce a variety of joint RWHEC cost functions and parameterizations that are less sensitive to measurement noise than two-stage closed-form RWHEC schemes. Using a dual quaternion (DQ) representation for poses, the authors of [6] cast the generalized RWHEC problem as a QCQP and use a methodology similar to ours to solve a convex relaxation of their problem formulation. One can also eschew pose sensor abstractions altogether and develop a graph-based approach to generalized RWHEC using camera reprojection error terms [7]. This formulation enables the algorithm to jointly estimate the base-target transform, end-effector-camera transform, camera intrinsic parameters, and camera distortion parameters. 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 joint estimation methods do not have a mechanism for incorporating knowledge of relative pose measurement accuracy.

Probabilistic RWHEC algorithms are presented in [8] and [9]. The method in [8] solves a nonlinear probabilistic formulation of for RWHEC, but does not use an on-manifold method like the local solver presented in this work. In [9], the authors 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.

B. Certifiably Correct Estimation

Convex semidefinite programming (SDP) relaxations of quadratically constrained quadratic programs (QCQPs) are a powerful tool for designing globally optimal algorithms to various geometric estimation problems in robotics and computer vision [10], [11]. Most notably, the SE-Sync algorithm introduced in [12] was the first efficient and globally optimal algorithm for simultaneous localization and mapping (SLAM). SE-Sync revealed and exploited the smooth manifold structure of an SDP relaxation of a QCQP formulation of pose-graph SLAM with a principled maximum likelihood (MLE) objective function. This approach spawned a host of solutions to spatial perception problems including point cloud registration [13], multiple variants of localization and mapping [14], [15], [16], [17], [18], hand-eye calibration [19], [20], [21], camera pose estimation [22], [23], as well as research into tools and optimization methods for user-specified problems [24], [25], [26].

An alternative approach for obtaining global optima for noisy geometric estimation problems is outlined in [27]. In contrast to the SDP relaxations, Wu et al. use a Gröbner basis method to solve polynomial optimization problems over a single pose variable. In [28], this symbolic method is extended to problems with unknown scale, including hand-eye calibration with a monocular camera. While promising,

TABLE I: 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	Representation
Li et al. [2]	✗	✗	✗	✗	✗	DQ
Shah [3]	✗	✗	✗	✗	✗	SE(3)
Wang et al. [4]	✗	✗	✗	$\mathbf{X}$s or $\mathbf{Y}$s	✗	SE(3)
Tabb and Yousef [5]	✓	✗	✗	✗	✗	Multiple
Horn et al. [6]	✓	✗	✓	$\mathbf{X}$s or $\mathbf{Y}$s	✗	DQ
Evangelista et al. [7]	✓	✗	✗	$\mathbf{X}$s	✗	SE(3)
Dornaika and Horaud [8]	✓	✓	✗	✗	✗	Multiple
Strobl and Hirzinger [9]	✓	✓	✗	✗	✗	SE(3)
Ours	✓	✓	✓	$\mathbf{X}$s and $\mathbf{Y}$s	✓	SE(3)

this approach has not been demonstrated to scale efficiently to problems with many unknown pose variables.

In this work, we extend the multi-frame RWHEC formulation of [4] and apply an SDP relaxation. This leads to an estimation problem over a graph similar in nature to SE-Sync [12], but without $\mathrm{SO}(d)$-invariance and rank-d solution structure required to exploit fast manifold optimization techniques. The QCQP formulation using a unit dual quaternion (DQ) representation of $\mathrm{SE}(3)$ in [6] is most similar to our approach in that the authors apply a semidefinite relaxation to a limited form of generalized RWHEC in the context of infrastructure calibration for intelligent transportation systems. However, we demonstrate in Section IX that the unit DQ representation leads to poorer performance than the homogeneous transformation matrices we use. This is most likely because unit quaternions are a *double cover* of $\mathrm{SO}(3)$ in that the quadratic mapping

$$\mathbf{R}(\mathbf{q}) = \mathbf{v}\mathbf{v}^\top + w^2\mathbf{I} + 2w\mathbf{v}^\wedge + (\mathbf{v}^\wedge)^2 \qquad (2)$$

from a unit quaternion $\mathbf{q} = [\mathbf{v}^\top \ w]^\top \in \mathbb{R}^4$ to $\mathbf{R} \in \mathrm{SO}(3)$ clearly maps $-\mathbf{q}$ to the same rotation matrix. 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. Finally, we also prove fundamental identifiability and global optimality theorems for our novel formulation of RWHEC.

III. 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 denoted $\mathbb{S}^n$ 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 directed graph $\vec{\mathcal{G}} = (\mathcal{V}, \vec{\mathcal{E}})$, we write each directed edges 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 Cartesian reference frame is denoted $\underrightarrow{\mathcal{F}}$. 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 rotation from $\underrightarrow{\mathcal{F}}_b$ to $\underrightarrow{\mathcal{F}}_a$ is denoted $\mathbf{R}_{ab} \in \mathrm{SO}(3)$. The rigid special Euclidean transformation from $\underrightarrow{\mathcal{F}}_b$ to $\underrightarrow{\mathcal{F}}_a$ is

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

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

$$\mathbf{p}^\wedge = \begin{bmatrix} 0 & -p_3 & p_2 \\ p_3 & 0 & -p_1 \\ -p_2 & p_1 & 0 \end{bmatrix}, \qquad (4)$$

and the operator $(\cdot)^\vee$ denotes its inverse. The Kronecker product $\otimes$ is

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} A_{1,1}\mathbf{B} & \cdots & A_{1,N}\mathbf{B} \\ \vdots & \ddots & \vdots \\ A_{M,1}\mathbf{B} & \cdots & A_{M,N}\mathbf{B} \end{bmatrix}, \qquad (5)$$

and the Kronecker sum $\oplus$ is

$$\mathbf{A} \oplus \mathbf{B} = \mathbf{A} \otimes \mathbf{I} + \mathbf{I} \otimes \mathbf{B}. \qquad (6)$$

The function $\mathrm{BlkDiag}(\mathbf{A}_1, \ldots, \mathbf{A}_N)$ creates a block diagonal matrix with its ordered matrix arguments on the diagonal. A vector that is Gaussian distributed with a mean of $\boldsymbol{\mu}$ and covariance of $\boldsymbol{\Sigma}$ is denoted $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Finally, a rotation matrix that is Langevin distributed with mode $\mathbf{M}$ and concentration $\kappa \geq 0$ is denoted $\mathrm{Lang}(\mathbf{M}, \kappa)$.

IV. GENERALIZED ROBOT-WORLD AND HAND-EYE CALIBRATION

In Section IV-A, we review the geometric constraints of the robot-world and hand-eye calibration problem. In Section IV-B, we formulated RWHEC as maximum likelihood estimation. In Section IV-C, we extend our problem formulation to support monocular cameras observing targets of unknown size. In Section IV-D, we convert our calibration problems into QCQPs in standard form. In Section IV-E, 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, Section IV-F presents a dimensionality reduction for our calibration problem.

A. 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 [4]), 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 $\overrightarrow{\mathscr{F}}_b, \overrightarrow{\mathscr{F}}_h, \overrightarrow{\mathscr{F}}_t, \overrightarrow{\mathscr{F}}_c$ 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 timestep t_i indexed by $i \in [N]$, these two measurements are related by

$$\mathbf{T}_{bh}(t_i)\mathbf{T}_{hc} = \mathbf{T}_{bt}\mathbf{T}_{tc}(t_i). \tag{7}$$

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

$$\begin{aligned} \mathbf{A}_i &\triangleq \mathbf{T}_{bh}(t_i) \\ \mathbf{X} &\triangleq \mathbf{T}_{hc} \\ \mathbf{Y} &\triangleq \mathbf{T}_{bt} \\ \mathbf{B}_i &\triangleq \mathbf{T}_{tc}(t_i), \end{aligned} \tag{8}$$

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

$$\mathbf{A}_i\mathbf{X} = \mathbf{Y}\mathbf{B}_i. \tag{9}$$

In Equation (9), the rotation and translation constraints can be separated into

$$\mathbf{R}_{\mathbf{A}_i}\mathbf{R}_{\mathbf{X}} = \mathbf{R}_{\mathbf{Y}}\mathbf{R}_{\mathbf{B}_i}, \tag{10}$$

$$\mathbf{R}_{\mathbf{A}_i}\mathbf{t}_{\mathbf{X}} + \mathbf{t}_{\mathbf{A}_i} = \mathbf{R}_{\mathbf{Y}}\mathbf{t}_{\mathbf{B}_i} + \mathbf{t}_{\mathbf{Y}}, \tag{11}$$

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.

B. Maximum Likelihood Estimation

Using Equations (10) and (11), 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 $\mathbf{x}$. Let us assume that

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) $\tilde{\mathbf{t}}_{\mathbf{B}_i} = \underline{\mathbf{t}}_{\mathbf{B}_i} + \mathbf{n}_i$, where $\mathbf{n}_i \sim \mathcal{N}\left(0, \sigma_i{}^2\mathbf{I}\right)$ and σ_i is the standard deviation of the translation component of measurements $\mathbf{B}_i$ $\forall\, i = 1, \dots, N$; and
3) $\tilde{\mathbf{R}}_{\mathbf{B}_i} = \underline{\mathbf{R}}_{\mathbf{B}_i}\mathbf{R}_m$, where $\mathbf{R}_m \sim \mathrm{Lang}(\mathbf{I}, \kappa_i)$, 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 scenarios [29], 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 (8) and henceforth use

$$\begin{aligned} \mathbf{A}_i &\triangleq \mathbf{T}_{bh}(t_i) = \begin{bmatrix} \mathbf{R}_{\mathbf{A}_i} & \mathbf{t}_{\mathbf{A}_i} \\ \mathbf{0}^\top & 1 \end{bmatrix} \\ \mathbf{B}_i &\triangleq \tilde{\mathbf{T}}_{tc}(t_i) = \begin{bmatrix} \tilde{\mathbf{R}}_{\mathbf{B}_i} & \tilde{\mathbf{t}}_{\mathbf{B}_i} \\ \mathbf{0}^\top & 1 \end{bmatrix}. \end{aligned} \tag{12}$$

We wish to find the solution to the following maximum likelihood problem:

$$\max_{\mathbf{X}, \mathbf{Y} \in \mathrm{SE}(3)} p(\{\tilde{\mathbf{B}}_1, \dots, \tilde{\mathbf{B}}_N\}|\mathbf{X}, \mathbf{Y}). \tag{13}$$

Since the random variables $\tilde{\mathbf{R}}_{\mathbf{B}_i}$ and $\tilde{\mathbf{t}}_{\mathbf{B}_i}$ are 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 (13) gives us the following equivalent minimization problem:

$$\min_{\mathbf{X}, \mathbf{Y} \in \mathrm{SE}(3)} -\sum_{i=1}^N (\log(p(\tilde{\mathbf{t}}_{\mathbf{B}_i}|\mathbf{X}, \mathbf{Y})) + \log(p(\tilde{\mathbf{R}}_{\mathbf{B}_i}|\mathbf{X}, \mathbf{Y}))). \tag{14}$$

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

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

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

$$\log p(\tilde{\mathbf{t}}_{\mathbf{B}_i}|\mathbf{X}, \mathbf{Y}) = -\frac{1}{2\sigma_i^2} \left\| \mathbf{R}_{\mathbf{A}_i}\mathbf{t}_{\mathbf{X}} + \underline{\mathbf{t}}_{\mathbf{A}_i} - \mathbf{t}_{\mathbf{Y}} - \mathbf{R}_{\mathbf{Y}}\tilde{\mathbf{t}}_{\mathbf{B}_i} \right\|_2^2, \tag{16}$$

which we recognize as a weighted error function for the translation constraint in Equation (11). Similarly, the log-likelihood of $p(\tilde{\mathbf{R}}_{\mathbf{B}_i}|\mathbf{X}, \mathbf{Y})$ is [12]

$$\log p(\tilde{\mathbf{R}}_{\mathbf{B}_i}|\mathbf{X}, \mathbf{Y}) = -c(\kappa_i) - \frac{\kappa_i}{2} \left\| \mathbf{R}_{\mathbf{A}_i}\mathbf{R}_{\mathbf{X}} - \mathbf{R}_{\mathbf{Y}}\tilde{\mathbf{R}}_{\mathbf{B}_i} \right\|_{\mathrm{F}}^2 + 3, \tag{17}$$

which is a weighted error function for the rotation constraint in Equation (10). Our MLE problem is the following QCQP:

Problem 1. *Maximum Likelihood Estimation for RWHEC*

$$\min_{\substack{\mathbf{R}_{\mathbf{X}}, \mathbf{R}_{\mathbf{Y}} \in \mathrm{SO}(3) \\ \mathbf{t}_{\mathbf{X}}, \mathbf{t}_{\mathbf{Y}} \in \mathbb{R}^3, s^2 = 1}} J_{\mathbf{t}} + J_{\mathbf{R}}, \tag{18}$$

where

$$J_{\mathbf{t}} \triangleq \frac{1}{2} \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \left\| \underline{\mathbf{R}}_{\mathbf{A}_i} \mathbf{t}_{\mathbf{X}} + s\underline{\mathbf{t}}_{\mathbf{A}_i} - \mathbf{t}_{\mathbf{Y}} - \mathbf{R}_{\mathbf{Y}} \tilde{\mathbf{t}}_{\mathbf{B}_i} \right\|_2^2 \quad (19)$$

$$J_{\mathbf{R}} \triangleq \frac{1}{2} \sum_{i=1}^{N} \kappa_i \left\| \underline{\mathbf{R}}_{\mathbf{A}_i} \mathbf{R}_{\mathbf{X}} - \mathbf{R}_{\mathbf{Y}} \tilde{\mathbf{R}}_{\mathbf{B}_i} \right\|_{\mathrm{F}}^2. \quad (20)$$

Note that to ensure all terms are quadratic and simplify the SDP relaxation in Section V, we have homogenized Equation (19) with the quadratically constrained variable $s^2 = 1$.

C. 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 IV-B, we replace the camera translation measurement model with

$$\tilde{\mathbf{t}}_{\mathbf{B}_i} = \alpha \mathbf{R}_{\mathbf{Y}}^\top (\underline{\mathbf{R}}_{\mathbf{A}_i} \mathbf{t}_{\mathbf{X}} + \underline{\mathbf{t}}_{\mathbf{A}_i} - \mathbf{t}_{\mathbf{Y}}) + \mathbf{n}_i, \quad (21)$$

where $\alpha \in \mathbb{R}$ is the unknown scale, $\mathbf{n}_i \sim \mathcal{N}\left(0, \sigma_i^2 \mathbf{I}\right)$, 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*

$$\min_{\substack{\mathbf{R}_{\mathbf{X}}, \mathbf{R}_{\mathbf{Y}} \in \mathrm{SO}(3) \\ \mathbf{t}_{\mathbf{X},\alpha}, \mathbf{t}_{\mathbf{Y},\alpha} \in \mathbb{R}^3, \, \alpha \in \mathbb{R}}} J_{\mathbf{t},\alpha} + J_{\mathbf{R}}, \quad (22)$$

where

$$J_{\mathbf{t},\alpha} \triangleq \frac{1}{2} \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \left\| \underline{\mathbf{R}}_{\mathbf{A}_i} \mathbf{t}_{\mathbf{X},\alpha} + \alpha\underline{\mathbf{t}}_{\mathbf{A}_i} - \mathbf{R}_{\mathbf{Y}} \tilde{\mathbf{t}}_{\mathbf{B}_i} - \mathbf{t}_{\mathbf{Y},\alpha} \right\|_2^2. \quad (23)$$

To maintain our assumption from Section IV-B 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 removed the need for the homogenizing variable s used in Problem 1. To recover the calibration parameters, we can correct the estimated system scale using $\frac{1}{\alpha} \mathbf{t}_{\mathbf{X},\alpha}$ and $\frac{1}{\alpha} \mathbf{t}_{\mathbf{Y},\alpha}$.

D. Quadratically Constrained Quadratic Programming

Herein we convert Problems 1 and 2 to a standard QCQP form with a vectorized decision variable and constraints defined by real symmetric matrices. The state vectors for the standard and monocular RWHEC problems are, respectively,

$$\mathbf{x}^\top = \begin{bmatrix} \mathbf{t}_{\mathbf{X}}^\top & \mathbf{t}_{\mathbf{Y}}^\top & \mathbf{r}_{\mathbf{X}}^\top & \mathbf{r}_{\mathbf{Y}}^\top & s \end{bmatrix}^\top, \quad (24)$$

$$\mathbf{x}_\alpha^\top = \begin{bmatrix} \mathbf{t}_{\mathbf{X},\alpha}^\top & \mathbf{t}_{\mathbf{Y},\alpha}^\top & \alpha & \mathbf{r}_{\mathbf{X}}^\top & \mathbf{r}_{\mathbf{Y}}^\top & s \end{bmatrix}^\top, \quad (25)$$

where $\mathbf{r}_{\mathbf{X}} = \mathrm{vec}\left(\mathbf{R}_{\mathbf{X}}\right)$ and $\mathbf{r}_{\mathbf{Y}} = \mathrm{vec}\left(\mathbf{R}_{\mathbf{Y}}\right)$. Importantly, the rotation matrix constraints require homogenization, so we have appended the homogenization variable to the state vector for the monocular RWHEC problem. With these state variables,

the standard and monocular RWHEC rotation cost terms can be written as

$$J_{\mathbf{R}} = \sum_{i=1}^{N} \kappa_i \, \mathbf{x}^\top \mathbf{M}_{\mathbf{R}_i}^\top \mathbf{M}_{\mathbf{R}_i} \mathbf{x}, \quad (26)$$

$$J_{\mathbf{R},\alpha} = \sum_{i=1}^{N} \kappa_i \, \mathbf{x}_\alpha^\top \mathbf{M}_{\mathbf{R}_i,\alpha}^\top \mathbf{M}_{\mathbf{R}_i,\alpha} \mathbf{x}_\alpha, \quad (27)$$

where

$$\mathbf{M}_{\mathbf{R}_i} = \begin{bmatrix} \mathbf{0}_{9 \times 6} & \mathbf{I} \otimes \underline{\mathbf{R}}_{\mathbf{A}_i} & -\tilde{\mathbf{R}}_{\mathbf{B}_i} \otimes \mathbf{I} & 0 \end{bmatrix}, \quad (28)$$

$$\mathbf{M}_{\mathbf{R}_i,\alpha} = \begin{bmatrix} \mathbf{0}_{9 \times 7} & \mathbf{I} \otimes \underline{\mathbf{R}}_{\mathbf{A}_i} & -\tilde{\mathbf{R}}_{\mathbf{B}_i} \otimes \mathbf{I} & 0 \end{bmatrix}. \quad (29)$$

These matrices and many to follow are obtained through a straightforward application of the column-major vectorization identity [30]

$$\mathrm{vec}\left(\mathbf{A}\mathbf{X}\mathbf{B}\right) = (\mathbf{B}^\top \otimes \mathbf{A})\mathrm{vec}\left(\mathbf{X}\right), \quad (30)$$

where $\mathbf{A}$, $\mathbf{X}$, and $\mathbf{B}$ are any compatible matrices. The standard and monocular RWHEC translation cost terms can now be written as

$$J_{\mathbf{t}} = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \, \mathbf{x}^\top \mathbf{M}_{\mathbf{t}_i}^\top \mathbf{M}_{\mathbf{t}_i} \mathbf{x}, \quad (31)$$

$$J_{\mathbf{t},\alpha} = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \, \mathbf{x}_\alpha^\top \mathbf{M}_{\mathbf{t}_i,\alpha}^\top \mathbf{M}_{\mathbf{t}_i,\alpha} \mathbf{x}_\alpha, \quad (32)$$

respectively, and

$$\mathbf{M}_{\mathbf{t}_i} = \begin{bmatrix} \underline{\mathbf{R}}_{\mathbf{A}_i} & -\mathbf{I} & \mathbf{0}_{3 \times 9} & -\tilde{\mathbf{t}}_{\mathbf{B}_i}^\top \otimes \mathbf{I} & \underline{\mathbf{t}}_{\mathbf{A}_i} \end{bmatrix}, \quad (33)$$

$$\mathbf{M}_{\mathbf{t}_i,\alpha} = \begin{bmatrix} \underline{\mathbf{R}}_{\mathbf{A}_i} & -\mathbf{I} & \underline{\mathbf{t}}_{\mathbf{A}_i} & \mathbf{0}_{3 \times 9} & -\tilde{\mathbf{t}}_{\mathbf{B}_i}^\top \otimes \mathbf{I} & \mathbf{0}_{3 \times 1} \end{bmatrix}. \quad (34)$$

The symmetric standard and monocular RWHEC cost matrices are

$$\mathbf{Q} = \sum_{i=1}^{N} \kappa_i \mathbf{M}_{\mathbf{R}_i}^\top \mathbf{M}_{\mathbf{R}_i} + \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \mathbf{M}_{\mathbf{t}_i}^\top \mathbf{M}_{\mathbf{t}_i}, \quad (35)$$

$$\mathbf{Q}_\alpha = \sum_{i=1}^{N} \kappa_i \mathbf{M}_{\mathbf{R}_i,\alpha}^\top \mathbf{M}_{\mathbf{R}_i,\alpha} + \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \mathbf{M}_{\mathbf{t}_i,\alpha}^\top \mathbf{M}_{\mathbf{t}_i,\alpha}. \quad (36)$$

Consequently, we can rewrite the standard and unscaled RWHEC optimization problems as

$$\min_{\mathbf{x}} \mathbf{x}^\top \mathbf{Q}\mathbf{x}, \\ \text{s.t. } \mathbf{R}_{\mathbf{X}}, \mathbf{R}_{\mathbf{Y}} \in \mathrm{SO}(3), \quad (37)$$

and

$$\min_{\mathbf{x}_\alpha} \mathbf{x}_\alpha^\top \mathbf{Q}_\alpha \mathbf{x}_\alpha, \\ \text{s.t. } \mathbf{R}_{\mathbf{X}}, \mathbf{R}_{\mathbf{Y}} \in \mathrm{SO}(3), \quad (38)$$

respectively.

E. Generalizing RWHEC for Multiple Frames

We can extend robot-world and hand-eye calibration to robots employing more than one sensor, and calibration procedures involving more than one target [4]. 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
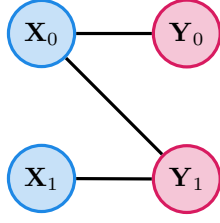
Fig. 2: An example of the bipartite graph structure of a simple generalized RWHEC problem with multiple $\mathbf{X}$s and $\mathbf{Y}$s. Each edge corresponds to measurements forming an instance of Problem 1 or Problem 2.

$\mathbf{Y}_1, \ldots, \mathbf{Y}_P$. 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 $\mathcal{V} \triangleq \mathcal{V}_{\mathbf{X}} \cup \mathcal{V}_{\mathbf{Y}} = [M + P]$, and each edge $e = (i, j) \in \vec{\mathcal{E}} \subseteq \mathcal{V}_{\mathbf{X}} \times \mathcal{V}_{\mathbf{Y}}$ represents a set of observations of target j by camera i. This multi-frame approach was introduced in [4] 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 [12]. One notable feature of our formulation is that each edge $e = (j, k) \in \vec{\mathcal{E}}$ is associated with a *set* $\mathcal{D}_e$ of $N_e \geq 1$ noisy observations involving unknown variables $\mathbf{X}_j$ and $\mathbf{Y}_k$:

$$\mathcal{D}_e \triangleq \{(\mathbf{A}_{e,i}, \mathbf{B}_{e,i}, \sigma_{e,i}, \kappa_{e,i}) \in \mathrm{SE}(3)^2 \times \mathbb{R}_+^2 \mid i \in [N_e]\}. \tag{39}$$

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 $\mathcal{V}$. In order to simplify the presentation of our formulation, we will assume that the uncertainty parameters $\sigma_{e,i}$ and $\kappa_{e,i}$ do not vary within the data $\mathcal{D}_e$ associated with each edge $e \in \vec{\mathcal{E}}$:[3]

$$\sigma_{e,i} = \sigma_e, \ \kappa_{e,i} = \kappa_e \ \forall i \in [N_e]. \tag{40}$$

We can simplify our notation by introducing several matrices structured by blocks mirroring the graph Laplacian (cf. the formulation of SE-Sync in [12]). Let $L(G^\tau) \in \mathbb{S}^{3(M+P)}$ be the symmetric (3×3)-block-structured matrix of the form

$$L(G^\tau)_{jk} = \begin{cases} \sum_{e \in \delta^-(j)} \frac{1}{\sigma_e^2} N_e \mathbf{I} & j = k \\ -\frac{1}{\sigma_{j,k}^2} \sum_{i=1}^{N_{j,k}} \mathbf{R}_{\mathbf{A}_{(j,k),i}}^\top & (j, k) \in \vec{\mathcal{E}} \\ -\frac{1}{\sigma_{k,j}^2} \sum_{i=1}^{N_{k,j}} \mathbf{R}_{\mathbf{A}_{(k,j),i}} & (k, j) \in \vec{\mathcal{E}} \\ \mathbf{0} & \text{otherwise.} \end{cases} \tag{41}$$

Similarly, let $L(G^\rho) \in \mathbb{S}^{9(M+P)}$ be the symmetric (9×9)-

block-structured matrix

$$L(G^\rho)_{jk} = \begin{cases} \sum_{e \in \delta^-(j)} \kappa_e N_e \mathbf{I} \otimes \mathbf{I} & j = k \\ -\kappa_{j,k} \sum_{i=1}^{N_{j,k}} \tilde{\mathbf{R}}_{\mathbf{B}_{(j,k),i}} \otimes \mathbf{R}_{\mathbf{A}_{(j,k),i}}^\top & (j, k) \in \vec{\mathcal{E}} \\ -\kappa_{k,j} \sum_{i=1}^{N_{k,j}} \tilde{\mathbf{R}}_{\mathbf{B}_{(k,j),i}}^\top \otimes \mathbf{R}_{\mathbf{A}_{(k,j),i}} & (k, j) \in \vec{\mathcal{E}} \\ \mathbf{0} & \text{otherwise.} \end{cases} \tag{42}$$

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

$$\mathbf{v}_l = \begin{cases} \sum_{e \in \delta^-(l)} \frac{1}{\sigma_e^2} \sum_{i=1}^{N_e} \mathbf{R}_{\mathbf{A}_{e,i}}^\top \underline{\mathbf{t}}_{\mathbf{A}_{e,i}} & l \leq M \\ -\sum_{e \in \delta^+(l)} \frac{1}{\sigma_e^2} \sum_{i=1}^{N_e} \underline{\mathbf{t}}_{\mathbf{A}_{e,i}} & \text{otherwise,} \end{cases} \tag{43}$$

and

$$v = \sum_{e \in \vec{\mathcal{E}}} \frac{1}{\sigma_e^2} \sum_{i=1}^{N_e} \underline{\mathbf{t}}_{\mathbf{A}_{e,i}}^\top \underline{\mathbf{t}}_{\mathbf{A}_{e,i}}. \tag{44}$$

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

$$\boldsymbol{\Sigma}_{jk} = \begin{cases} \sum_{e \in \delta^+(j)} \frac{1}{\sigma_e^2} \sum_{i=1}^{N_e} (\tilde{\mathbf{t}}_{\mathbf{B}_{e,i}} \tilde{\mathbf{t}}_{\mathbf{B}_{e,i}}^\top) \otimes \mathbf{I} & j = k \\ \mathbf{0} & \text{otherwise.} \end{cases} \tag{45}$$

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

$$\mathbf{U}_{jk} = \begin{cases} \sum_{e \in \delta^+(j)} \frac{1}{\sigma_e^2} \sum_{i=1}^{N_e} \tilde{\mathbf{t}}_{\mathbf{B}_{e,i}}^\top \otimes \mathbf{I} & j = k \\ -\frac{1}{\sigma_e^2} \sum_{i=1}^{N_{jk}} \tilde{\mathbf{t}}_{\mathbf{B}_{(j,k),i}}^\top \otimes \mathbf{R}_{\mathbf{A}_{(j,k),i}}^\top & (j, k) \in \vec{\mathcal{E}}, \\ \mathbf{0} & \text{otherwise} \end{cases} \tag{46}$$

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

$$\mathbf{u}_l = \begin{cases} \mathbf{0} & l \leq M \\ -\sum_{e \in \delta^+(l)} \sum_{i=1}^{N_e} \tilde{\mathbf{t}}_{\mathbf{B}_{e,i}} \otimes \underline{\mathbf{t}}_{\mathbf{A}_{e,i}} & \text{otherwise.} \end{cases} \tag{47}$$

Using the matrices in Equations (41) to (47), we can define the standard and monocular RWHEC objective function matrices as follows:

$$\mathbf{Q} = \begin{bmatrix} L(G^\tau) & \mathbf{U} & \mathbf{v} \\ \mathbf{U}^\top & \boldsymbol{\Sigma} + L(G^\rho) & \mathbf{u} \\ \mathbf{v}^\top & \mathbf{u}^\top & v \end{bmatrix}, \tag{48}$$

$$\mathbf{Q}_\alpha = \begin{bmatrix} L(G^\tau) & \mathbf{v} & \mathbf{U} & \mathbf{0} \\ \mathbf{v}^\top & v & \mathbf{u}^\top & \mathbf{0} \\ \mathbf{U}^\top & \mathbf{u} & \boldsymbol{\Sigma} + L(G^\rho) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \tag{49}$$

Note that the standard and monocular matrices share most of their entries, differing only by values inserted in a few rows and columns of $\mathbf{Q}_\alpha$. With this new notation, the standard and monocular RWHEC optimization problems for multiple $\mathbf{X}$s and $\mathbf{Y}$s can be stated as QCQPs with objective functions described by quadratic forms:

Problem 3. *QCQP Formulation of the Standard Generalized RWHEC Problem*

$$\min_{\mathbf{x}} \ \mathbf{x}^\top \mathbf{Q} \mathbf{x}$$
$$\text{s.t. } \mathbf{R}_{\mathbf{X}_j} \in \mathrm{SO}(3) \ \forall j \in [M] \tag{50}$$
$$\mathbf{R}_{\mathbf{Y}_k} \in \mathrm{SO}(3) \ \forall k \in [P],$$

and

---

[3]The Julia implementation of our method used in the experiments of Sections VIII and IX supports the use of inhomogeneous measurement precision within edge data $\mathcal{D}_e$.

Problem 4. *QCQP Formulation of the Monocular Generalized RWHEC Problem*

$$\min_{\mathbf{x}_\alpha} \mathbf{x}_\alpha^\top \mathbf{Q}_\alpha \mathbf{x}_\alpha$$
$$\text{s.t. } \mathbf{R}_{\mathbf{X}_j} \in \mathrm{SO}(3) \; \forall \, j \in [M] \tag{51}$$
$$\mathbf{R}_{\mathbf{Y}_k} \in \mathrm{SO}(3) \; \forall \, k \in [P].$$

The state vectors are

$$\mathbf{x}^\top = \begin{bmatrix} \mathbf{t}^\top & \mathbf{r}^\top & s \end{bmatrix}$$
$$\mathbf{x}_\alpha^\top = \begin{bmatrix} \mathbf{t}_\alpha^\top & \alpha & \mathbf{r}^\top & s \end{bmatrix}, \tag{52}$$

where

$$\mathbf{t}^\top = \begin{bmatrix} \mathbf{t}_x^\top & \mathbf{t}_y^\top \end{bmatrix}$$
$$\mathbf{t}_x^\top = \begin{bmatrix} \mathbf{t}_{\mathbf{X}_1}^\top & \cdots & \mathbf{t}_{\mathbf{X}_M}^\top \end{bmatrix}$$
$$\mathbf{t}_y^\top = \begin{bmatrix} \mathbf{t}_{\mathbf{Y}_1}^\top & \cdots & \mathbf{t}_{\mathbf{Y}_P}^\top \end{bmatrix}$$
$$\mathbf{t}_\alpha^\top = \begin{bmatrix} \mathbf{t}_{x,\alpha}^\top & \mathbf{t}_{y,\alpha}^\top \end{bmatrix} \tag{53}$$
$$\mathbf{t}_{x,\alpha}^\top = \begin{bmatrix} \mathbf{t}_{\mathbf{X}_1,\alpha}^\top & \cdots & \mathbf{t}_{\mathbf{X}_M,\alpha}^\top \end{bmatrix}$$
$$\mathbf{t}_{y,\alpha}^\top = \begin{bmatrix} \mathbf{t}_{\mathbf{Y}_1,\alpha}^\top & \cdots & \mathbf{t}_{\mathbf{Y}_P,\alpha}^\top \end{bmatrix},$$

and

$$\mathbf{r}^\top = \begin{bmatrix} \mathbf{r}_x^\top & \mathbf{r}_y^\top \end{bmatrix}$$
$$\mathbf{r}_x^\top = \begin{bmatrix} \mathrm{vec}\left(\mathbf{R}_{\mathbf{X}_1}\right)^\top & \cdots & \mathrm{vec}\left(\mathbf{R}_{\mathbf{X}_M}\right)^\top \end{bmatrix} \tag{54}$$
$$\mathbf{r}_y^\top = \begin{bmatrix} \mathrm{vec}\left(\mathbf{R}_{\mathbf{Y}_1}\right)^\top & \cdots & \mathrm{vec}\left(\mathbf{R}_{\mathbf{Y}_P}\right)^\top \end{bmatrix}.$$

F. Reducing the Dimension of the QCQP

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

$$\mathbf{t}^\star = -L(G^\tau)^{-1} \mathbf{U} \mathbf{r}^\star, \tag{55}$$

$$\begin{bmatrix} \mathbf{t}_\alpha^{\star\top} & \alpha^\star \end{bmatrix}^\top = -\begin{bmatrix} L(G^\tau) & \mathbf{v} \\ \mathbf{v}^\top & v \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{U} \\ \mathbf{u}^\top \end{bmatrix} \mathbf{r}^\star. \tag{56}$$

Using the Schur complement, we reduce the cost matrices to

$$\mathbf{Q}_{red} = \begin{bmatrix} \boldsymbol{\Sigma} + L(G^\rho) & \mathbf{u} \\ \mathbf{u}^\top & v \end{bmatrix} - \begin{bmatrix} \mathbf{U}^\top \\ \mathbf{v}^\top \end{bmatrix} L(G^\tau)^{-1} \begin{bmatrix} \mathbf{U} & \mathbf{v} \end{bmatrix}, \tag{57}$$

$$\mathbf{Q}_{\alpha,red} = \begin{bmatrix} \boldsymbol{\Sigma} + L(G^\rho) & \mathbf{0} \\ \mathbf{0}^\top & 0 \end{bmatrix} - \begin{bmatrix} \mathbf{U}^\top \\ \mathbf{u}^\top \end{bmatrix} \begin{bmatrix} L(G^\tau) & \mathbf{v} \\ \mathbf{v}^\top & v \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{U} & \mathbf{u} \end{bmatrix}. \tag{58}$$

The reduced form of the problem is

Problem 5. *Reduced QCQP Formulation of RWHEC.*

$$\min_{\mathbf{r},s} \mathbf{r}^\top \mathbf{Q}_{red} \mathbf{r},$$
$$\text{s.t. } \mathbf{R}_{\mathbf{X}_j} \in \mathrm{SO}(3) \; \forall \, j \in [M], \tag{59}$$
$$\mathbf{R}_{\mathbf{Y}_k} \in \mathrm{SO}(3) \; \forall \, k \in [P],$$
$$s^2 = 1.$$

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

$$\min_{\mathbf{r},s} \mathbf{r}^\top \mathbf{Q}_{\alpha,red} \mathbf{r},$$
$$\text{s.t. } \mathbf{R}_{\mathbf{X}_j} \in \mathrm{SO}(3) \; \forall \, j \in [M], \tag{60}$$
$$\mathbf{R}_{\mathbf{Y}_k} \in \mathrm{SO}(3) \; \forall \, k \in [P],$$
$$s^2 = 1.$$

The computational cost of solving the MLE formulation in Section IV-B depends on assigning states and measurements to frames of reference such that measurements $\mathbf{B}_i$ are noisy and measurements $\mathbf{A}_i$ are noiseless (see [29] for a detailed discussion). Consider the two scenarios shown in Figure 3. By appropriately labelling states, either graph could represent eight cameras observing an arm with a target of unknown scale (e.g., consider the "eye-on-hand" and "eye-to-base" scenarios described in [4]). 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. This sparse case corresponds to the multiple eye-on-hand scenario. We recognize that some RWHEC problems may be 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 VIII and Section IX, we use the conic operator splitting method COSMO [31], which is able to exploit sparsity in $\mathbf{Q}_{red}$ and $\mathbf{Q}_{\alpha,red}$. Finally, the impact of interpreting the camera extrinsics as either $\mathbf{X}$s or $\mathbf{Y}$s is explored in [29], alongside a detailed discussion of anisotropic noise distributions beyond the scope of this work.
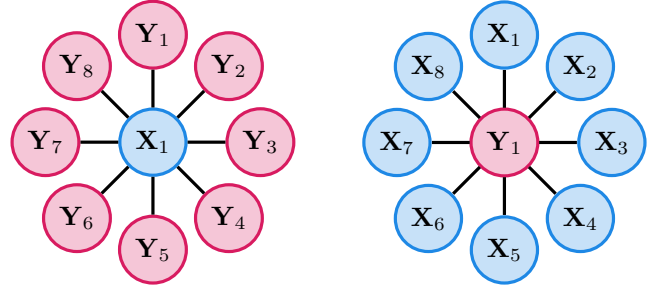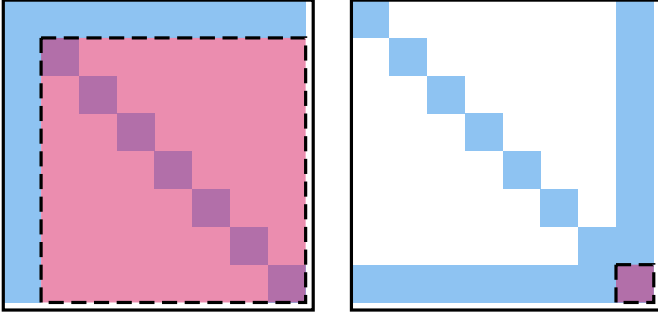

(a) A graph with multiple $\mathbf{Y}$s. (b) A graph with multiple $\mathbf{X}$s.

Fig. 3: Two alternate formulations or labelling conventions of the same generalized RWHEC scenario. Depending on which nodes ($\mathbf{X}$s or $\mathbf{Y}$s) represent sensor extrinsics, only one of these conventions lends itself to the MLE interpretation in Section IV-B. Furthermore, the convention in Figure 3b leads to a sparse reduced QCQP as shown in Figure 4b.

V. CERTIFIABLY GLOBALLY OPTIMAL EXTRINSIC CALIBRATION

In this section, we derive a convex SDP relaxation of our calibration problem. Before performing the relaxation, let us

(a) Sparsity pattern of $\mathbf{Q}_{\alpha,red}$ for a formulation with multiple $\mathbf{Y}$s.

(b) Sparsity pattern of $\mathbf{Q}_{\alpha,red}$ for a formulation with multiple $\mathbf{X}$s.

Fig. 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 cost function with the homogenization term, which only exists in the constraints in the unscaled case.

refine our SO(3) constraints to a more familiar QCQP form. To constrain $\mathbf{R} \in \mathrm{SO}(3)$, we use

$$\mathbf{R}\,\mathbf{R}^\top = s^2\mathbf{I},$$
$$\mathbf{R}^\top\mathbf{R} = s^2\mathbf{I}, \tag{61}$$
$$\mathbf{R}^{(i)} \times \mathbf{R}^{(j)} = s\mathbf{R}^{(k)}, \ \mathrm{cyclic}(1,2,3),$$

where $s^2 = 1$ is our homogenizing variable, $\mathrm{cyclic}(\cdot)$ gives the cyclic permutations of its arguments, and $\mathbf{R}^{(i)}$ denotes the ith column of $\mathbf{R}$.

Using the homogenized constraints and the theory of duality for generalized inequalities [32], the Lagrangians for Problems 3 and 4 are

$$\mathcal{L}\,(\mathbf{x},\boldsymbol{\nu}) = \nu_s + \mathbf{x}^\top\mathbf{Z}(\boldsymbol{\nu})\mathbf{x}, \tag{62}$$
$$\mathbf{Z}(\boldsymbol{\nu}) = \mathbf{Q}_{red} + \mathbf{P}_1(\boldsymbol{\nu}) + \mathbf{P}_2(\boldsymbol{\nu}), \tag{63}$$
$$\mathcal{L}_\alpha(\mathbf{x},\boldsymbol{\nu}) = \nu_s + \mathbf{x}^\top\mathbf{Z}_\alpha(\boldsymbol{\nu})\mathbf{x}, \tag{64}$$
$$\mathbf{Z}_\alpha(\boldsymbol{\nu}) = \mathbf{Q}_{\alpha,red} + \mathbf{P}_1(\boldsymbol{\nu}) + \mathbf{P}_2(\boldsymbol{\nu}), \tag{65}$$

where

$$\mathbf{P}_1(\boldsymbol{\mathcal{V}}_1,\ldots,\boldsymbol{\mathcal{V}}_{2(M+P)}) = \mathrm{Diag}(-(\boldsymbol{\mathcal{V}}_1\oplus\boldsymbol{\mathcal{V}}_2),$$
$$\ldots,-(\boldsymbol{\mathcal{V}}_{2(M+P)-1}\oplus\boldsymbol{\mathcal{V}}_{2(M+P)}), \overset{2(M+P)}{\underset{i=1}{\sum}}\mathrm{tr}(\boldsymbol{\mathcal{V}}_i)) \tag{66}$$

and

$$\mathbf{P}_2 = \begin{bmatrix} \mathbf{P}_D & \mathbf{p} \\ \mathbf{p}^\top & -\nu_s \end{bmatrix} \tag{67}$$
$$\mathbf{P}_D = \mathrm{diag}\left((\mathbf{P}_h(\boldsymbol{\nu}_1),\ldots,\mathbf{P}_h(\boldsymbol{\nu}_{(M+P)}))\right) \tag{68}$$
$$\mathbf{p} = \begin{bmatrix} -\mathbf{P}_o\boldsymbol{\nu}_1 \\ \vdots \\ -\mathbf{P}_o\boldsymbol{\nu}_{(M+P)} \end{bmatrix} \tag{69}$$
$$\boldsymbol{\mathcal{V}}_1,\ldots,\boldsymbol{\mathcal{V}}_{2(M+P)} \in \mathbb{S}^3, \tag{70}$$
$$\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_{(M+P)} \in \mathbb{R}^9. \tag{71}$$

We can subdivide $\boldsymbol{\nu}_l$, such that $\boldsymbol{\nu}_l^\top = \begin{bmatrix} \boldsymbol{\nu}_{l,ijk}^T & \boldsymbol{\nu}_{l,jki}^T & \boldsymbol{\nu}_{l,kij}^T \end{bmatrix}$ and $\boldsymbol{\nu}_{l,ijk},\boldsymbol{\nu}_{l,jki},\boldsymbol{\nu}_{l,kij} \in \mathbb{R}^3 \ \forall \ l = 1,\ldots,M+P$. Further, the matrices $\mathbf{P}_o$ and $\mathbf{P}_h$ are

$$\mathbf{P}_o = \begin{bmatrix} \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \\ \mathbf{I} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \tag{72}$$

$$\mathbf{P}_h(\boldsymbol{\nu}_l) = \begin{bmatrix} \mathbf{0}_{3\times3} & -\boldsymbol{\nu}_{l,ijk}^\wedge & \boldsymbol{\nu}_{l,kij}^\wedge \\ \boldsymbol{\nu}_{l,ijk}^\wedge & \mathbf{0}_{3\times3} & -\boldsymbol{\nu}_{l,jki}^\wedge \\ -\boldsymbol{\nu}_{l,kij}^\wedge & \boldsymbol{\nu}_{l,jki}^\wedge & \mathbf{0}_{3\times3} \end{bmatrix}. \tag{73}$$

The minimum of each Lagrangian function is only defined if $\mathbf{Z}(\boldsymbol{\nu}),\mathbf{Z}_\alpha(\boldsymbol{\nu}) \succeq 0$. As a result, the Lagrangian dual problem is

Problem 7. *Dual of RWHEC Problem*

$$\begin{aligned} \max_{\boldsymbol{\nu}} \ & \nu_s, \\ \mathrm{s.t.} \ & \mathbf{Z}(\boldsymbol{\nu}) \succeq 0, \end{aligned} \tag{74}$$

where $\boldsymbol{\nu} \in \mathbb{R}^{22(M+P)+1}$.

Problem 8. *Dual of Monocular RWHEC Problem*

$$\begin{aligned} \max_{\boldsymbol{\nu}} \ & \nu_s, \\ \mathrm{s.t.} \ & \mathbf{Z}_\alpha(\boldsymbol{\nu}) \succeq 0, \end{aligned} \tag{75}$$

where $\boldsymbol{\nu} \in \mathbb{R}^{22(M+P)+1}$.

Given a solution to these SDPs, $\boldsymbol{\nu}^\star$, the solution, $\mathbf{x}^\star$, to the primal problem is the null vector of $\mathbf{Z}(\boldsymbol{\nu}^\star)$. Further, we must divide $\mathbf{x}^\star$ by $s^\star$ to recover $\mathbf{r}^\star$. Finally, we can determine the scale, α, and translations, $\mathbf{t}$ using Equations (55) and (56). If we have estimated the scale of the system, we may correct the translation vectors as mentioned in Section IV-C.

VI. UNIQUENESS OF SOLUTIONS

In this section, we derive conditions on measurement data which ensure that Problem 3 has a unique solution in the absence of noise (i.e., whether a problem instance is *identifiable*). In addition to precisely characterizing which robot and sensor motions lead to a well-posed calibration problem with identifiable parameters, the results in this section are used in Section VII to prove that SDP relaxations of our QCQP are tight, even when noisy measurements are used. A similar analysis is conducted for the monocular hand-eye calibration problem in [33] and the standard RWHEC problem in [3], but the multi-sensor and multi-target case has not been addressed until now.

A. The Rotation-Only Case

We begin with a theorem on the uniqueness of RWHEC solutions for exact measurements and SO(3)-valued data.

Theorem 1. *Consider a connected bipartite graph $\vec{\mathcal{G}}$ with exact rotation measurements $\bar{\mathbf{R}}_{\mathbf{A}_{e,i}},\bar{\mathbf{R}}_{\mathbf{B}_{e,i}} \in \mathrm{SO}(3)$ for $i \in [N_e]$ associated with each edge $e \in \vec{\mathcal{E}}$. If there exists a subset of the vertices $\mathcal{V}^\star \subseteq \mathcal{V}$ inducing subgraph $\vec{\mathcal{G}}^\star = (\mathcal{V}^\star,\vec{\mathcal{E}}^\star)$ such that*

$$\bar{\mathbf{R}}_{\mathbf{A}_{e,i}}\mathbf{R}_{\mathbf{X}_j} = \mathbf{R}_{\mathbf{Y}_k}\bar{\mathbf{R}}_{\mathbf{B}_{e,i}} \quad \forall e = (j,k) \in \vec{\mathcal{E}}^\star, \ i \in [N_e] \tag{76}$$

has a unique solution for $\mathbf{R}_{\mathbf{X}_j}, \mathbf{R}_{\mathbf{Y}_k} \in \mathrm{SO}(3)$, *then the entire generalized RWHEC problem for rotation matrices given by Equation* (76) *over $\vec{\mathcal{E}}$ instead of $\vec{\mathcal{E}}^\star$ has a unique solution.*

Proof. Equation (76) is equivalent to

$$\bar{\mathbf{R}}_{\mathbf{A}_{e,i}} \mathbf{R}_{\mathbf{X}} \bar{\mathbf{R}}_{\mathbf{B}_{e,i}}^\top = \mathbf{R}_{\mathbf{Y}_k}, \qquad (77)$$

which can be vectorized with Equation (30) as

$$\mathbf{L}_{e,i} \mathbf{x}_j = \mathbf{y}_k, \qquad (78)$$

where $\mathbf{x}_j \triangleq \mathrm{vec}\left(\mathbf{R}_{\mathbf{X}_j}\right)$, $\mathbf{y}_k \triangleq \mathrm{vec}\left(\mathbf{R}_{\mathbf{Y}_k}\right)$ for $e = (j, k)$, and

$$\mathbf{L}_{e,i} \triangleq \bar{\mathbf{R}}_{\mathbf{B}_{e,i}} \otimes \bar{\mathbf{R}}_{\mathbf{A}_{e,i}} \in \mathbb{R}^{9 \times 9}. \qquad (79)$$

Since rotation matrices are full rank and the rank of a Kronecker product is the product of its constituent matrices' ranks [34], $\mathbf{L}_{e,i}$ is nonsingular. Since $\vec{\mathcal{G}}$ is connected, for each vertex $v \in \mathcal{V} \setminus \mathcal{V}^\star$ there exists a path $p = (e_1, \ldots, e_{|p|})$ from v to some $v^\star \in \mathcal{V}^\star$. Without loss of generality, we assume $v, v^\star \in \mathcal{V}_{\mathbf{X}}$ and observe that

$$\mathbf{x}_v = \left(\prod_{l=1}^{|p|} \mathbf{L}_{e_l, i_l}^{(-1)^l} \right) \mathbf{x}_{v^\star} \qquad (80)$$

for any choice of $i_l \in [N_{e_l}]$. Since each $\mathbf{L}_{e,i}$ is full rank, this gives us a unique expression for each $\mathbf{x}_j$ and $\mathbf{y}_k$ as desired. $\square$

Theorem 1 tells us that in the idealized noise-free case, finding a "subproblem" of an instance of generalized RWHEC with a unique solution suffices to prove that the entire problem has a unique solution. The following corollary is a direct consequence of our result and Theorem 2.3 in [3]:

Corollary 2. *If there exists a single edge $e^\star \in \vec{\mathcal{E}}$ with $i_1, i_2, i_3 \in [N_e^\star]$ such that $\bar{\mathbf{R}}_{\mathbf{A}_{e^\star, i_2}}^\top \bar{\mathbf{R}}_{\mathbf{A}_{e^\star, i_1}}$ and $\bar{\mathbf{R}}_{\mathbf{A}_{e^\star, i_3}}^\top \bar{\mathbf{R}}_{\mathbf{A}_{e^\star, i_1}}$ have distinct principal axes, then the generalized RWHEC with exact measurements has a unique solution.*

Proof. Theorem 2.3 in [3] tells us that the subproblem defined by the subgraph induced by the vertices incident to e has a unique solution, satisfying the requirements of Theorem 1. $\square$

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 rotations about two distinct axes, then the entire problem has a unique solution.

B. The Full SE(3) Case

Theorem 3. *Consider a connected bipartite graph which has 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 there exists a subset of the vertices $\mathcal{V}^\star \subseteq \mathcal{V}$ inducing subgraph $\vec{\mathcal{G}}^\star = (\mathcal{V}^\star, \vec{\mathcal{E}}^\star)$ such that*

$$\mathbf{A}_{e,i} \mathbf{X}_i = \mathbf{Y}_j \mathbf{B}_{e,i} \quad \forall e = (j, k) \in \vec{\mathcal{E}}^\star, \; i \in [N_e] \qquad (81)$$

has a unique solution, then the entire generalized RWHEC problem given by Equation (81) *over $\vec{\mathcal{E}}$ instead of $\vec{\mathcal{E}}^\star$ has a unique solution.*

Proof. Theorem 1 gives us a unique solution for the rotation component of each $\mathbf{X}_i$ and $\mathbf{Y}_j$. The translational component of Equation (81) is

$$\mathbf{t}_{\mathbf{X}_i} = \bar{\mathbf{R}}_{\mathbf{A}_{e,i}}^\top (\mathbf{R}_{\mathbf{Y}_k} \bar{\mathbf{t}}_{\mathbf{B}_{e,i}} + \mathbf{t}_{\mathbf{Y}_j} - \bar{\mathbf{t}}_{\mathbf{A}_{e,i}}), \qquad (82)$$

which relates the translation components via an affine equation. Since rotation matrices are full rank, any translation $\mathbf{t}_{\mathbf{X}_i}$ or $\mathbf{t}_{\mathbf{Y}_j}$ related to a unique translation via Equation (82) is itself uniquely determined. Therefore, the fact that $\vec{\mathcal{G}}$ is connected allows us to uniquely specify each translation through a procedure analogous to the one used to prove Theorem 1. $\square$

Once again, the results in [3] provide us with a geometrically interpretable sufficient condition that is analogous to the rotation-only case in Corollary 2:

Corollary 4. *If there exists an edge $e^\star \in \vec{\mathcal{E}}$ with $k_1, k_2, k_3 \in [N_{e^\star}]$ such that $\mathbf{R}_{\mathbf{A}_{k_2}}^\top \mathbf{R}_{\mathbf{A}_{k_1}}$ and $\mathbf{R}_{\mathbf{A}_{k_3}}^\top \mathbf{R}_{\mathbf{A}_{k_1}}$ have distinct principal axes, then the generalized RWHEC with exact measurements has a unique solution.*

Proof. Theorem 3.2 in [3] tells us that the subproblem defined by the subgraph induced by the vertices incident to $e^\star$ has a unique solution, satisfying the requirements of Theorem 3. $\square$

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 in [1]. Theorem 3 reveals that satisfying this condition for a single $\mathbf{X}_i$-$\mathbf{Y}_j$ pair is sufficient to ensure generalized RWHEC has a unique solution in the idealized case without noise. These results indicates that practitioners aiming to calibrate a multi-sensor rig can ensure all parameters are identifiable by exciting their platform about two axes when taking measurements of a single target with a single sensor, so long as the graph of measurements is connected.

VII. GLOBAL OPTIMALITY GUARANTEES

A solution to the convex SDP relaxation described in Section V provides a post-hoc certificate of global optimality in the form of 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 threshold. This is achieved by applying the following theorem:

Theorem 5 (Thm. 3.2 in [11])**.** *Consider the family of parametric QCQPs of the form*

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \quad & \mathbf{x}^\top \mathbf{F}(\boldsymbol{\theta}) \mathbf{x} \\ s.t. \quad & \mathbf{x}^\top \mathbf{G}_i \mathbf{x} + c_i = 0, \; i = 1, \ldots, m, \end{aligned} \qquad (83)$$

where $\mathbf{F} : \Theta \to \mathbb{S}^n$ is a continuous function of a parameter $\boldsymbol{\theta}$ in some parameter space $\Theta \subseteq \mathbb{R}^d$ and $\exists c_i \neq 0$. Let $\bar{\boldsymbol{\theta}}$ be such that $\mathbf{F}(\bar{\boldsymbol{\theta}}) \succeq \mathbf{0}$ has corank one and an optimal objective value of zero. 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 (83) *are tight.*

For simplicity, we apply Theorem 5 to Problem 3 with $\sigma_{e,i} = \kappa_{e,i} = 1$ for all indices $i \in [N_e]$ and each edge $e \in \vec{\mathcal{E}}$:

Theorem 6. *Let $\bar{\boldsymbol{\theta}} \triangleq \text{vec}\left(\{\mathbf{A}_{e,k}, \mathbf{B}_{e,k}\}_{e \in \bar{\mathcal{E}}, k \in [N_e]}\right)$ consist of exact measurements for which the conditions of Theorem 3 hold. Then there is a neighbourhood of $\bar{\boldsymbol{\theta}}$ in which the SDP relaxation of Problem 3 is tight.*

Proof. We begin by noting that Problem 3 takes the form of Equation (83) with $\mathbf{F}(\boldsymbol{\theta}) = \mathbf{Q}(\boldsymbol{\theta})$, where we explicitly note the dependence of Equation (48) on our new parameter $\boldsymbol{\theta}$. The cost function is PSD because it is the sum of squared norms, and the optimal objective value is zero because it comes from exact measurements. The ACQ holds because the quadratic variety in Equation (61) used to describe $\text{SO}(3)^{2 \sum_{e \in \bar{\mathcal{E}}} N_e}$ is smooth and the ideal is radical [11]. The inclusion of the homogenizing variable $s^2 = 1$ ensures that some $c_i \neq 0$. Finally, Theorem 3 tells us that the linear system of equations whose summed squared residual norms form the cost function encoded in $\mathbf{Q}(\boldsymbol{\theta}) \succeq 0$ has a unique solution. Since a sum of squared norms is zero if and only if its arguments are zero, this implies that $\mathbf{Q}(\bar{\boldsymbol{\theta}})$ is corank zero and that the unique minimizer is in its nullspace. Therefore, Theorem 5 applies and there exists a neighbourhood around the noiseless measurements $\bar{\boldsymbol{\theta}}$ within which the SDP relaxation is tight. $\square$

Theorem 6 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. This global optimality result is the first of its kind for a multi-sensor calibration problem.

VIII. 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 time t_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 t_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 computed

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 experiment, we consider translation noise with standard deviations of $\sigma = 1$ cm and $\sigma = 5$ cm, and rotation noise concentrations of $\kappa = 125$ and $\kappa = 12$. 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 [3] and [4], the certifiable method in [6], the probabilistic method in [8], and a local on-manifold probabilistic solver 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 the local probabilistic solver. When a scenario has more than one $\mathbf{X}$ to solve for, we compare our methods to [4], [6], and the local solver.

For the solvers that require initialization, we provide random initial values, the solution from [4], or parameters close to the ground truth. For each experiment, our method and the method in [6] are randomly initialized. We attempted randomly initializing our local solver, but this scheme resulted in the local solver returning local minima. Consequently, we also initialize the solver in [8] and our local solver using the solution from the method in [4]. However, the method in [4] cannot solve monocular RWHEC problems. In the monocular RWHEC studies, we initialize the local solver with calibration parameters that are within 10 cm and 10° of the ground truth values. Additionally, we set the initial scale to 1. Our globally optimal solver was implemented in Julia with the JuMP modelling language [35]. For all experiments in this section and Section IX, the COSMO SDP solver was used with default parameters except for residual norm convergence tolerances of $\epsilon_{\text{abs}} = \epsilon_{\text{rel}} = 5 \times 10^{-11}$, initial step size $\rho = 10^{-4}$, and a maximum of one million iterations to ensure convergence [31]. Finally, our local on-manifold solver was implemented in Ceres [36], 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.

A. 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 $\mathbf{Y} = \mathbf{T}_{bw}$, where $\underrightarrow{\mathcal{F}}_b, \underrightarrow{\mathcal{F}}_w, \underrightarrow{\mathcal{F}}_h, \underrightarrow{\mathcal{F}}_c$, 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
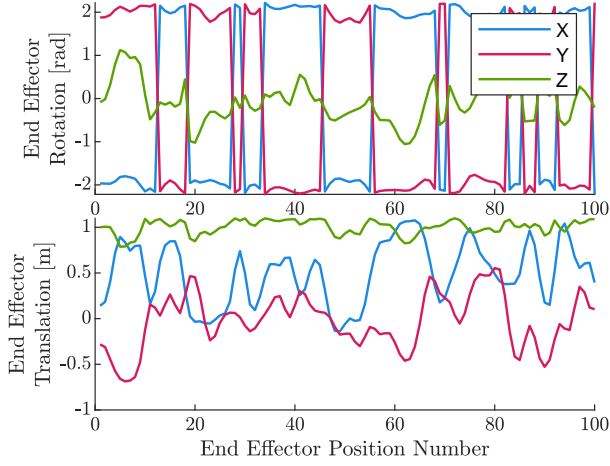
Fig. 5: Simulated end-effector pose relative to the robot base for the robot poses collected on a sphere in support of RWHEC experiments with known scale. Data collection is simulated as if the robot stops at each pose to take measurements.

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 (see Figure 5 for a representative example trajectory). Trajectories generated in this manner for the RWHEC study were empirically observed to not have a unique solution for the monocular RWHEC problem: to create *identifiable* monocular RWHEC problem instances, the camera travels on a sphere of radius 0.25 m after the 50th manipulator pose (see Figure 6 for the exact trajectory).

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 II. Generally, the least accurate methods are the two-stage closed-form solvers from [3] and [4], while our method and the local solver are the most accurate. Often, our method and the local method return similar solutions, which outperform the other methods by reducing the error by up to 50 mm and 3 degrees. Additionally, these two methods are more susceptible to translation measurement noise than rotation measurement noise. Finally, all algorithms have roughly the same accuracy when $\kappa = 125$ and $\sigma = 5$ cm. In
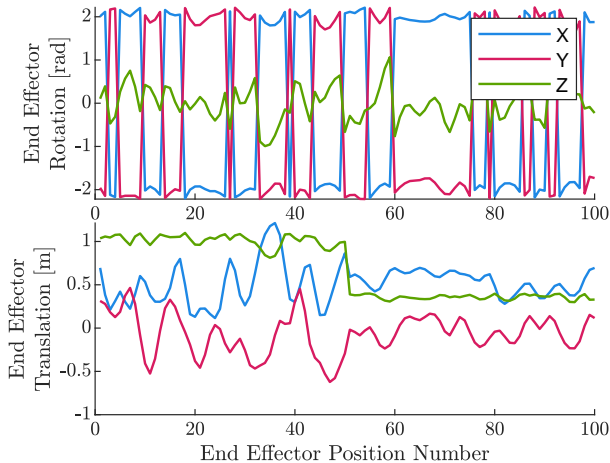

Fig. 6: Plot of end effector pose relative to the base for the robot poses on sphere experiment with unknown scale. We treat the data as if the robot stops at each position.

this case, the rotation measurement noise is low, so the two-stage closed-form solvers are more likely to return accurate rotation estimates.

Table III 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 the local method 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., [37] for more insight into the effect of sensor trajectory on calibration accuracy).

B. 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 7 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 leads to a problem instance with a unique solution (see Figure 8 for an example trajectory).

The results of standard and monocular RWHEC studies are shown in Tables IV and V, 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 V also displays the mean and standard deviation of the estimated scale error (α_{err}). As with the previous experiments, our method and the local method produced the most accurate parameter estimates. The method of Wang et al. [4] 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. [6]. Recalling the double cover from quaternions to SE(3) in Equation (2), we hypothesize that the use of dual quaternions by the method in [6] leads to its deteriorated performance in the multi-$\mathbf{X}$ case, as each measurement (i.e., $\mathbf{A}$ or $\mathbf{B}$) must be assigned one of two equivalent unit DQ representations of SE(3). Finally, as in our previous monocular RWHEC study, our algorithm and the local method initialized to within 10 cm and 10° of the true solution achieved similar accuracy.

TABLE II: Calibration results for the experiments in Section VIII-A 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 the local solvers are initialized with the method in [4]. Initializing the local method 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
	Local (Wang)	11.1 ± 5.82	0.78 ± 0.37	3.69 ± 2.12	0.63 ± 0.38
	Ours	10.9 ± 5.6	0.77 ± 0.36	3.71 ± 2.13	0.62 ± 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	1.42 ± 0.60	19.9 ± 9.2	1.36 ± 0.59
	Local (Wang)	28.5 ± 13.4	1.45 ± 0.62	18.5 ± 8.8	1.39 ± 0.62
	Ours	28.4 ± 13.0	1.42 ± 0.63	18.5 ± 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
	Local (Wang)	15.2 ± 10.0	1.84 ± 0.84	3.4 ± 1.8	0.88 ± 0.59
	Ours	15.1 ± 10.0	1.81 ± 0.83	3.4 ± 1.8	0.87 ± 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
	Local (Wang)	48.3 ± 27.3	3.18 ± 1.67	18.7 ± 9.04	2.72 ± 1.64
	Ours	47.7 ± 27.0	3.12 ± 1.67	18.8 ± 9.02	2.68 ± 1.65

TABLE III: Calibration results for the experiments in Section VIII-A 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. The local method is initialized with parameters $10°$ and 10 cm from the ground truth values and a scale of 1. As in Table II, the local method 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	Local (close)	4.61 ± 2.39	0.51 ± 0.21	3.82 ± 1.93	0.22 ± 0.12	$1.78\text{e-}3 \pm 1.54\text{e-}3$
	Ours	4.71 ± 2.43	0.513 ± 0.218	3.82 ± 1.91	0.250 ± 0.113	$1.76\text{e-}3 \pm 1.54\text{e-}3$
$\kappa = 125$, $\sigma = 5$cm	Local (close)	23.0 ± 11.2	0.94 ± 0.44	19.1 ± 11.4	0.85 ± 0.40	$8.45\text{e-}3 \pm 7.03\text{e-}3$
	Ours	23.8 ± 11.3	1.11 ± 0.53	19.0 ± 11.4	1.06 ± 0.52	$8.41\text{e-}3 \pm 7.02\text{e-}3$
$\kappa = 12$, $\sigma = 1$cm	Local (close)	4.52 ± 1.97	1.57 ± 0.59	3.78 ± 1.89	0.20 ± 0.10	$1.88\text{e-}3 \pm 1.37\text{e-}3$
	Ours	4.50 ± 1.99	1.54 ± 0.57	3.77 ± 1.88	0.206 ± 0.098	$1.87\text{e-}3 \pm 1.37\text{e-}3$
$\kappa = 12$, $\sigma = 5$cm	Local (close)	23.2 ± 14.1	1.90 ± 0.81	20.0 ± 11.0	1.07 ± 0.54	$9.24\text{e-}3 \pm 7.69\text{e-}3$
	Ours	24.4 ± 14.3	1.99 ± 0.81	19.6 ± 11.0	1.25 ± 0.58	$9.16\text{e-}3 \pm 7.69\text{e-}3$

IX. REAL-WORLD EXPERIMENT

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

A. 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 $\mathbf{A}_{ijk}$ and fiducial markers in the environment with unknown poses $\mathbf{Y}_j$ to form the RWHEC problem shown in Figure 9. Figures 10 and 11 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 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

TABLE IV: Calibration results for the multi-sensor experiments in Section VIII-B 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. The local solver is initialized with the results from Wang. Note we have no a priori estimates of $\mathbf{t_{X}}_i$. Once again, we have not included results from experiments initializing the local solver 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
	Local (Wang)	0.994 ± 0.416	0.454 ± 0.190	0.572 ± 0.296	0.022 ± 0.010
	Ours	0.992 ± 0.416	0.454 ± 0.189	0.570 ± 0.294	0.021 ± 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
	Local (Wang)	4.57 ± 1.96	0.460 ± 0.191	2.74 ± 1.34	0.120 ± 0.056
	Ours	4.56 ± 1.96	0.459 ± 0.191	2.69 ± 1.32	0.111 ± 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
	Local (Wang)	0.933 ± 0.370	1.52 ± 0.64	0.537 ± 0.262	0.021 ± 0.010
	Ours	0.933 ± 0.370	1.50 ± 0.62	0.537 ± 0.262	0.021 ± 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
	Local (Wang)	4.57 ± 1.97	1.51 ± 0.682	2.93 ± 1.34	0.104 ± 0.047
	Ours	4.57 ± 1.97	1.48 ± 0.67	2.92 ± 1.32	0.103 ± 0.046

TABLE V: Calibration results for the multi-sensor experiments in Section VIII-B 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. The local method 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 the local solver 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	Local (close)	1.05 ± 0.45	0.469 ± 0.185	0.631 ± 0.249	0.030 ± 0.013	2.71e-4 $\pm$ 1.98e-4
	Ours	1.05 ± 0.45	0.468 ± 0.185	0.597 ± 0.256	0.026 ± 0.012	2.71e-4 $\pm$ 1.98e-4
$\kappa = 125$, $\sigma = 5$cm	Local (close)	5.25 ± 2.20	0.457 ± 0.195	3.15 ± 1.26	0.179 ± 0.079	1.37e-3 $\pm$ 1.03e-3
	Ours	5.21 ± 2.18	0.454 ± 0.194	3.03 ± 1.24	0.159 ± 0.071	1.37e-3 $\pm$ 1.03e-3
$\kappa = 12$, $\sigma = 1$cm	Local (close)	1.02 ± 0.44	1.50 ± 0.62	0.530 ± 0.284	0.021 ± 0.010	2.58e-4 $\pm$ 2.02e-4
	Ours	1.02 ± 0.44	1.46 ± 0.61	0.530 ± 0.284	0.021 ± 0.010	2.57e-4 $\pm$ 2.02e-4
$\kappa = 12$, $\sigma = 5$cm	Local (close)	5.20 ± 2.27	1.48 ± 0.64	3.13 ± 1.56	0.156 ± 0.075	1.18e-3 $\pm$ 8.63e-4
	Ours	5.18 ± 2.26	1.44 ± 0.626	3.02 ± 1.51	0.144 ± 0.069	1.18e-3 $\pm$ 8.62e-4

estimated camera calibration parameters, we leverage the IMU mounted on the system to define the ground truth extrinsic transformations from the cameras to a sensor rig reference frame $\mathcal{F}_r$ fixed to the IMU using the Kalibr toolbox [38]. A total of sixteen AprilTags [39] are mounted in our experimental environment for use as fiducial markers. To evaluate the accuracy of estimated AprilTag poses $\mathbf{X}_k$, we place OptiTrack markers on a single AprilTag to establish ground truth measurements.

Figure 12 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 perpendicularly 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 13 is a grid where row j corresponds to camera j, and column i corresponds to AprilTag i. If the grid square at coordinate i, j 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 a camera-target pair did not contain sufficient excitation, while a white square indicates that a camera did not observe a particular target. Figure 13 indicates that each camera observed at least one target, and that the overall generalized RWHEC problem is described by a connected bipartite graph.

Our data preprocessing pipeline for this experiment consists of four steps. First, we rectify the images using the camera intrinsic parameters estimated with Kalibr (see [38]). Second, we measure the camera-to-AprilTag transformations using an AprilTag detector. Third, we synchronize the camera and OptiTrack measurements by linearly interpolating the OptiTrack system measurements to the camera measurement timestamps.
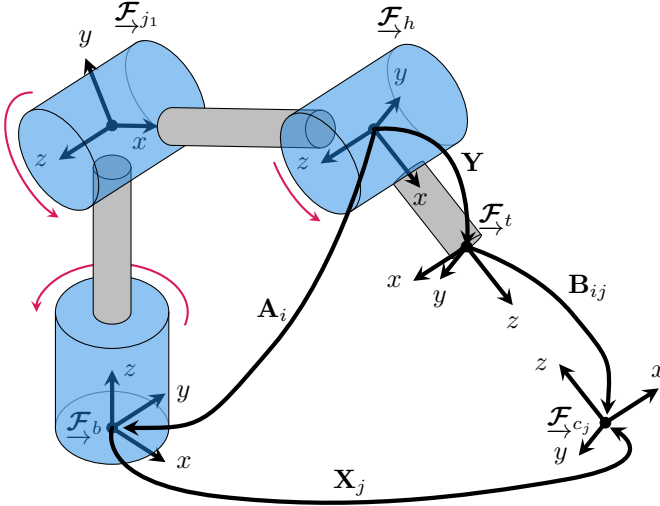
Fig. 7: A diagram of a robot arm with a hand-mounted target that is observed by a stationary target. In this diagram, the base, joint 1, hand, camera j, and target reference frames are labelled $\underrightarrow{\mathcal{F}}_b$, $\underrightarrow{\mathcal{F}}_{j_1}$, $\underrightarrow{\mathcal{F}}_h$, $\underrightarrow{\mathcal{F}}_{c_j}$, and $\underrightarrow{\mathcal{F}}_t$, respectively. The red arrows indicate the axis of joint rotation. At time t_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, $\mathbf{B}_{ij}$.

Finally, we use the following RANSAC [40] scheme to reject gross outliers: if each $\mathbf{Y}_k$ is approximately known, then the RWHEC geometric constraint at time t_i becomes

$$\mathbf{X}_j = \mathbf{A}_{ijk}^{-1}\mathbf{Y}_k\mathbf{B}_{ijk}. \tag{84}$$

Consequently, we can use RANSAC on the model in Equation (84) to determine the $\mathbf{A}_{ijk}$-$\mathbf{B}_{ijk}$ pairs that result in a consistent $\mathbf{X}_j$ transform. For this RANSAC procedure, our minimum inlier set size is one third of the number of $\mathbf{A}_{ijk}$-$\mathbf{B}_{ijk}$ pairs for camera j and target k. Further, data from camera-target pairs with only one $\mathbf{A}_{ijk}$-$\mathbf{B}_{ijk}$ pair is rejected because it cannot be validated using this scheme. An $\mathbf{A}_{ijk}$-$\mathbf{B}_{ijk}$ pair is considered an inlier if it is within 0.6 m and 60° of the estimated pose $\mathbf{X}_j$. The inlier $\mathbf{A}_{ijk}$-$\mathbf{B}_{ijk}$ pairs are then
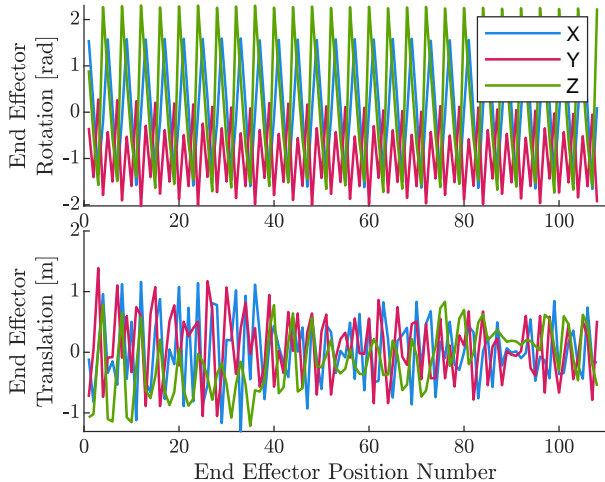

Fig. 8: End effector motion for the standard and monocular RWHEC simulation experiment with many cameras observing a hand-mounted target described in Section VIII-B.

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.

B. Experimental Results

Using the preprocessed data, we obtain the calibration results shown in Table VI. 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 [4]. 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 method in [4] returned the least accurate rotation estimates since the estimated rotations were projected onto SO(3). As in our simulation studies, our method and the local method compute parameters with similar accuracy. Interestingly, 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 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 estimates, and future work can investigate a maximum a posteriori (MAP) estimation scheme that leverages prior estimates of parameters in a principled fashion.

C. 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 VI. COSMO solved the synthetic problems with the sparsity pattern in Figure 3b substantially faster, taking at most seven seconds. Unsurprisingly, the Ceres implementation of our local solver took at most one second.

X. 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
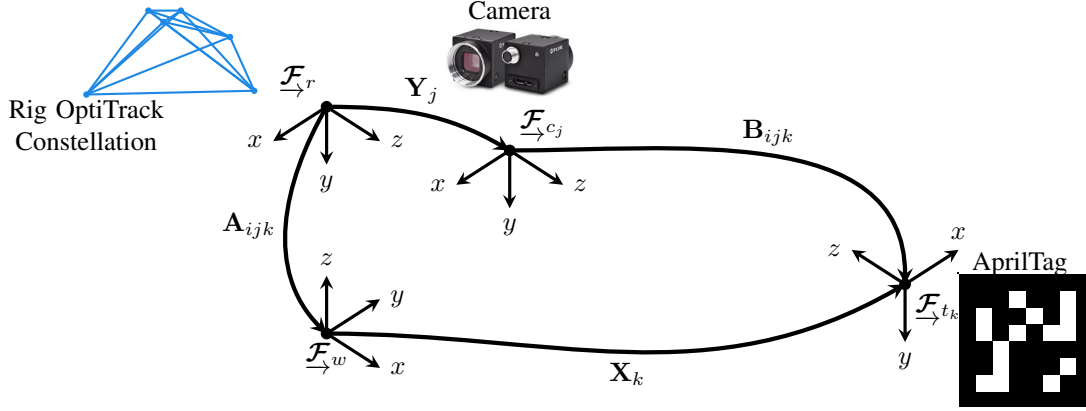
Fig. 9: A diagram of the measurements for camera i and target j at time t_i. The reference frames for the camera, target, OptiTrack world, and rig reference frames are $\underrightarrow{\mathcal{F}}_{c_j}$, $\underrightarrow{\mathcal{F}}_{t_k}$, $\underrightarrow{\mathcal{F}}_w$, and $\underrightarrow{\mathcal{F}}_r$. The OptiTrack rig constellation enables estimation of the transform $\mathbf{A}_{ijk}$. The monocular camera observing the the AprilTag enables estimation of the transform $\mathbf{B}_{ijk}$.

TABLE VI: 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 are given in each cells corresponding to the $\mathbf{t}_{c_0}^{c_i c_0}$ and $\mathbf{R}_{c_0 c_i}$ errors. The error in the estimated transform between AprilTag 20 the OptiTrack reference frame is also provided. The local method is initialized with the solution from [4].

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 et al. [4]	3.46 ± 1.69	1.38 ± 0.81	11.7	5.30
Local (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
Local with Scale (Wang)	3.28 ± 2.42	0.86 ± 0.47	7.92	4.50
Ours with Scale	2.87 ± 1.94	0.99 ± 0.52	7.94	4.72

monocular cameras and arbitrary multi-sensor and target configurations. Additionally, we have presented a novel on-manifold local solver, and used it to demonstrate the existence of numerous local minima in multiple instances of the generalized RWHEC problem, highlighting the importance of global optimality. We have also characterized the subset of
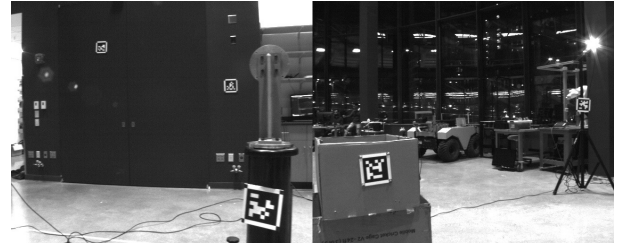

Fig. 11: Images from the real-world experiment. These images are from camera 0 and show a subset of the AprilTags in the environment. The bottom left AprilTag in the image on the right has OptiTrack markers, so we can determine the ground truth pose of the AprilTag frame relative to the OptiTrack world frame.
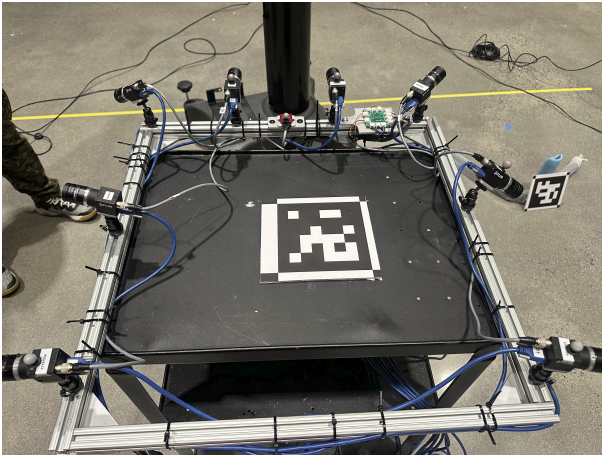

Fig. 10: Image of the real-world data collection rig. The data collection rig consists of eight hardware synchronized cameras facing a variety of different directions. Further, the data collection rig includes an IMU and opti track markers. The OptiTrack markers enable us to estimate the rig pose relative to the OptiTrack reference frame.

generalized 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 a key step towards truly "power-on-and-go" sensor calibration algorithms for multi-sensor robotic systems. This will require a RWHEC solver which does not depend on specialized calibration targets and is therefore able to handle outliers caused by errors in data association. One potential solution to this problem is the modi-

fication of RWHEC to use a robust cost function like truncated least squares [13]. Additionally, our current MLE formulation assumes that translation and rotation noise is isotropic, which limits the fidelity of sensor models employed. Future work can remedy this shortcoming with insights from the study of anisotropic SLAM in [41]. Finally, truly autonomous sensor calibration will require an active perception strategy which can design trajectories that are information-theoretically optimal [42] or seek out measurements that render parameters identifiable [43].

APPENDIX A
ON-MANIFOLD RWHEC SOLVER

In this appendix, we derive a local solver for the RWHEC and unscaled RWHEC problems. Given a bipartite graph with nodes $\{\mathbf{X}_1, \ldots, \mathbf{X}_{N_i}, \mathbf{Y}_1, \ldots, \mathbf{Y}_{N_j}\}$, let $\mathcal{D}_{i,j}$ be the data that forms the connection between $\mathbf{X}_i$ and $\mathbf{Y}_j$, specifically $\mathcal{D}_{i,j} = \{(\mathbf{A}_{i,j,k}, \mathbf{B}_{i,j,k}) \ \forall \ k = 1, \ldots, N_{i,j}\}$. The noisy measurement models for a given $(\mathbf{A}_{i,j,k}, \mathbf{B}_{i,j,k})$ pair are

$$\begin{aligned}
\mathbf{R}_{\mathbf{B}_{i,j,k}} &= \mathbf{R}_{\mathbf{Y}_j}^\top \mathbf{R}_{\mathbf{A}_{i,j,k}} \mathbf{R}_{\mathbf{X}_i} \exp\left(\mathbf{n}_R^\wedge\right), \\
\mathbf{n}_R &\sim \mathcal{N}\left(0, \sigma_R^2 \mathbf{I}\right), \\
\mathbf{t}_{\mathbf{B}_{i,j,k}} &= \mathbf{R}_{\mathbf{Y}_j}^\top \left(\mathbf{R}_{\mathbf{A}_{i,j,k}} \mathbf{t}_{\mathbf{X}_i} + \mathbf{t}_{\mathbf{A}_{i,j,k}} - \mathbf{t}_{\mathbf{Y}_j}\right) + \mathbf{n}_t, \\
\mathbf{n}_R &\sim \mathcal{N}\left(0, \sigma_t^2 \mathbf{I}\right).
\end{aligned} \tag{85}$$

This model leverages the right-perturbation noise framework of [44] to ensure that the model is an on-manifold analogue of the MLE model in Section IV-B. The associated error distributions are

$$\begin{aligned}
\mathbf{e}_{R_{i,j,k}} &= \log\left(\mathbf{R}_{\mathbf{X}_i}^\top \mathbf{R}_{\mathbf{A}_{i,j,k}}^\top \mathbf{R}_{\mathbf{Y}_j} \mathbf{R}_{\mathbf{B}_{i,j,k}}\right)^\vee \\
&\sim \mathcal{N}\left(0, \sigma_R^2 \mathbf{I}\right), \\
\mathbf{e}_{t_{i,j,k}} &= \mathbf{R}_{\mathbf{A}_{i,j,k}} \mathbf{t}_{\mathbf{X}_i} + \mathbf{t}_{\mathbf{A}_{i,j,k}} - \mathbf{R}_{\mathbf{Y}_j} \mathbf{t}_{\mathbf{B}_{i,j,k}} - \mathbf{t}_{\mathbf{Y}_j} \\
&\sim \mathcal{N}\left(0, \sigma_t^2 \mathbf{I}\right).
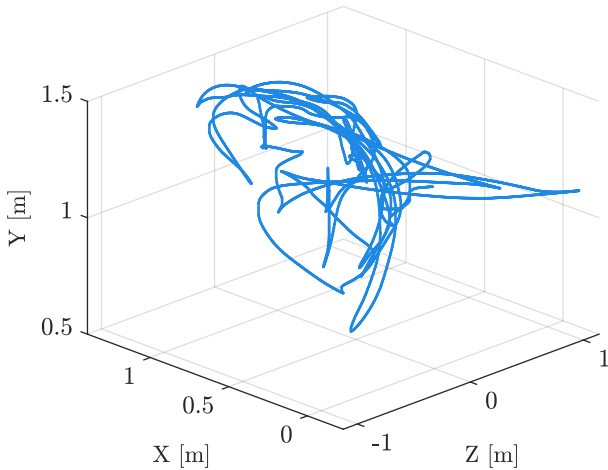\end{aligned} \tag{86}$$

Fig. 12: Trajectory of the rig in the real-world experiment. Initially, the platform rotates about the y-axis follows a planar motion in the xz-plane. Following the planar motion, the system follows an unconstrained trajectory, which allows for rotation about the x and z axes.
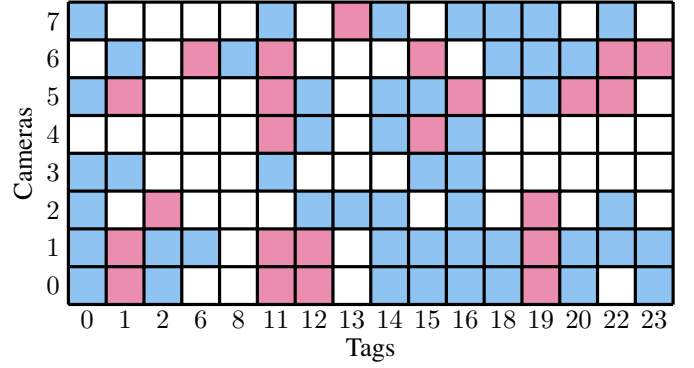

Fig. 13: A grid describing the observability of each connection in the bipartite graph generated by our real-world experiment. Blue squares indicate that measurements between tag i and camera j are sufficient for an identifiable RWHEC problem. Red squares indicate that there is a connection between tag i and camera j, 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 i by camera j.

The associated error Jacobians are

$$\begin{aligned}
\frac{\partial \mathbf{e}_{R_{i,j,k}}}{\partial \psi_{\mathbf{X}_i}} &= -\mathbf{R}_{\mathbf{B}_{i,j,k}}^\top \mathbf{R}_{\mathbf{Y}_j}^\top \mathbf{R}_{\mathbf{A}_{i,j,k}}, \\
\frac{\partial \mathbf{e}_{R_{i,j,k}}}{\partial \psi_{\mathbf{Y}_j}} &= \mathbf{R}_{\mathbf{B}_{i,j,k}}^\top \mathbf{R}_{\mathbf{Y}_j}^\top, \\
\frac{\partial \mathbf{e}_{t_{i,j,k}}}{\partial \mathbf{t}_{\mathbf{X}_i}} &= \mathbf{R}_{\mathbf{A}_{i,j,k}}, \\
\frac{\partial \mathbf{e}_{t_{i,j,k}}}{\partial \psi_{\mathbf{Y}_j}} &= \left(\mathbf{R}_{\mathbf{Y}_j} \mathbf{t}_{\mathbf{B}_{i,j,k}}\right)^\wedge, \\
\frac{\partial \mathbf{e}_{t_{i,j,k}}}{\partial \mathbf{t}_{\mathbf{Y}_j}} &= -\mathbf{I}.
\end{aligned} \tag{87}$$

If the measurement $\mathbf{B}_{i,j,k}$ is from an unscaled pose sensor, then the translation measurement model and error distribution are

$$\begin{aligned}
\mathbf{t}_{\mathbf{B}_{i,j,k}} &= \mathbf{R}_{\mathbf{Y}_j}^\top \left(\mathbf{R}_{\mathbf{A}_{i,j,k}} \mathbf{t}_{\mathbf{X}_i,\alpha} + \alpha \mathbf{t}_{\mathbf{A}_{i,j,k}} - \mathbf{t}_{\mathbf{Y}_j,\alpha}\right) + \mathbf{n}_t, \\
\mathbf{n}_R &\sim \mathcal{N}\left(0, \sigma_t^2 \mathbf{I}\right), \\
\mathbf{e}_{t_{i,j,k,\alpha}} &= \mathbf{R}_{\mathbf{A}_{i,j,k}} \mathbf{t}_{\mathbf{X}_i,\alpha} + \alpha \mathbf{t}_{\mathbf{A}_{i,j,k}} - \mathbf{R}_{\mathbf{Y}_j} \mathbf{t}_{\mathbf{B}_{i,j,k}} - \mathbf{t}_{\mathbf{Y}_j,\alpha} \\
&\sim \mathcal{N}\left(0, \sigma_t^2 \mathbf{I}\right).
\end{aligned} \tag{88}$$

The associated error Jacobians are

$$\begin{aligned}
\frac{\partial \mathbf{e}_{t_{i,j,k,\alpha}}}{\partial \mathbf{t}_{\mathbf{X}_i,\alpha}} &= \mathbf{R}_{\mathbf{A}_{i,j,k}}, \\
\frac{\partial \mathbf{e}_{t_{i,j,k,\alpha}}}{\partial \psi_{\mathbf{Y}_j}} &= \left(\mathbf{R}_{\mathbf{Y}_j} \mathbf{t}_{\mathbf{B}_{i,j,k}}\right)^\wedge, \\
\frac{\partial \mathbf{e}_{t_{i,j,k,\alpha}}}{\partial \mathbf{t}_{\mathbf{Y}_j,\alpha}} &= -\mathbf{I}, \\
\frac{\partial \mathbf{e}_{t_{i,j,k,\alpha}}}{\partial \alpha} &= \mathbf{t}_{\mathbf{A}_{i,j,k}}.
\end{aligned} \tag{89}$$

Consequently, the optimization problem is

Problem 9. *Local RWHEC Optimization Problem*

$$\min_{\mathbf{X}_1,\ldots,\mathbf{X}_{N_i},\mathbf{Y}_1,\ldots,\mathbf{Y}_{N_j}} \sum_{i,j,k} \frac{1}{\sigma_R^2}\mathbf{e}_{R_{i,j,k}}^\top \mathbf{e}_{R_{i,j,k}} + \frac{1}{\sigma_t^2}\mathbf{e}_{t_{i,j,k}}^\top \mathbf{e}_{t_{i,j,k}}. \tag{90}$$

The optimization problem for the unscaled case is

Problem 10. *Local Unscaled RWHEC Optimization Problem*

$$\min_{\mathbf{X}_1,\ldots,\mathbf{X}_{N_i},\mathbf{Y}_1,\ldots,\mathbf{Y}_{N_j},\alpha} \sum_{i,j,k} \frac{1}{\sigma_R^2}\mathbf{e}_{R_{i,j,k}}^\top \mathbf{e}_{R_{i,j,k}} + \frac{1}{\sigma_t^2}\mathbf{e}_{t_{i,j,k,\alpha}}^\top \mathbf{e}_{t_{i,j,k,\alpha}}. \tag{91}$$

We solve these two optimization problems using the Ceres solver.

REFERENCES

[1] H. Zhuang, Z. Roth, and R. Sudhakar, "Simultaneous robot/world and tool/flange calibration by solving homogeneous transformation equations of the form AX=YB," *IEEE Transactions on Robotics and Automation*, vol. 10, no. 4, pp. 549–554, Aug./1994.

[2] A. G. Li, L. Wang, and D. Wu, "Simultaneous robot-world and hand-eye calibration using dual-quaternions and kronecker product," *International Journal of Physical Sciences*, vol. 5, pp. 1530–1536, 2010.

[3] M. Shah, "Solving the Robot-World/Hand-Eye Calibration Problem Using the Kronecker Product," *Journal of Mechanisms and Robotics*, vol. 5, no. 3, p. 031007, Aug. 2013.

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

[5] A. Tabb and K. M. A. Yousef, "Solving the robot-world hand-eye(s) calibration problem with iterative methods," *Machine Vision and Applications*, vol. 28, pp. 569–590, 2017.

[6] M. Horn, T. Wodtko, M. Buchholz, and K. Dietmayer, "Extrinsic Infrastructure Calibration Using the Hand-Eye Robot-World Formulation," May 2023.

[7] D. Evangelista, E. Olivastri, D. Allegro, E. Menegatti, and A. Pretto, "A graph-based optimization framework for hand-eye calibration for multi-camera setups," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 11 474–11 480.

[8] F. Dornaika and R. Horaud, "Simultaneous robot-world and hand-eye calibration," *IEEE transactions on Robotics and Automation*, vol. 14, no. 4, pp. 617–622, 1998.

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

[10] D. M. Rosen, K. J. Doherty, A. Terán Espinoza, and J. J. Leonard, "Advances in Inference and Representation for Simultaneous Localization and Mapping," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 4, no. 1, pp. 215–242, May 2021.

[11] D. Cifuentes, S. Agarwal, P. A. Parrilo, and R. R. Thomas, "On the local stability of semidefinite relaxations," *Mathematical Programming*, vol. 193, no. 2, pp. 629–663, Jun. 2022.

[12] D. M. Rosen, L. Carlone, A. S. Bandeira, and J. J. Leonard, "SE-Sync: A certifiably correct algorithm for synchronization over the special Euclidean group," *The International Journal of Robotics Research*, vol. 38, no. 2-3, pp. 95–125, Mar. 2019.

[13] H. Yang, J. Shi, and L. Carlone, "TEASER: Fast and Certifiable Point Cloud Registration," *IEEE Transactions on Robotics*, vol. 37, no. 2, pp. 314–333, Apr. 2021.

[14] C. Holmes and T. D. Barfoot, "An Efficient Global Optimality Certificate for Landmark-Based SLAM," *IEEE Robotics and Automation Letters*, vol. 8, no. 3, pp. 1539–1546, Mar. 2023.

[15] T. Fan, H. Wang, M. Rubenstein, and T. Murphey, "CPL-SLAM: Efficient and Certifiably Correct Planar Graph-Based SLAM Using the Complex Number Representation," *IEEE Transactions on Robotics*, vol. 36, no. 6, pp. 1719–1737, Dec. 2020.

[16] Y. Tian, K. Khosoussi, D. M. Rosen, and J. P. How, "Distributed Certifiably Correct Pose-Graph Optimization," *IEEE Transactions on Robotics*, vol. 37, no. 6, pp. 2137–2156, Dec. 2021.

[17] A. Papalia, J. Morales, K. J. Doherty, D. M. Rosen, and J. J. Leonard, "SCORE: A Second-Order Conic Initialization for Range-Aided SLAM," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. London, United Kingdom: IEEE, May 2023, pp. 10 637–10 644.

[18] F. Dumbgen, C. Holmes, and T. D. Barfoot, "Safe and Smooth: Certified Continuous-Time Range-Only Localization," *IEEE Robotics and Automation Letters*, vol. 8, no. 2, pp. 1117–1124, Feb. 2023.

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

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

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

[22] M. Garcia-Salguero, J. Briales, and J. Gonzalez-Jimenez, "Certifiable relative pose estimation," *Image and Vision Computing*, vol. 109, p. 104142, May 2021.

[23] J. Zhao, "An Efficient Solution to Non-Minimal Case Essential Matrix Estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–1, 2020.

[24] F. Dümbgen, C. Holmes, B. Agro, and T. D. Barfoot, "Toward Globally Optimal State Estimation Using Automatically Tightened Semidefinite Relaxations," Aug. 2023.

[25] D. M. Rosen, "Scalable Low-Rank Semidefinite Programming for Certifiably Correct Machine Perception," in *Algorithmic Foundations of Robotics XIV*, S. M. LaValle, M. Lin, T. Ojala, D. Shell, and J. Yu, Eds. Cham: Springer International Publishing, 2021, vol. 17, pp. 551–566.

[26] H. Yang and L. Carlone, "One ring to rule them all: Certifiably robust geometric perception with outliers," *Advances in neural information processing systems*, vol. 33, pp. 18 846–18 859, 2020.

[27] J. Wu, Y. Zheng, Z. Gao, Y. Jiang, X. Hu, Y. Zhu, J. Jiao, and M. Liu, "Quadratic Pose Estimation Problems: Globally Optimal Solutions, Solvability/Observability Analysis, and Uncertainty Description," *IEEE Transactions on Robotics*, vol. 38, no. 5, pp. 3314–3335, Oct. 2022.

[28] B. Xue, Y. Zhu, T. Liu, J. Wu, J. Jiao, Y. Jiang, C. Zhang, X. Jiang, and Z. He, "S QPEP: Global Optimal Solutions to Scaled Quadratic Pose Estimation Problems," *IEEE Transactions on Instrumentation and Measurement*, pp. 1–1, 2025.

[29] J. Ha, "Probabilistic Framework for Hand–Eye and Robot–World Calibration $AX = YB$," *IEEE Transactions on Robotics*, vol. 39, no. 2, pp. 1196–1211, Apr. 2023.

[30] H. V. Henderson and S. R. Searle, "The vec-permutation matrix, the vec operator and Kronecker products: A review," *Linear and Multilinear Algebra*, vol. 9, no. 4, pp. 271–288, Jan. 1981.

[31] M. Garstka, M. Cannon, and P. Goulart, "COSMO: A Conic Operator Splitting Method for Convex Conic Problems," *Journal of Optimization Theory and Applications*, vol. 190, no. 3, pp. 779–810, Sep. 2021.

[32] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK ; New York: Cambridge University Press, 2004.

[33] N. Andreff, R. Horaud, and B. Espiau, "Robot Hand-Eye Calibration Using Structure-from-Motion," *The International Journal of Robotics Research*, vol. 20, no. 3, pp. 228–248, Mar. 2001.

[34] K. Schacke, "On the kronecker product," *Master's Thesis, University of Waterloo*, 2004.

[35] I. Dunning, J. Huchette, and M. Lubin, "JuMP: A Modeling Language for Mathematical Optimization," *SIAM Review*, vol. 59, no. 2, pp. 295–320, Jan. 2017.

[36] S. Agarwal, K. Mierle, and T. C. S. Team, "Ceres Solver," 10 2023. [Online]. Available: https://github.com/ceres-solver/ceres-solver

[37] C. Grebe, "A study of observability-aware trajectory optimization," Master's thesis, University of Toronto, Toronto, Ontario, Canada, January 2022.

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

[39] E. Olson, "AprilTag: A robust and flexible visual fiducial system," in *2011 IEEE International Conference on Robotics and Automation*. Shanghai, China: IEEE, May 2011, pp. 3400–3407.

[40] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[41] C. Holmes, F. Dümbgen, and T. Barfoot, "On Semidefinite Relaxations for Matrix-Weighted State-Estimation Problems in Robotics," *IEEE Transactions on Robotics*, vol. 40, pp. 4805–4824, 2024.

[42] C. Grebe, E. Wise, and J. Kelly, "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, Sep. 2021, pp. 1–8.

[43] J. Yang, J. Rebello, and S. L. Waslander, "Next-best-view selection for robot eye-in-hand calibration," in *2023 20th Conference on Robots and Vision (CRV)*.   IEEE, 2023, pp. 161–168.

[44] T. D. Barfoot, *State Estimation for Robotics*, 2nd ed.   Cambridge University Press, Jan. 2024.