

On the covariance of X in $AX = XB$

Huy Nguyen and Quang-Cuong Pham

Abstract—Hand-eye calibration, which consists in identifying the rigid-body transformation between a camera mounted on the robot end-effector and the end-effector itself, is a fundamental problem in robot vision. Mathematically, this problem can be formulated as: solve for X in $AX = XB$. In this paper, we provide a rigorous derivation of the covariance of the solution X , when A and B are randomly perturbed matrices. This fine-grained information is critical for applications that require a high degree of perception precision. Our approach consists in applying covariance propagation methods in $SE(3)$. Experiments involving synthetic and real calibration data confirm that our approach can predict the covariance of the hand-eye transformation with excellent precision.

Index Terms—Hand-eye calibration, Uncertainty, Calibration and Identification

I. INTRODUCTION

Hand-eye calibration, which consists in identifying the rigid-body transformation between a camera (eye) mounted on the robot end-effector and the end-effector (hand) itself, is a fundamental problem in robot vision. Mathematically, this problem can be formulated as: solve for X in $AX = XB$, where X is the unknown 4×4 hand-eye transformation matrix and A and B are known 4×4 transformation matrices (see details in Section II-A). Starting from the late 1980's, a large amount of literature has been devoted to this problem, and a number of efficient methods have been developed, see *e.g.* [1], [2], [3], [4], [5], [6].

In this paper, we are interested, not merely in solving for X , but more comprehensively, in evaluating the *covariance* of X from those of A and B , where A and B are now *randomly perturbed* transformation matrices. This fine-grained information is critical in high-precision robotics applications for several reasons.

Motivations

The uncertainty of the object pose estimation comes from three main sources: (i) the uncertainty of the object pose estimation in the camera frame, (ii) the uncertainty of the hand-eye calibration, and (iii) the uncertainty of the robot end-effector positioning. In practice, source (ii) arguably contributes the most: for instance, a tiny orientation error of 0.05 degrees in the hand-eye calibration already implies an error of 0.6 mm in object position if the latter is 700 mm away from the camera (typical viewing distance for commodity 3D cameras). In turn, having a precise knowledge of the uncertainty of the object pose estimation is critical:

- In high-precision manufacturing, it is important, not only to know the pose of an object, but also to *guarantee* that the pose estimation error is within some *tolerance*. For instance, when drilling holes in the fuselage of an aircraft, the hole position tolerance is 0.5 mm – which would be violated by an error of 0.05 degrees in the hand-eye calibration, even when assuming that the object pose estimation in the camera frame is perfect (see above);
- The precise knowledge of the object pose covariance *matrix* allows one to intelligently refine the object pose estimation by other perception modes. For instance, in visuo-tactile sensor fusion [7], knowing that the covariance of the object pose is comparatively large in the translation along, say, the X-axis will

The authors are with the School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore 639798. Corresponding author: Huy Nguyen (email: huy.nguyendinh09@gmail.com).

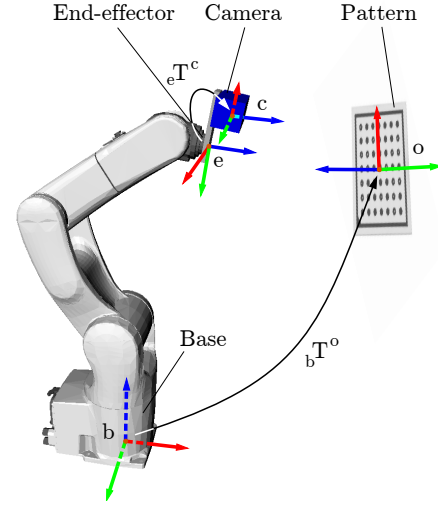


Figure 1. The hand-eye calibration problem consists in identifying the rigid-body transformation ${}_eT^c$ between a camera mounted on the end-effector of a robot and the end-effector itself.

prompt us to touch the object along that axis in order to best reduce the uncertainty.

In addition, knowing the covariance of X allows improving the calibration process itself, by *e.g.* choosing the appropriate number of measurements to achieve a desired level of precision, or choosing the appropriate matrices A and B that minimize the covariance of X .

Related works

Finding the covariance of X is challenging for several reasons. First, as X , A and B represent rigid-body transformations, they live in $SE(3)$, a subset of the space of 4×4 matrices endowed with a non-trivial Lie group structure [8]. Second, how to represent and calculate uncertainties in $SE(3)$ is by itself a complex issue, which has prompted advanced mathematical developments [9]. Finally, merely solving for X in $AX = XB$ is already a difficult problem [3], [4], [5], let alone evaluating the uncertainty of the solution.

There are a number of works dealing with the uncertainty of hand-eye calibration. In [10], based on a sensitivity analysis of closed-form solutions, some critical factors and criteria influencing the accuracy of the result are analyzed. For instance, one may try to maximize the angle between rotation axes of relative movement to reduce the influence on error in rotation, or to minimize the distance between the optical center of the camera and the calibration pattern to reduce the influence on error in translation. Based on this analysis, Shi *et al.* [11] present a algorithm to select movement pairs automatically from a series of measurements to reduce the error of the estimate. Schmidt *et al.* also introduce similar approach based on a vector quantization method [12]. In [13], Aron *et al.* present an error estimation method of the rotation part of X based on an Euler angles parameterization. The authors do not discuss how that error propagates to the translation part of X and their vision tracking measurements are also assumed to be noise-free. More fundamentally, the Euler angles formulation, as opposed to the $SE(3)$ formulation, is well-known to involve singularities.

In [6], [14], [15], authors introduced the concept of using mean and covariance of distributions on $SE(3)$ to formulate the hand-eye calibration problem. An information-theoretic approach is proposed in [14] by viewing the problem in terms of distributions on the group $SE(3)$, and minimizing the Kullback-Leibler divergence of these distributions with respect to the unknown X . This allows them to handle the problem without a priori knowledge of the

correspondence of the given data sets. In [15] the algorithm is further improved by introducing two new probabilistic approaches built on top of the Batch method. Our work exploits the benefits of applying optimization techniques directly on $\mathbf{SE}(3)$ and focuses on the derivation of the covariance of the hand-eye transformation, which is different from the mentioned works as they only focus on finding a solution to \mathbf{X} .

The idea of estimating explicitly uncertainties in the system is also by no means new. While many have studied the problem of uncertainty in the camera model (intrinsic and extrinsic parameters) [16] and the propagation of uncertainties through the camera model [9], we place emphasis on the hand-eye transformation and its uncertainty in this work.

Contribution and organization of the paper

It can be noted that none of the aforementioned works has provided a derivation of the covariance of \mathbf{X} , which is ultimately the most generic and relevant quantification of the uncertainty of the hand-eye calibration process. The goal of this paper is to rigorously work out such a derivation. Specifically, we transpose methods for forward and backward propagation of covariance [16] into the framework of uncertainty in $\mathbf{SE}(3)$ ¹. The structure of the hand-eye calibration equation raises specific technical difficulties, which we shall address in detail.

The remainder of the paper is organized as follows. In Section II, we state the hand-eye calibration problem and introduce the mathematical background of the work, which includes the representation of uncertainty in $\mathbf{SE}(3)$, and methods for forward and backward propagation of covariance. In Section III, we present our method to estimate the rotation and translation parts of the hand-eye transformation matrix and their associated covariance matrices. In Section IV, we show that the method can indeed predict with excellent precision these covariances in synthetic and real calibration datasets, and uses this information to compute the covariance of the object pose estimation in a real setting. Finally, in Section V, we conclude by discussing the advantages and drawbacks of our approach and sketch some future research directions.

II. BACKGROUND

A. Formulation of the hand-eye calibration problem

The classical hand-eye calibration method consists in looking at a fixed pattern from two different viewpoints, say 1 and 2, giving rise to the following equation

$${}_b\mathbf{T}_1^e {}_e\mathbf{T}^c {}_c\mathbf{T}_1^o = {}_b\mathbf{T}^o = {}_b\mathbf{T}_2^e {}_e\mathbf{T}^c {}_c\mathbf{T}_2^o, \quad (1)$$

where

- ${}_b\mathbf{T}_i^e$ is the transformation of the end-effector with respect to the fixed robot base at configuration i ;
- ${}_e\mathbf{T}^c$ is the constant transformation of the camera with respect to the end-effector;
- ${}_c\mathbf{T}_i^o$ is the transformation of the pattern (object) with respect to the camera at configuration i ;
- ${}_b\mathbf{T}^o$ is the constant transformation of the pattern with respect to the robot base (see Fig. 1).

Next, one can transform the above equation into

$${}_b\mathbf{T}_2^{e-1} {}_b\mathbf{T}_1^e {}_e\mathbf{T}^c = {}_e\mathbf{T}^c {}_c\mathbf{T}_2^o {}_c\mathbf{T}_1^{o-1}, \quad (2)$$

¹Many works have provided rigorous treatments for representing and associating uncertainty in $\mathbf{SE}(3)$, $\mathbf{SO}(3)$. Common notations, terminologies and theoretical tools can be found in [9], [17], [18], [19], [20]. In this work, to model the uncertainty on $\mathbf{SE}(3)$, we choose to follow the approach proposed in [9] as we find it accessible and easy to implement.

which has the form of $\mathbf{AX} = \mathbf{XB}$, where $\mathbf{X} := {}_e\mathbf{T}^c$ is the unknown hand-eye transformation, and $\mathbf{A} := {}_b\mathbf{T}_2^{e-1} {}_b\mathbf{T}_1^e$ and $\mathbf{B} := {}_c\mathbf{T}_2^o {}_c\mathbf{T}_1^{o-1}$ can be computed from respectively the robot kinematics and pattern pose estimation [3]. Next, if the fixed pattern is viewed from a large number of viewpoints, one can collect many different \mathbf{A} 's and \mathbf{B} 's. Suppose that we have a set of k measurements $(\mathbf{A}_1, \mathbf{B}_1), (\mathbf{A}_2, \mathbf{B}_2), \dots, (\mathbf{A}_k, \mathbf{B}_k)$. Since in practice these measurements are perturbed by actuator/sensor noise, the exact solution for the set of k equations $\mathbf{A}_i\mathbf{X} = \mathbf{XB}_i$ will not exist. Instead, the problem is commonly framed as an optimization problem in which \mathbf{X} is found as the transformation that “best” fits the k equalities.

In some other cases, the camera is not mounted on the end-effector but on a fixed stand. In these cases, finding the relative transformation between the camera and the robot base can also be formulated as the $\mathbf{AX} = \mathbf{XB}$ problem and can be treated by the same method.

B. Representation of rigid-body transformations and of their uncertainties

We first recall some facts and formulae involving the rotation group $\mathbf{SO}(3)$ and the special Euclidean group $\mathbf{SE}(3)$ [8]. Elements of $\mathbf{SO}(3)$ are given by the 3×3 real matrices \mathbf{R} satisfying $\mathbf{R}^\top \mathbf{R} = \mathbb{I}$ and $\det \mathbf{R} = 1$. The group $\mathbf{SE}(3)$ is defined to be

$$\mathbf{SE}(3) = \left\{ \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \middle| \mathbf{R} \in \mathbf{SO}(3), \mathbf{t} \in \mathbb{R}^3 \right\}. \quad (3)$$

$\mathbf{SO}(3)$ is a matrix Lie group, and its associated Lie algebra, denoted $\mathfrak{so}(3)$, is given by the set of 3×3 real skew-symmetric matrices. The $[\cdot]$ operator turns $\phi \in \mathbb{R}^3$ into a 3×3 member of the Lie algebra $\mathfrak{so}(3)$:

$$[\phi] := \begin{bmatrix} 0 & -\phi_3 & \phi_2 \\ \phi_3 & 0 & -\phi_1 \\ -\phi_2 & \phi_1 & 0 \end{bmatrix}. \quad (4)$$

We will use \vee as the inverse operation of $[\cdot]$. By using the matrix exponential $\exp(\cdot)$ on elements of $\mathfrak{so}(3)$, we can obtain group elements of $\mathbf{SO}(3)$ as $\mathbf{R} = \exp([\phi])$. The inverse of the exponential, or logarithm, is used to obtain the vector of exponential coordinates ϕ from the group element $\mathbf{R} \in \mathbf{SO}(3)$ as $\phi = (\log \mathbf{R})^\vee$.

To model the uncertainty on $\mathbf{SE}(3)$, we follow the approach proposed in [9]. Since there is in general no bi-invariant distance on $\mathbf{SE}(3)$ [21], finding the rotation and translation components of \mathbf{X} simultaneously would require a non-trivial rotation/translation weighting in any cases. Instead, we choose to solve them separately, which entails a number of simplifications [3]. As a consequence, the uncertainties of the rotation and the translation parts are also modeled separately.

Specifically, we assume that the rotation parts of the observations \mathbf{A}_i and \mathbf{B}_i are corrupted as follows

$$\mathbf{R}_{A_i} = \exp([\xi_{R_{A_i}}]) \bar{\mathbf{R}}_{A_i}, \quad \mathbf{R}_{B_i} = \exp([\xi_{R_{B_i}}]) \bar{\mathbf{R}}_{B_i}, \quad (5)$$

where $\bar{\mathbf{R}}_{A_i}, \bar{\mathbf{R}}_{B_i} \in \mathbf{SO}(3)$ are the means of $\mathbf{R}_{A_i}, \mathbf{R}_{B_i}$, and $\xi_{R_{A_i}}, \xi_{R_{B_i}} \in \mathbb{R}^3$ are zero-mean Gaussian perturbations with covariance matrices $\Sigma_{R_{A_i}}, \Sigma_{R_{B_i}}$, respectively.

The translation parts of the \mathbf{A}_i and \mathbf{B}_i are corrupted as follows

$$\mathbf{t}_{A_i} = \xi_{t_{A_i}} + \bar{\mathbf{t}}_{A_i}, \quad \mathbf{t}_{B_i} = \xi_{t_{B_i}} + \bar{\mathbf{t}}_{B_i}, \quad (6)$$

where $\bar{\mathbf{t}}_{A_i}, \bar{\mathbf{t}}_{B_i} \in \mathbb{R}^3$ are the means of $\mathbf{t}_{A_i}, \mathbf{t}_{B_i}$, and $\xi_{t_{A_i}}, \xi_{t_{B_i}} \in \mathbb{R}^3$ are zero-mean Gaussian perturbations with covariance matrices $\Sigma_{t_{A_i}}, \Sigma_{t_{B_i}}$, respectively.

Note that the above assumptions imply that rotation and translation noises are independent, which on some occasions (e.g., manipulator

with large deflections on the robot's links) might require further considerations.

C. Forward and backward propagation of covariance

Forward propagation. Let \mathbf{P} be a random vector in \mathbb{R}^M with mean $\bar{\mathbf{P}}$ and covariance matrix Σ . Consider a function $\mathbf{f} : \mathbb{R}^M \rightarrow \mathbb{R}^N$ that is differentiable in a neighbourhood of $\bar{\mathbf{P}}$. Then, at the first order of approximation, $\mathbf{f}(\mathbf{P})$ is a random variable with mean $\mathbf{f}(\bar{\mathbf{P}})$ and covariance matrix

$$\Sigma_f = \mathcal{J} \Sigma \mathcal{J}^\top, \quad (7)$$

where \mathcal{J} is the Jacobian matrix of \mathbf{f} at $\bar{\mathbf{P}}$.

Backward propagation. Assume now that \mathbf{P} (the parameter) is unknown, but that $\mathbf{V} := \mathbf{f}(\mathbf{P})$ (the measurement) is known and determined to be a random variable with mean $\bar{\mathbf{V}}$ and covariance matrix Σ_V . Then the best estimate for \mathbf{P} is given by

$$\mathbf{P}^* := \min_{\mathbf{P}} \|\mathbf{V} - \mathbf{f}(\mathbf{P})\|_{\Sigma_V}.$$

To estimate the covariance of \mathbf{P} , one can approximate \mathbf{f} by an affine function $\mathbf{f}(\mathbf{P}) = \mathbf{f}(\bar{\mathbf{P}}) + \mathcal{J}(\mathbf{P} - \bar{\mathbf{P}})$, which yields

$$\|\mathbf{V} - \mathbf{f}(\mathbf{P})\|_{\Sigma_V} = \|(\mathbf{V} - \bar{\mathbf{V}}) - \mathcal{J}(\mathbf{P} - \bar{\mathbf{P}})\|_{\Sigma_V}. \quad (8)$$

Using the weighted pseudo-inverse, one has

$$\mathbf{P}^* - \bar{\mathbf{P}} = (\mathcal{J}^\top \Sigma_V^{-1} \mathcal{J})^{-1} \mathcal{J}^\top \Sigma_V^{-1} (\mathbf{V} - \bar{\mathbf{V}}). \quad (9)$$

From (7), the covariance of \mathbf{P} can now be approximated at the first order by

$$\begin{aligned} \Sigma^* &= [(\mathcal{J}^\top \Sigma_V^{-1} \mathcal{J})^{-1} \mathcal{J}^\top \Sigma_V^{-1}] \Sigma_V \\ &\quad [(\mathcal{J}^\top \Sigma_V^{-1} \mathcal{J})^{-1} \mathcal{J}^\top \Sigma_V^{-1}]^\top \\ &= (\mathcal{J}^\top \Sigma_V^{-1} \mathcal{J})^{-1}. \end{aligned} \quad (10)$$

In practice, when performing an iterative least-squares optimization, one can use (10) at the last iteration to obtain the estimation of the covariance of \mathbf{P} .

Note that the quality of the approximations given by Equations (7) and (10) depends in particular on the quality of the linear approximation of \mathbf{f} .

III. DERIVATION OF THE COVARIANCE OF \mathbf{X}

Equation $\mathbf{A}_i \mathbf{X} = \mathbf{X} \mathbf{B}_i$ can be decomposed as

$$\mathbf{R}_{A_i} \mathbf{R} = \mathbf{R} \mathbf{R}_{B_i}, \quad (11)$$

$$\mathbf{R}_{A_i} \mathbf{t} + \mathbf{t}_{A_i} = \mathbf{R} \mathbf{t}_{B_i} + \mathbf{t}, \quad (12)$$

where \mathbf{R}, \mathbf{t} denote respectively the rotation and translation parts of \mathbf{X} .

A. Covariance of the rotation part of \mathbf{X}

We first consider the rotation part \mathbf{R} of \mathbf{X} . Let $[\alpha_i], [\beta_i] \in \mathfrak{so}(3)$ denote the logarithms of \mathbf{R}_{A_i} and \mathbf{R}_{B_i} respectively, i.e.

$$[\alpha_i] := \log \mathbf{R}_{A_i}, \quad [\beta_i] := \log \mathbf{R}_{B_i}. \quad (13)$$

Note that the covariance matrices of α_i and β_i can be obtained by applying the forward propagation of covariance

$$\Sigma_{\alpha_i} = \mathbf{J}(\alpha_i)^{-1} \Sigma_{\mathbf{R}_{A_i}} \mathbf{J}(\alpha_i)^{-1\top}, \quad (14)$$

$$\Sigma_{\beta_i} = \mathbf{J}(\beta_i)^{-1} \Sigma_{\mathbf{R}_{B_i}} \mathbf{J}(\beta_i)^{-1\top}, \quad (15)$$

where $\mathbf{J}(\alpha_i)$ denotes the (left) Jacobian of $\mathfrak{so}(3)$ at α_i , see [9] for more details.

Next, via logarithm mapping, equation (11) can be written as

$$\log \mathbf{R}_{A_i} = \log \mathbf{R} \mathbf{R}_{B_i} \mathbf{R}^\top = \mathbf{R} [\beta_i] \mathbf{R}^\top. \quad (16)$$

Applying the rule $\mathbf{R}[\omega] \mathbf{R}^\top = [\mathbf{R}\omega]$ for $\mathbf{R} \in \mathfrak{SO}(3)$ and $[\omega] \in \mathfrak{so}(3)$, one has

$$\alpha_i = \mathbf{R} \beta_i. \quad (17)$$

In order to use the uncertainty model in $\mathfrak{SO}(3)$, we define a random variable ξ_R that represents the difference between \mathbf{R} and the current estimate $\hat{\mathbf{R}}$ by

$$\mathbf{R} = \exp([\xi_R]) \hat{\mathbf{R}}.$$

Next, to apply the backward propagation of covariance, one needs the measurement vectors α_i and β_i to appear on the same side of the equation. To achieve this without making it too complex, we use a trick from [16], which consists in “copying” the β_i ’s on both sides, as follows

$$\underbrace{\begin{pmatrix} \beta_1 \\ \alpha_1 \\ \vdots \\ \beta_k \\ \alpha_k \end{pmatrix}}_{\mathbf{V}} = \underbrace{\begin{pmatrix} \beta_1 \\ \exp([\xi_R]) \hat{\mathbf{R}} \beta_1 \\ \vdots \\ \beta_k \\ \exp([\xi_R]) \hat{\mathbf{R}} \beta_k \end{pmatrix}}_{\mathbf{f}(\mathbf{P})}. \quad (18)$$

Now, the measurement vector is given by $\mathbf{V} := (\mathbf{V}_1, \dots, \mathbf{V}_k)$, where $\mathbf{V}_i := (\beta_i, \alpha_i)$, and the parameter vector is given by $\mathbf{P} := (\xi_R, \beta_1, \dots, \beta_k)$.

Since the noise of α_i ’s and β_i ’s are independent (α is caused by robot kinematics while β_i is caused by object pose estimation in the camera frame), the covariance matrix of the measurement vector is given by

$$\Sigma_V := \text{diag}(\Sigma_{V_1}, \Sigma_{V_2}, \dots, \Sigma_{V_k}), \quad (19)$$

with $\Sigma_{V_i} := \text{diag}(\Sigma_{\beta_i}, \Sigma_{\alpha_i})$.

Now, the covariance-weighted minimization is given by

$$\min_{\mathbf{P}} \|\mathbf{V} - \mathbf{f}(\mathbf{P})\| = \min_{\mathbf{P}} \sum_i^k (\mathbf{V}_i - \mathbf{f}(\mathbf{P})_i)^\top \Sigma_{V_i}^{-1} (\mathbf{V}_i - \mathbf{f}(\mathbf{P})_i).$$

This minimization problem can be solved by iteratively updating the estimate of the parameter vector by the rules

$$\hat{\mathbf{R}}_{(j+1)} = \exp([\xi_R]) \hat{\mathbf{R}}_{(j)}, \quad (20)$$

$$\hat{\beta}_{i(j+1)} = \hat{\beta}_{i(j)} + \delta \beta_i, \quad (21)$$

where at each step (j) the update vector $\delta := (\xi_R, \delta \beta_1, \dots, \delta \beta_k)$ is found by solving the normal equation

$$\mathcal{J}_f^\top \Sigma_V^{-1} \mathcal{J}_f \delta = \mathcal{J}_f^\top \Sigma_V^{-1} (\mathbf{V} - \mathbf{f}(\mathbf{P})). \quad (22)$$

The Jacobian of \mathbf{f} has the form

$$\mathcal{J}_f = \begin{bmatrix} \mathcal{J}^{\xi_R} | \mathcal{J}^\beta \\ \mathcal{J}_1^{\xi_R} & \mathcal{J}_1^\beta & & \\ \mathcal{J}_2^{\xi_R} & & \mathcal{J}_2^\beta & \\ \vdots & & & \ddots \\ \mathcal{J}_k^{\xi_R} & & & \mathcal{J}_k^\beta \end{bmatrix}, \quad (23)$$

$$\text{where } \mathcal{J}_i^{\xi_R} := \begin{bmatrix} \mathbf{0} \\ -[\hat{\mathbf{R}} \hat{\beta}_i] \end{bmatrix}, \mathcal{J}_i^\beta := \begin{bmatrix} \mathbb{I} \\ \hat{\mathbf{R}} \end{bmatrix}. \quad (24)$$

The set of equations (22) may now written in block form as

$$\begin{bmatrix} \mathcal{J}^{\xi_R\top} \Sigma_V^{-1} \mathcal{J}^{\xi_R} & \mathcal{J}^{\xi_R\top} \Sigma_V^{-1} \mathcal{J}^\beta \\ \mathcal{J}^{\beta\top} \Sigma_V^{-1} \mathcal{J}^{\xi_R} & \mathcal{J}^{\beta\top} \Sigma_V^{-1} \mathcal{J}^\beta \end{bmatrix} \begin{pmatrix} \xi_R \\ \delta \beta \end{pmatrix} = \begin{pmatrix} \mathcal{J}^{\xi_R\top} \Sigma_V^{-1} (\mathbf{V} - \mathbf{f}(\mathbf{P})) \\ \mathcal{J}^{\beta\top} \Sigma_V^{-1} (\mathbf{V} - \mathbf{f}(\mathbf{P})) \end{pmatrix}. \quad (25)$$

To simplify the left-hand side of (25), let

$$\mathbf{U} := \mathcal{J}^{\xi_R\top} \Sigma_V^{-1} \mathcal{J}^{\xi_R} = \sum_i^k \mathcal{J}_i^{\xi_R\top} \Sigma_{V_i}^{-1} \mathcal{J}_i^{\xi_R}, \quad (26)$$

$$\mathbf{W} := \mathcal{J}^{\xi_R\top} \Sigma_V^{-1} \mathcal{J}^\beta = [\mathbf{W}_1, \dots, \mathbf{W}_k], \quad (27)$$

$$\mathbf{Z} := \mathcal{J}^{\beta\top} \Sigma_V^{-1} \mathcal{J}^\beta = \text{diag}(\mathbf{Z}_1, \dots, \mathbf{Z}_k), \quad (28)$$

$$\text{where } \mathbf{W}_i := \mathcal{J}_i^{\xi_R\top} \Sigma_{V_i}^{-1} \mathcal{J}_i^\beta \text{ and } \mathbf{Z}_i := \mathcal{J}_i^{\beta\top} \Sigma_{V_i}^{-1} \mathcal{J}_i^\beta. \quad (29)$$

As for the right-hand side of (25), let

$$\begin{aligned}\epsilon_{\xi_R} &:= \mathcal{J}^{\xi_R \top} \Sigma_V^{-1} (V - f(P)) \\ &= \sum_i^k \mathcal{J}_i^{\xi_R \top} \Sigma_{V_i}^{-1} (V_i - f(P)_i),\end{aligned}\quad (30)$$

$$\epsilon_{\beta} := \mathcal{J}^{\beta \top} \Sigma_V^{-1} (V - f(P)) = (\epsilon_{\beta_1}, \dots, \epsilon_{\beta_k}), \quad (31)$$

$$\text{where } \epsilon_{\beta_i} := \mathcal{J}_i^{\beta \top} \Sigma_{V_i}^{-1} (V_i - f(P)_i). \quad (32)$$

To solve equations (25), one can left-multiply both sides by $\begin{bmatrix} \mathbb{I} & \mathbf{W}\mathbf{Z}^{-1} \\ \mathbf{0} & \mathbb{I} \end{bmatrix}$, which yields

$$(U - \mathbf{W}\mathbf{Z}^{-1}\mathbf{W}^\top) \xi_R = \epsilon_{\xi_R} - \mathbf{W}\mathbf{Z}^{-1} \epsilon_{\beta}, \quad (33)$$

$$\mathbf{Z} \delta_{\beta} = \epsilon_{\beta} - \mathbf{W}^\top \xi_R. \quad (34)$$

The above equations can now be solved to find the updating vectors ξ_R and δ_{β} .

Applying backward propagation of covariance, a first-order approximation of the covariance of P is given the following matrix, taken at the last iteration,

$$\begin{aligned}\Sigma^* &= (\mathcal{J}_f^\top \Sigma_V^{-1} \mathcal{J}_f)^{-1} \\ &= \begin{bmatrix} \Sigma_R & -\Sigma_R \mathbf{W}\mathbf{Z}^{-1} \\ -(\Sigma_R \mathbf{W}\mathbf{Z}^{-1})^\top & (\mathbf{W}\mathbf{Z}^{-1})^\top \Sigma_R \mathbf{W}\mathbf{Z}^{-1} + \mathbf{Z}^{-1} \end{bmatrix}.\end{aligned}$$

Truncation of the covariance matrix Σ^* gives covariance matrices for parameters ξ_R and β_i^* separately. The covariance of ξ_R is given by the top-left block of Σ^* , that is:

$$\Sigma_R = (U - \mathbf{W}\mathbf{Z}^{-1}\mathbf{W}^\top)^{-1} = (U - \sum_i^k \mathbf{W}_i \mathbf{Z}_i^{-1} \mathbf{W}_i^\top)^{-1}. \quad (35)$$

The covariance of β_i^* is given by:

$$\Sigma_{\beta_i^*} = (\mathbf{W}_i \mathbf{Z}_i^{-1})^\top \Sigma_R \mathbf{W}_i \mathbf{Z}_i^{-1} + \mathbf{Z}_i^{-1}. \quad (36)$$

The cross-covariances of ξ_R and β_i^* are $\Sigma_{R\beta_i^*} = -\Sigma_R \mathbf{W}_i \mathbf{Z}_i^{-1}$.

B. Covariance of the translation part of X

We now consider the translation part t of X . Let $q_i := R t_{B_i} - t_{A_i}$. Equations (12) can be written as

$$q_i = (R_{A_i} - \mathbb{I})t. \quad (37)$$

Note that the covariance matrices of q_i can be approximated by applying the forward propagation of covariance

$$\Sigma_{q_i} = \Sigma_{t_A} + R^* \Sigma_{t_B} R^{*\top} + [R^* t_{B_i}] \Sigma_R [R^* t_{B_i}]^\top, \quad (38)$$

where R^* is the optimal rotation found in the previous section, and Σ_R is the corresponding covariance.

We then update the α_i by substituting the optimal values of R and β_i found in last section into equation (17). The covariance matrices of the corresponding $R_{A_i}^*$ can be approximated by forward propagation of covariance

$$\begin{aligned}\Sigma_{R_{A_i}^*} &= J(\alpha_i^*) \{ [R^* \beta_i^*] \Sigma_R [R^* \beta_i^*]^\top - R^* \Sigma_{R\beta_i^*}^\top [R^* \beta_i^*]^\top \\ &\quad - [R^* \beta_i^*] \Sigma_{R\beta_i^*} R^{*\top} + R^* \Sigma_{\beta_i^*} R^{*\top} \} J(\alpha_i^*)^\top.\end{aligned}\quad (39)$$

Applying the same trick as previously, we “copy” the $R_{A_i}^*$ ’s on both sides of the equation, as follows

$$\underbrace{\begin{pmatrix} R_{A_1}^* \\ q_1 \\ \vdots \\ R_{A_k}^* \\ q_k \end{pmatrix}}_V = \underbrace{\begin{pmatrix} R_{A_1}^* \\ (R_{A_1}^* - \mathbb{I})\hat{t} \\ \vdots \\ R_{A_k}^* \\ (R_{A_k}^* - \mathbb{I})\hat{t} \end{pmatrix}}_{f(P)}. \quad (40)$$

Now the measurement vector is given by $V := (V_1, \dots, V_k)$, where $V_i := (R_{A_i}^*, q_i)$, and the parameter vector is given by $P := (t, R_{A_1}^*, \dots, R_{A_k}^*)$.

The covariance matrix of the measurement vector is given by

$$\Sigma_V := \text{diag}(\Sigma_{V_1}, \Sigma_{V_2}, \dots, \Sigma_{V_k}), \quad (41)$$

where

$$\Sigma_{V_i} := \begin{bmatrix} \Sigma_{R_{A_i}^*} & \Sigma_{R_{A_i}^* q_i} \\ \Sigma_{q_i R_{A_i}^*} & \Sigma_{q_i} \end{bmatrix}. \quad (42)$$

The cross-covariance matrices $\Sigma_{q_i R_{A_i}^*}$ can be estimated by applying forward propagation of covariance:

$$\begin{aligned}\Sigma_{q_i R_{A_i}^*} &= -[R^*, t_A] \Sigma_{R R_{A_i}^*} \\ &= [R^*, t_A] (\Sigma_R [R^* \beta_i^*]^\top - \Sigma_{R\beta_i^*} R^{*\top}) J(\alpha_i^*)^\top\end{aligned}\quad (43)$$

Now, the covariance-weighted minimization is given by

$$\min_P \|V - f(P)\| = \min_P \sum_i^k (V_i - f(P)_i)^\top \Sigma_{V_i}^{-1} (V_i - f(P)_i).$$

We solve this by iteratively updating the estimate of the parameter vector by the rules

$$\hat{t}_{(j+1)} = \hat{t}_{(j)} + \delta_t, \quad (45)$$

$$\hat{R}_{A_i(j+1)}^* = \exp([\xi_{R_{A_i}}]) \hat{R}_{A_i(j)}^*, \quad (46)$$

where at each step (j) the update vector $\delta := (\delta_t, \xi_{R_A}) = (\delta_t, \xi_{R_{A_1}}, \dots, \xi_{R_{A_k}})$ is found by solving the normal equation

$$\mathcal{J}_f^\top \Sigma_V^{-1} \mathcal{J}_f \delta = \mathcal{J}_f^\top \Sigma_V^{-1} (V - f(P)). \quad (47)$$

The Jacobian matrix has the form

$$\mathcal{J}_f = [\mathcal{J}^t | \mathcal{J}^{\xi_{R_A}}], \quad (48)$$

$$\text{where } \mathcal{J}^t := (\mathcal{J}_1^t, \dots, \mathcal{J}_k^t), \quad (49)$$

$$\mathcal{J}^{\xi_{R_A}} := \text{diag}(\mathcal{J}_1^{\xi_{R_A}}, \dots, \mathcal{J}_k^{\xi_{R_A}}), \quad (50)$$

$$\mathcal{J}_i^t := \begin{bmatrix} \mathbf{0} \\ \hat{R}_{A_i}^* - \mathbb{I} \end{bmatrix}, \quad (51)$$

$$\mathcal{J}_i^{\xi_{R_A}} := \begin{bmatrix} \mathbb{I} \\ -[\hat{R}_{A_i}^* \hat{t}_X] \end{bmatrix}. \quad (52)$$

For the rest of the derivation, we following the same procedure as previously derived. One thus can obtain the update vectors from

$$(U - \mathbf{W}\mathbf{Z}^{-1}\mathbf{W}^\top) \delta_t = \epsilon_t - \mathbf{W}\mathbf{Z}^{-1} \epsilon_{\xi_{R_A}}, \quad (53)$$

$$\mathbf{Z} \xi_{R_A} = \epsilon_{\xi_{R_A}} - \mathbf{W}^\top \delta_t, \quad (54)$$

$$\text{where } U := \sum_i^k \mathcal{J}_i^{t \top} \Sigma_{V_i}^{-1} \mathcal{J}_i^t, \quad (55)$$

$$\mathbf{W} := [\mathbf{W}_1, \dots, \mathbf{W}_k] \text{ with } \mathbf{W}_i := \mathcal{J}_i^{t \top} \Sigma_{V_i}^{-1} \mathcal{J}_i^{\xi_{R_A}}, \quad (56)$$

$$\mathbf{Z} := \text{diag}(\mathbf{Z}_1, \dots, \mathbf{Z}_k) \text{ with } \mathbf{Z}_i := \mathcal{J}_i^{\xi_{R_A} \top} \Sigma_{V_i}^{-1} \mathcal{J}_i^{\xi_{R_A}}, \quad (57)$$

$$\epsilon_t := \sum_i^k \mathcal{J}_i^{t \top} \Sigma_{V_i}^{-1} (V_i - f(P)_i), \quad (58)$$

$$\epsilon_{\xi_{R_A}} := (\epsilon_{\xi_{R_{A_1}}}, \dots, \epsilon_{\xi_{R_{A_k}}}), \quad (59)$$

$$\text{with } \epsilon_{\xi_{R_{A_i}}} = \mathcal{J}_i^{\xi_{R_A} \top} \Sigma_{V_i}^{-1} (V_i - f(P)_i). \quad (60)$$

At the last iteration, a first-order approximation of the covariance matrix of t is given by

$$\Sigma_t = (U - \mathbf{W}\mathbf{Z}^{-1}\mathbf{W}^\top)^{-1} = (U - \sum_i^k \mathbf{W}_i \mathbf{Z}_i^{-1} \mathbf{W}_i^\top)^{-1}. \quad (61)$$

IV. EXPERIMENTS

We now validate the proposed method in predicting the covariance of the hand-eye transformation using synthetic and real calibration data. Using the covariance of \mathbf{X} , we are then in a position to compute the covariance of the object pose estimation in a real setting. Our implementation is open-source and is available at <https://github.com/dinhhuu2109/python-cope>.

A. Synthetic calibration data

To generate synthetic data, we start by selecting a random transformation matrix $\bar{\mathbf{X}} = (\bar{\mathbf{R}}, \bar{\mathbf{t}})$, which serves as the true hand-eye transformation. We then generate $M = 250$ dataset, each dataset comprising $k = 30$ corrupted pairs $(\mathbf{A}_i, \mathbf{B}_i)_{i \in [1, k]}$. Each corrupted pair is generated as follows. First, we generate a random *uncorrupted* pair $(\bar{\mathbf{A}}_i, \bar{\mathbf{B}}_i)$, which are drawn from a simulation model of the system in Fig. 1 and verifies $\bar{\mathbf{A}}_i \bar{\mathbf{X}} = \bar{\mathbf{X}} \bar{\mathbf{B}}_i$ exactly. Next, we add noise to $\bar{\mathbf{A}}_i$ and $\bar{\mathbf{B}}_i$ as explained in Section II-B. The covariance matrices of the noise are chosen arbitrarily as

$$\Sigma_{\mathbf{R}_{A_i}} := \lambda \text{diag}(5, 2, 3), \quad (62)$$

$$\Sigma_{\mathbf{R}_{B_i}} := \lambda \text{diag}(7, 2, 8), \quad (63)$$

$$\Sigma_{\mathbf{t}_{A_i}} := \lambda \text{diag}(0.1, 0.2, 0.5), \quad (64)$$

$$\Sigma_{\mathbf{t}_{B_i}} := \lambda \text{diag}(0.7, 0.8, 0.1), \quad (65)$$

where $\lambda \in \mathbb{R}$ is a scaling parameter that allows us to change the magnitude of the uncertainties.

At each noise level λ , we evaluate the covariance of \mathbf{X} following two methods

- *Our method*: For some dataset $m \in [1, M]$, we compute the covariance matrices $\Sigma_{\mathbf{R}_m}, \Sigma_{\mathbf{t}_m}$ using the proposed method (PRedicted). In fact, $\Sigma_{\mathbf{R}_m}, \Sigma_{\mathbf{t}_m}$ are slightly different across the M datasets. This reflects the influence of the data distribution on the estimation of the hand-eye transformation. However, since the differences are very small compared to the standard deviations, we use the averages for our comparison. Also note that the accuracy analyses performed here only focus on the covariance matrices since our experiments show that the standard deviations are always much larger than the biases.;
- *Empirical Estimation*: For each dataset $m \in [1, M]$, we find the rotation and translation $\hat{\mathbf{R}}_m, \hat{\mathbf{t}}_m$ that optimally fit the k hand-eye equations following the proposed method. We then compute the covariance matrices $\Sigma_{\mathbf{R}}^{\text{ES}}, \Sigma_{\mathbf{t}}^{\text{ES}}$ across the M datasets as

$$\Sigma_{\mathbf{R}}^{\text{ES}} := \frac{1}{M} \sum_m \xi_{\mathbf{R}_m} \xi_{\mathbf{R}_m}^\top, \quad (66)$$

$$\Sigma_{\mathbf{t}}^{\text{ES}} := \frac{1}{M} \sum_m \xi_{\mathbf{t}_m} \xi_{\mathbf{t}_m}^\top, \quad (67)$$

where $\xi_{\mathbf{R}_m} := (\log(\hat{\mathbf{R}}_m \bar{\mathbf{R}}^{-1}))^\vee$, $\xi_{\mathbf{t}_m} := \hat{\mathbf{t}}_m - \bar{\mathbf{t}}$.

Fig. 2 shows projections of the covariance ellipsoids on pairs of axes for the Empirical Estimation and our method, when $\lambda = 10^{-4}$. It can be noted that the proposed method provides an excellent estimation of the covariance of the hand-eye transformation.

To gauge the performance at different noise levels, we use the following metrics

$$\varepsilon = \frac{\sqrt{\text{tr}((\Sigma^{\text{PR}} - \Sigma^{\text{ES}})^\top (\Sigma^{\text{PR}} - \Sigma^{\text{ES}}))}}{\sqrt{\text{tr}(\Sigma^{\text{ES}}^\top \Sigma^{\text{ES}})}}. \quad (68)$$

Fig 3 shows that our algorithm can cope well with increasing magnitudes of the measurement uncertainty. The estimation errors remain low overall, and increases slightly with the magnitude of the noise, since larger noise levels increase the number of local minima

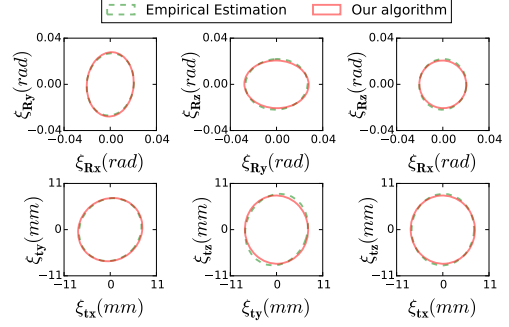


Figure 2. Projection of the one-standard-deviation covariance ellipsoids on pairs of axes for $\Sigma_{\mathbf{R}}$ and $\Sigma_{\mathbf{t}}$, shown for Empirical Estimation and our method on synthetic data when $\lambda = 10^{-4}$, where $\xi_{\mathbf{R}_x}, \xi_{\mathbf{R}_y}, \xi_{\mathbf{R}_z}$ are errors in rotation around X, Y, Z-axis and $\xi_{\mathbf{t}_x}, \xi_{\mathbf{t}_y}, \xi_{\mathbf{t}_z}$ are errors in translation along X, Y, Z-axis.

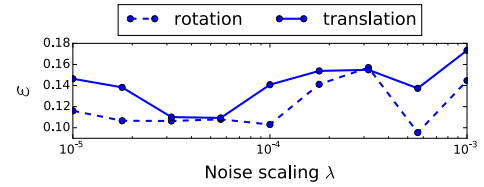


Figure 3. Difference between the covariances obtained from Empirical Estimation and predicted by our method, at different noise levels.

at each iteration. Regarding the computation cost, our method is naturally several magnitude faster than the Empirical Estimation.

It is also worth noting that the closed form solution in [3] always yields slightly higher covariance compared to our method. This is because our method does optimally minimize the uncertainty of the estimated transformation by taking into account the measurement noise.

B. Real calibration data

We now validate the proposed method on actual calibration data obtained from our robot system, which consists of a 3D camera mounted on a 6-DOF industrial manipulator, as shown in Fig. 1.

1) *Covariances of \mathbf{A} and \mathbf{B} in the actual system*: We first need to estimate the covariances of the \mathbf{A}_i 's and \mathbf{B}_i 's in our system, so that we can give them as inputs to our method.

As our industrial manipulator has a very high accuracy compared to that of our camera system², we assume that the noise on the \mathbf{A}_i 's is negligible.

Regarding the \mathbf{B}_i 's, the $\mathbf{B}_1, \dots, \mathbf{B}_k$ are assumed to have the same noise distributions: $\forall i, \Sigma_{\mathbf{R}_{B_i}} = \Sigma_{\mathbf{R}_B}, \Sigma_{\mathbf{t}_{B_i}} = \Sigma_{\mathbf{t}_B}$. We experimentally collect randomly 500 pairs of \mathbf{A}_i and \mathbf{B}_i from our system, with the only restriction that distances from pattern to camera are from 700 mm to 1100 mm. Next, we generate $M = 300$ datasets, each dataset comprising $k = 30$ pairs $(\mathbf{A}_i, \mathbf{B}_i)$ randomly selected from the collected pairs.

The rotation and translation errors of \mathbf{B}_i 's are then computed as

$$\xi_{\mathbf{R}_{B_i}} = (\log(\mathbf{R}_{B_i} \bar{\mathbf{R}}^{-1}))^\vee, \quad \xi_{\mathbf{t}_{B_i}} = \mathbf{t}_{B_i} - \bar{\mathbf{t}}, \quad (69)$$

where the ground truth is $\bar{\mathbf{B}}_i = \bar{\mathbf{X}}^{-1} \mathbf{A}_i \bar{\mathbf{X}}$. Note that we assumed rotation and translation noises are independent.

²Using a well-calibrated tooltip and a calibrating plate with well-manufactured ground truth points, our experiments show that the tooltip can achieve up to 0.3mm in error of the translation part. This can be considered as negligible compared to the uncertainty of the pose estimation using our camera system.

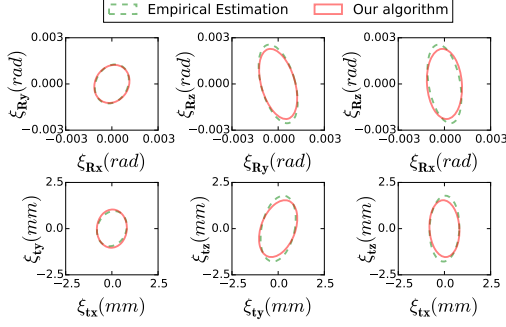


Figure 4. One-standard-deviation covariance ellipsoids of the hand-eye transformation Σ_R and Σ_t , shown for Empirical Estimation and our algorithm with real data, axes' labels are similar as to Fig. 2.

Since the true transformation \bar{X} is unknown in the real system, we use

$$R^{\text{avg}} = \exp \left(\left[\frac{1}{M} \sum_m (\log \hat{R}_m)^\vee \right] \right), \quad t^{\text{avg}} = \frac{1}{M} \sum_m \hat{t}_m, \quad (70)$$

as the ground truth³. Estimating \hat{R}_m, \hat{t}_m using our method would require information of A_i 's, B_i 's noise, hence we use [3] instead.

Then the empirical covariance matrices of B_i can be estimated similarly to Equations (66, 67):

$$\Sigma_{R_{B_i}} = \begin{bmatrix} 5.7 & -2.3 & -2.8 \\ -2.3 & 18.1 & 3.8 \\ -2.8 & 3.8 & 1.8 \end{bmatrix} 10^{-6} \text{rad}^2, \quad (71)$$

$$\Sigma_{t_{B_i}} = \begin{bmatrix} 13.1 & 0.7 & 0.9 \\ 0.7 & 2.1 & -0.3 \\ 0.9 & -0.3 & 0.8 \end{bmatrix} 10^{-6} \text{mm}^2. \quad (72)$$

2) *Validation*: To validate our method, we collect another 500 pairs of A_i and B_i from our system. We constrain the robot motion so that it covers the same area as that used for determining the noise on B_i . We then generate $M = 300$ datasets, each dataset comprising $k = 30$ pairs (A_i, B_i) randomly selected from the collected pairs. The covariance matrices are computed from these datasets using the Empirical Estimation and our method, in the same manner as previously.

Fig. 4 provides projections of the covariance ellipsoids on pairs of axes shown for two methods. We see that the proposed method delivers a good estimation of the covariances. We do not believe there has been another methods of estimating uncertainty of the hand-eye transformation. Moreover, the proposed method is also relatively easy to replicate and use in practical applications.

3) *Assessing a possible systematic bias*: Note that the study of the covariance would not be meaningful if there exists a systematic bias in the estimation of X . However, direct evaluation of the systematic bias is difficult – most papers on hand-eye calibration indeed do not contain such a direct evaluation. We, therefore, focus on the following indirect evaluation.

Consider first the average X^{avg} computed across the M datasets. We put the object to a fixed position, i.e. ${}_bT^o$ is constant. We then observed the object from 300 robot poses, yielding 300 pairs of ${}_bT_i^e, {}_cT_i^o$. If there is no systematic bias in the estimation of X^{avg} and no uncertainty in ${}_cT_i^o$, then we shall have ${}_bT_i^e X^{\text{avg}} {}_cT_i^o$ constant for all i (they are all equal to ${}_bT^o$). Thus, the “standard deviation” of ${}_bT_i^e X^{\text{avg}} {}_cT_i^o$ is a possible indirect measure of the systematic bias.

³Note that there are other ways to calculate the mean R_{mean} of a rotation distribution (e.g. by the condition $\sum_m^M \log(R_{\text{mean}}^\top R_m) = 0$ [20]). However, as we noticed that the rotations of X found in our experiment are only distributed around the identity, equation (70) is used to simplify the calculation.

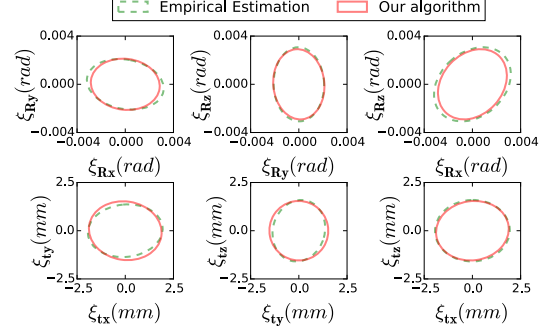


Figure 5. One-standard-deviation covariance ellipsoids of the object pose (w.r.t the robot base) Σ_{R_Y} and Σ_{t_Y} , shown for Empirical Estimation and our algorithm, axes' labels are similar as to Fig. 2. The distances from the pattern to the camera are from 700 mm to 1100 mm.

In our setting, we found that the “standard deviation” of the translation part of ${}_bT_i^e X^{\text{avg}} {}_cT_i^o$ was 0.7 mm. Since this also contains the uncertainty of ${}_cT_i^o$, the systematic bias must be quite small, compared to the translation part of the covariance of X , which was found to be around 2 mm. As another element of comparison, in case one considers the X resulting from a random dataset m , then the “standard deviation” of the translation part of ${}_bT_i^e X {}_cT_i^o$ would be around 2 mm.

C. Covariance of the object pose estimation

Using the covariance of X previously obtained, we are now in a position to predict the covariance of the object pose estimation, which is our ultimate goal. Here, we demonstrate the propagation of uncertainties to the object pose estimation using the same robot system as previously (see Fig. 1).

Recall that the constant transformation of the pattern (object) with respect to the robot base is given by

$$Y := {}_bT^o = {}_bT_i^e {}_cT_i^o. \quad (73)$$

The covariances of ${}_bT_i^e$ are assumed to be negligible, and the covariance of ${}_cT^o (= X)$ can be estimated by using the procedure proposed in Section IV-B. Thus, to predict the mean and the covariance of Y , one needs now to estimate the covariances of ${}_cT_i^o$.

Suppose that ${}_cT_1^o, \dots, {}_cT_k^o$ have the same noise distribution, i.e. $\forall i, \Sigma_{R_{{}_cT_i^o}} = \Sigma_{R_{{}_cT_1^o}}, \Sigma_{t_{{}_cT_i^o}} = \Sigma_{t_{{}_cT_1^o}}$. We begin by experimentally collecting 500 pairs of ${}_bT_i^e, {}_cT_i^o$ from our system. The rotation and translation errors of ${}_cT_i^o$'s are then computed similar to (69), where the ground truth is ${}_cT_i^o = (\bar{Y}^{-1} {}_bT_i^e \bar{X})^{-1}$. As discussed earlier, we will use (70) instead of the true value of \bar{X} . Regarding \bar{Y} , one can transform (73) into

$${}_bT_p^e {}_bT_q^{e-1} Y = Y {}_cT_p^{o-1} {}_cT_q^o, p \neq q, \quad (74)$$

which has the form of $A'_i Y = Y B'_i$, where $A' := {}_bT_p^e {}_bT_q^{e-1}$ and $B' := {}_cT_p^{o-1} {}_cT_q^o$. Hence, \bar{Y} can also be computed in the same manner as computing \bar{X} .

We now collect 500 pairs of ${}_bT_i^e, {}_cT_i^o$ from our system. We then generate $M = 300$ datasets, each dataset comprising $k = 30$ (A_i, B_i) pairs computed from $({}_bT_i^e, {}_cT_i^o)$ pairs randomly selected from the 500 collected pairs.

Next, we evaluate the covariance of Y following two methods

- *Our method*: For some dataset $m \in [1, M]$, we compute the covariance matrices $\Sigma_{R_{Y_m}}, \Sigma_{t_{Y_m}}$ using the propagation method described in Appendix A (PRedicted). Since the $\Sigma_{R_{Y_m}}, \Sigma_{t_{Y_m}}$ are only slightly different across the M datasets, we use the averages for our comparison;
- *Empirical Estimation*: For each dataset $m \in [1, M]$, we compute $Y_m = {}_bT_m^e = {}_b\hat{T}_m^e {}_c\hat{T}_m^o$, where ${}_b\hat{T}_m^e$ and ${}_c\hat{T}_m^o$

are randomly selected from the collected pairs; the rotation and translation of ${}^e\hat{T}_m^c$ together with their covariances are computed using our method as proposed in Section III (see equations (36, 61)). Next, the covariances are computed similarly to Equations (66, 67).

Fig. 5 shows the one-standard-deviation covariance ellipsoids shown for two methods. One can see that our prediction matches very well the covariances from the Empirical Estimation.

In absolute values, the covariance of the hand-eye calibration compounds with that of the object pose estimation in the camera frame, resulting in a relatively large overall covariance for the object pose estimation in the robot frame, around 2 mm in standard deviation. This again emphasizes the need of having access to the covariance of the hand-eye transformation. This fine-grained information tells us how confident we can be regarding the object pose estimation and shall also enable us to design new perception algorithms and methods for reaching higher precision, by *e.g.* visuo-tactile sensor fusion.

V. CONCLUSION

In this paper, we have presented a rigorous derivation of the covariance of the solution \mathbf{X} , when \mathbf{A} and \mathbf{B} are randomly perturbed matrices. Our approach consists in transposing methods for forward and backward propagation of covariance into the framework of uncertainty in $\mathbf{SE}(3)$. Experiments involving synthetic and real calibration data show that our approach can predict the covariance of the hand-eye transformation with excellent precision.

While these estimates could also be provided by the Empirical Estimation, such a method would require collecting a large number of samples, which is not practical. Furthermore, the Empirical Estimation yields no insights into how the uncertainties on the measurements of \mathbf{A} and \mathbf{B} propagate to the uncertainty of the hand-eye transformation. By contrast, in our method, by analyzing critical factors influencing the covariance of \mathbf{X} , for instance, based on the formulae (36) and (61), one may be able to refine the calibration process to achieve a higher precision, by *e.g.* determining the appropriate number of sample viewpoints or choosing their optimal distribution, which is the object of our future research.

ACKNOWLEDGMENT

This work was supported in part by NTUitive Gap Fund NGF-2016-01-028 and SMART Innovation Grant NG000074-ENG.

APPENDIX A

PROPAGATING UNCERTAINTIES WHEN ROTATION AND TRANSLATION ARE DECOUPLED

In this Section, we present our extension of the covariance propagation method in [9] to the case where rotation and translation are decoupled. Interested readers may also refer to [17], [18], [19], [20] for other treatments for representing and propagating uncertainty on Lie groups.

Consider two noisy poses \mathbf{T}_1 and \mathbf{T}_2 , whose nominal values and associated uncertainties are $\{\bar{\mathbf{R}}_1, \Sigma_{R_1}\}, \{\bar{\mathbf{t}}_1, \Sigma_{t_1}\}$ and $\{\bar{\mathbf{R}}_2, \Sigma_{R_2}\}, \{\bar{\mathbf{t}}_2, \Sigma_{t_2}\}$ respectively.

Let $\mathbf{T}_{12} = \mathbf{T}_1\mathbf{T}_2$ be the compounded pose, we have

$$\bar{\mathbf{R}}_{12} = \bar{\mathbf{R}}_1\bar{\mathbf{R}}_2, \quad \bar{\mathbf{t}}_{12} = \bar{\mathbf{R}}_1\bar{\mathbf{t}}_2 + \bar{\mathbf{t}}_1. \quad (75)$$

Similar to [9] (Section III), the covariance matrix of the rotation can be estimated by:

$$\Sigma_{R_{12}} \approx \Sigma_{R_1} + \Sigma'_{R_2} + \frac{1}{12}(\mathbf{A}_1\Sigma'_{R_2} + \Sigma'_{R_2}\mathbf{A}_1^\top + \Sigma_{R_1}\mathbf{A}_2 + \Sigma_{R_1}\mathbf{A}_2^\top) + \frac{\mathbf{B}}{4}, \quad (76)$$

$$\text{where} \quad \Sigma'_{R_2} := \bar{\mathbf{R}}_1\Sigma_{R_2}\bar{\mathbf{R}}_1^\top, \quad (77)$$

$$\mathbf{A}_1 := \langle\langle\Sigma_{R_1}\rangle\rangle, \mathbf{A}_2 := \langle\langle\Sigma'_{R_2}\rangle\rangle, \mathbf{B} := \langle\langle\Sigma_{R_1}, \Sigma'_{R_2}\rangle\rangle, \quad (78)$$

$$\text{and} \quad \langle\langle\mathbf{M}\rangle\rangle := -\text{tr}(\mathbf{M})\mathbf{I} + \mathbf{M}, \quad (79)$$

$$\langle\langle\mathbf{M}, \mathbf{N}\rangle\rangle := \langle\langle\mathbf{M}\rangle\rangle\langle\langle\mathbf{N}\rangle\rangle + \langle\langle\mathbf{NM}\rangle\rangle, \quad (80)$$

with $\mathbf{M}, \mathbf{N} \in \mathbb{R}^{n \times n}$.

Regarding the translation vector, its covariance matrix can be estimated simply by using the forward propagation method of Section II-C:

$$\Sigma_{t_{12}} \approx \Sigma_{t_1} + \bar{\mathbf{R}}_1\Sigma_{t_2}\bar{\mathbf{R}}_1^\top + [\bar{\mathbf{R}}_1\bar{\mathbf{t}}_2]\Sigma_{R_1}[\bar{\mathbf{R}}_1\bar{\mathbf{t}}_2]^\top. \quad (81)$$

In summary, to compound two poses, we propagate the means using (75) and the covariances using (76,81).

REFERENCES

- [1] Y. C. Shiu and S. Ahmad, "Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form $\mathbf{ax}=\mathbf{xb}$," *IEEE Transactions on Robotics and Automation*, vol. 5, no. 1, 1989.
- [2] C. C. Wang, "Extrinsic calibration of a vision sensor mounted on a robot," *IEEE Transactions on Robotics and Automation*, vol. 8, no. 2, pp. 161–175, 1992.
- [3] F. C. Park and B. J. Martin, "Robot sensor calibration: solving $\mathbf{ax}=\mathbf{xb}$ on the euclidean group," *IEEE Transactions on Robotics and Automation*, vol. 10, no. 5, 1994.
- [4] R. Horaud and F. Dornaika, "Hand-eye calibration," *The international journal of robotics research*, vol. 14, no. 3, pp. 195–210, 1995.
- [5] K. H. Strobl and G. Hirzinger, "Optimal hand-eye calibration," in *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*. IEEE, 2006, pp. 4647–4653.
- [6] M. K. Ackerman and G. S. Chirikjian, "A probabilistic solution to the $\mathbf{ax}=\mathbf{xb}$ problem: Sensor calibration without correspondence," in *Geometric Science of Information*. Springer, 2013, pp. 693–701.
- [7] A. Petrovskaya and O. Khatib, "Global localization of objects via touch," *IEEE Transactions on Robotics*, vol. 27, no. 3, pp. 569–585, 2011.
- [8] R. M. Murray, Z. Li, S. S. Sastry, and S. S. Sastry, *A mathematical introduction to robotic manipulation*. CRC press, 1994, pp. 403–433.
- [9] T. D. Barfoot and P. T. Furgale, "Associating uncertainty with three-dimensional poses for use in estimation problems," *IEEE Transactions on Robotics*, vol. 30, no. 3, pp. 679 – 693, June 2014.
- [10] R. Tsai and R. K. Lenz, "A new technique for fully autonomous and efficient 3d robotics hand/eye calibration," *IEEE Transactions on Robotics and Automation*, vol. 5, no. 3, pp. 345–358, 1989.
- [11] F. Shi, J. Wang, and Y. Liu, "An approach to improve online hand-eye calibration," in *Iberian Conference on Pattern Recognition and Image Analysis*. Springer, 2005, pp. 647–655.
- [12] J. Schmidt and H. Niemann, "Data selection for hand-eye calibration: a vector quantization approach," *The International Journal of Robotics Research*, vol. 27, no. 9, pp. 1027–1053, 2008.
- [13] M. Aron, G. Simon, and M.-O. Berger, "Handling uncertain sensor data in vision-based camera tracking," in *Mixed and Augmented Reality, 2004. ISMAR 2004. Third IEEE and ACM International Symposium on*. IEEE, 2004, pp. 58–67.
- [14] M. K. Ackerman, A. Cheng, and G. Chirikjian, "An information-theoretic approach to the correspondence-free $\mathbf{ax}=\mathbf{xb}$ sensor calibration problem," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 2014, pp. 4893–4899.
- [15] Q. Ma, H. Li, and G. S. Chirikjian, "New probabilistic approaches to the $\mathbf{ax}=\mathbf{xb}$ hand-eye calibration without correspondence," in *Robotics and Automation (ICRA), 2016 IEEE International Conference on*. IEEE, 2016, pp. 4365–4371.
- [16] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge university press, 2003, pp. 132–150.
- [17] S. F. Su and C. Lee, "Uncertainty manipulation and propagation and verification of applicability of actions in assembly tasks," in *Robotics and Automation, 1991. Proceedings., 1991 IEEE International Conference on*. IEEE, 1991, pp. 2471–2476.

- [18] P. Smith, T. Drummond, and K. Roussopoulos, "Computing map trajectories by representing, propagating and combining pdfs over groups," in *null*. IEEE, 2003, p. 1275.
- [19] Y. Wang and G. S. Chirikjian, "Error propagation on the euclidean group with applications to manipulator kinematics," *IEEE Transactions on Robotics*, vol. 22, no. 4, pp. 591–602, 2006.
- [20] —, "Nonparametric second-order theory of error propagation on motion groups," *The International journal of robotics research*, vol. 27, no. 11-12, pp. 1258–1273, 2008.
- [21] F. C. Park and B. Ravani, "Smooth invariant interpolation of rotations," *ACM Transactions on Graphics (TOG)*, vol. 16, no. 3, pp. 277–295, 1997.