

Associating Uncertainty With Three-Dimensional Poses for Use in Estimation Problems

Timothy D. Barfoot, *Member, IEEE*, and Paul T. Furgale, *Member, IEEE*

Abstract—In this paper, we provide specific and practical approaches to associate uncertainty with 4×4 transformation matrices, which is a common representation for pose variables in 3-D space. We show constraint-sensitive means of perturbing transformation matrices using their associated exponential-map generators and demonstrate these tools on three simple-yet-important estimation problems: 1) propagating uncertainty through a compound pose change, 2) fusing multiple measurements of a pose (e.g., for use in pose-graph relaxation), and 3) propagating uncertainty on poses (and landmarks) through a nonlinear camera model. The contribution of the paper is the presentation of the theoretical tools, which can be applied in the analysis of many problems involving 3-D pose and point variables.

Index Terms—Exponential maps, homogeneous points, matrix Lie groups, pose uncertainty, transformation matrices.

I. INTRODUCTION

THE main contribution of this paper is to provide simple and practical techniques to associate uncertainty with 4×4 transformation matrices (commonly used to represent pose variables) for use in robotic estimation problems. The challenge in choosing a representation of pose and its associated uncertainty is that representations either have singularities or constraints; this results from the fact that rotation variables are not *vectors* but rather members of a *noncommutative group*. The advantage of the technique that we will present is that it is free of singularities but avoids the need to enforce constraints when solving optimal estimation problems. This is accomplished by storing the mean as a (singularity-free) 4×4 transformation matrix and using a 6×1 (constraint-sensitive) perturbation of the pose (with an associated 6×6 covariance matrix). As we will see, our approach amounts to getting the best of both worlds as it facilitates rather straightforward implementations of common estimation problems such as bundle adjustment [1].

Transformation matrices are by no means new, having risen to popularity through manipulator robotics [2] and computer vision [3]. We will use standard definitions for the *special orthogonal*

Manuscript received August 8, 2013; revised November 17, 2013; accepted December 20, 2013. This paper was recommended for publication by Associate Editor R. Eustice and Editor D. Fox upon evaluation of the reviewers' comments. This work was supported by the Natural Sciences and Engineering Research Council of Canada and the EU FP7 2007–2013 Programme, Challenge 2, Cognitive Systems, Interaction, Robotics, under Grant 269916, V-Charge.

T. D. Barfoot is with the Autonomous Space Robotics Lab, University of Toronto, ON M5S 2J7, Canada (e-mail: tim.barfoot@utoronto.ca).

P. T. Furgale is with the Autonomous Systems Lab, ETH Zurich, 8092 Zurich, Switzerland (e-mail: paul.furgale@mavt.ethz.ch).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TRO.2014.2298059

group that represents rotation

$$SO(3) := \{ \mathbf{C} \in \mathbb{R}^{3 \times 3} \mid \mathbf{C}\mathbf{C}^T = \mathbf{1}, \det \mathbf{C} = 1 \}$$

where $\mathbf{1}$ is the identity matrix, and the *special Euclidean group* that represents rotation and translation

$$SE(3) := \left\{ \mathbf{T} = \begin{bmatrix} \mathbf{C} & \mathbf{r} \\ \mathbf{0}^T & 1 \end{bmatrix} \in \mathbb{R}^{4 \times 4} \mid \{\mathbf{C}, \mathbf{r}\} \in SO(3) \times \mathbb{R}^3 \right\}. \quad (1)$$

Both are examples of *matrix Lie groups*, for which Stillwell [4] provides an accessible introduction. We will avoid rehashing the basics of group theory here but stress that we cannot apply the usual approach of additive uncertainty for such quantities as they are not members of a *vector space*. In other words

$$\mathbf{x} = \bar{\mathbf{x}} + \boldsymbol{\delta} \quad (2)$$

where $\mathbf{x} \in \mathbb{R}^n$ is a random variable, $\bar{\mathbf{x}}$ is a ‘large,’ noise-free value, and $\boldsymbol{\delta}$ is a ‘small,’ uncertain perturbation (i.e., zero-mean noise), which does not directly apply to members of $SO(3)$ and $SE(3)$. This paper investigates one means to overcome this problem.

In robotics, it has been well established that estimating uncertain spatial relationships is a fundamentally important problem. For example, the early works of Brooks [5], Smith and Cheeseman [6], Durrant-Whyte [7], and Smith *et al.* [8] spurred the popular subfield of *simultaneous localization and mapping* [9]. Our approach is similar to the work of Chirikjian [10] and [11], who provide a rigorous treatment of representing and propagating uncertainty on Lie groups. It is also related to and extends the ideas presented by Su and Lee [12], [13] and Smith *et al.* [14]. Finally, in a recent paper, Hertzberg *et al.* [15] take a similar approach to associate uncertainty with general manifolds, while we provide more details on handling $SE(3)$ specifically. One could view our paper as providing an accessible and specific implementation of a means of associating and manipulating uncertainties for $SE(3)$ [and, thus, collaterally with $SO(3)$]. The differences between our approach and the previously published works are subtle but important. Rather than list them here in detail, we will take care to point out specific differences throughout the paper, then summarize them in the conclusion.

We also demonstrate the usefulness of our method on three key problems: 1) propagating uncertainty through a compounding of two transformation matrices, 2) fusing multiple uncertain measurements of a pose into a single estimate, and 3) propagating pose (and landmark) uncertainty through a nonlinear (stereo) camera model. In our calculations, we will retain as many terms in the noise variables as possible, thereby demonstrating the power of our approach.

This paper is organized as follows. Section II presents the mathematical machinery that we will use throughout the paper. Section III discusses how to compound two poses and their uncertainties, Section IV shows how to fuse multiple estimates of pose together, and Section V shows how to propagate pose and landmark uncertainty through a nonlinear measurement model. A conclusion rounds out the paper.

II. MATHEMATICAL PRELIMINARIES

A. Random Variables and Probability Density Functions

In contrast with (2), we will define random variables for $SE(3)$ according to

$$\mathbf{T} := \exp(\boldsymbol{\xi}^\wedge) \bar{\mathbf{T}} \quad (3)$$

where $\bar{\mathbf{T}} \in SE(3)$ is a ‘large,’ noise-free value, and $\boldsymbol{\xi} \in \mathbb{R}^6$ is a ‘small,’ noisy perturbation. The \wedge operator¹ turns $\boldsymbol{\xi}$ into a 4×4 member of the *Lie algebra* $\mathfrak{se}(3)$ according to

$$\boldsymbol{\xi}^\wedge := \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{\phi} \end{bmatrix}^\wedge = \begin{bmatrix} \boldsymbol{\phi}^\wedge & \boldsymbol{\rho}^\wedge \\ \mathbf{0}^T & 0 \end{bmatrix} \quad (4)$$

where $\boldsymbol{\rho}, \boldsymbol{\phi} \in \mathbb{R}^3$ are 3×1 , and \wedge also turns $\boldsymbol{\phi}$ into a member of the *Lie algebra* $\mathfrak{so}(3)$:

$$\boldsymbol{\phi}^\wedge := \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix}^\wedge = \begin{bmatrix} 0 & -\phi_3 & \phi_2 \\ \phi_3 & 0 & -\phi_1 \\ -\phi_2 & \phi_1 & 0 \end{bmatrix}. \quad (5)$$

We will use \vee as the inverse operation of \wedge . This perturbation approach ensures that $\mathbf{T} \in SE(3)$ and is sometimes referred to as an *injection* of noise onto the group $SE(3)$ [20]. Additional notation and a closed-form expression for the matrix exponential in (3) can be found in the Appendix.

In this paper, we will define the small perturbation variable to be zero-mean Gaussian $p(\boldsymbol{\xi}) = \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is a 6×6 covariance matrix. We stress that we are directly defining this probability density function (PDF) in the vectorspace, \mathbb{R}^6 . This choice in turn induces a PDF over $SE(3)$, $p(\mathbf{T})$. To see what this PDF is, we note that by our definition

$$\int_{\mathbb{R}^6} p(\boldsymbol{\xi}) d\boldsymbol{\xi} = \int_{\mathbb{R}^6} \eta \exp\left(-\frac{1}{2}\boldsymbol{\xi}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\xi}\right) d\boldsymbol{\xi} = 1 \quad (6)$$

where $\eta = \frac{1}{\sqrt{(2\pi)^6 \det(\boldsymbol{\Sigma})}}$ is the (constant) normalization factor. To switch variables and write a PDF over \mathbf{T} , we must define what we mean by an infinitesimal volume of $SE(3)$, which is written as $d\mathbf{T}$. Along the lines of Chirikjian [11], we will relate perturbations in $\boldsymbol{\xi}$ to perturbations in \mathbf{T} according to

$$d\mathbf{T} := \ln(\mathbf{T}' \mathbf{T}^{-1})^\vee \quad (7)$$

¹The \wedge operator is ‘overloaded’ in the sense that it can be applied to both 6×1 and 3×1 columns [2] [11] [16]. For the 6×1 case, \wedge is similar to the \boxplus operator defined by [17] but without the negative sign in front of the $\boldsymbol{\rho}$ component. For the 3×1 case, $\boldsymbol{\phi}^\wedge$ is sometimes written as $\boldsymbol{\phi}^\times$ [18] or $[[\boldsymbol{\phi}]]$ [19]; sometimes, the \wedge is written above the symbol, as in $\hat{\boldsymbol{\phi}}$; however, we shall stick to the original definition in Murray *et al.* [2, pp. 26 and 41], which has it off to the right, since we often apply it to expressions involving multiple variables.

where

$$\mathbf{T}' := \exp((\boldsymbol{\xi} + \delta\boldsymbol{\xi})^\wedge) \bar{\mathbf{T}}, \quad \mathbf{T} := \exp(\boldsymbol{\xi}^\wedge) \bar{\mathbf{T}}. \quad (8)$$

It can then be shown that

$$\delta\mathbf{T} = \mathcal{J} \delta\boldsymbol{\xi} \quad (9)$$

where

$$\mathcal{J} := \int_0^1 \exp(\boldsymbol{\xi}^\wedge)^\alpha d\alpha \quad (10)$$

is referred to as the *Jacobian* for $SE(3)$ [21]; we will discuss this in more detail in the next section. We define the \wedge operator² to be

$$\boldsymbol{\xi}^\wedge = \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{\phi} \end{bmatrix}^\wedge = \begin{bmatrix} \boldsymbol{\phi}^\wedge & \boldsymbol{\rho}^\wedge \\ \mathbf{0} & \boldsymbol{\phi}^\wedge \end{bmatrix} \quad (12)$$

which is 6×6 . We will also use \vee as the inverse operation of \wedge , for convenience.

The infinitesimal volume elements can then be related [21] by

$$d\mathbf{T} = |\det(\mathcal{J})| d\boldsymbol{\xi}. \quad (13)$$

Using $\boldsymbol{\xi} = \ln(\mathbf{T}' \bar{\mathbf{T}}^{-1})^\vee$ and (13) in (6), we may find the PDF over \mathbf{T} by a straightforward change of coordinates³

$$\begin{aligned} 1 &= \int_{\mathbb{R}^6} \eta \exp\left(-\frac{1}{2}\boldsymbol{\xi}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\xi}\right) d\boldsymbol{\xi} \\ &= \int_{SE(3)} \underbrace{\beta \exp\left(-\frac{1}{2} \ln(\mathbf{T}' \bar{\mathbf{T}}^{-1})^{T^V} \boldsymbol{\Sigma}^{-1} \ln(\mathbf{T}' \bar{\mathbf{T}}^{-1})^V\right)}_{p(\mathbf{T})} d\mathbf{T} \\ &= \int_{SE(3)} p(\mathbf{T}) d\mathbf{T} \end{aligned} \quad (14)$$

where

$$\beta = \frac{\eta}{|\det(\mathcal{J})|} = \frac{1}{\sqrt{(2\pi)^6 \det(\mathcal{J} \boldsymbol{\Sigma} \mathcal{J}^T)}}. \quad (15)$$

Importantly, $p(\mathbf{T})$ is not Gaussian because of the fact that β depends on \mathbf{T} via \mathcal{J} . Again, we stress that our $p(\mathbf{T})$ is only indirectly defined via $p(\boldsymbol{\xi})$.

Our approach to define uncertainty on $SE(3)$ via $\boldsymbol{\xi}$ is very similar to that introduced by Su and Lee [12] and [13], although we apply the perturbation in (3) on the left rather than the right (both are valid). However, our paper works out the consequences

²Our \wedge operator is similarly defined as \times by D’Eleuterio [22] and [23] for use in manipulator kinematics/dynamics. In relation to the other notation in the literature [11], we have

$$\text{Ad}(\exp(\boldsymbol{\xi}^\wedge)) = \exp(\boldsymbol{\xi}^\wedge) = \exp(\text{ad}(\boldsymbol{\xi}^\wedge)) \quad (11)$$

where $\text{Ad}(\cdot)$ and $\text{ad}(\cdot)$ are the *adjoints* for $SE(3)$ and $\mathfrak{se}(3)$, respectively. Thus, by introducing \wedge , we have a means to compute both adjoints.

³Technically speaking, to carry out this change of coordinates rigorously, we would like the integration limits on the $\boldsymbol{\phi}$ part of $\boldsymbol{\xi}$ to be such that $\|\boldsymbol{\phi}\| < \pi$. However, a Gaussian has infinite tails, and therefore, we are effectively allowing the tails to ‘fold’ around by integrating over all of \mathbb{R}^6 . In most practical situations, the Gaussian is not too spread out, and there is very little probability mass beyond the π limit. When a global representation of uncertainty on $SO(3)$ is required, we may turn to the approach of Lee *et al.* [24].

of this parameterization further (and to higher accuracy) than Su and Lee, particularly the compounding and fusing of uncertain transformations.

Our approach is also similar but subtly different from the approach of Chirikjian [11], Wang and Chirikjian [25], [26], and Wolfe *et al.* [27]. These works start by defining $p(\mathbf{T})$ directly and switch to parameterizing \mathbf{T} via exponential coordinates to carry out computations. In terms of our notation, they define a Gaussian PDF over \mathbf{T} as

$$p'(\mathbf{T}) = \beta' \exp \left(-\frac{1}{2} \ln (\mathbf{T}\bar{\mathbf{T}}^{-1})^{\vee T} \Sigma^{-1} \ln (\mathbf{T}\bar{\mathbf{T}}^{-1})^{\vee} \right) \quad (16)$$

where β' is a (constant) normalization factor chosen to make $p'(\mathbf{T})$ a valid PDF [27, eq. 14]. By contrast, we define $p(\xi)$ first and induce $p(\mathbf{T})$ through (3). The difference between these works and ours is essentially between β' (a constant) and β (dependent on \mathbf{T}); both are in fact valid PDFs, but we believe our approach has merit in terms of its relative ease of computation.

In practice, we simply work with $p(\xi)$ and never need to compute $p(\mathbf{T})$. For example, if we want to compute the expected value $E[\cdot]$ of some nonlinear function of a transformation $f(\mathbf{T})$, we can simply use $p(\xi)$ as follows:

$$\begin{aligned} E[f(\mathbf{T})] &= \int_{SE(3)} f(\mathbf{T}) p(\mathbf{T}) d\mathbf{T} \\ &= \int_{SE(3)} f(\mathbf{T}) \underbrace{\beta \exp \left(-\frac{1}{2} \ln (\mathbf{T}\bar{\mathbf{T}}^{-1})^{\vee T} \Sigma^{-1} \ln (\mathbf{T}\bar{\mathbf{T}}^{-1})^{\vee} \right)}_{p(\mathbf{T})} d\mathbf{T} \\ &= \int_{\mathbb{R}^6} f(\exp(\xi^\wedge)\bar{\mathbf{T}}) \underbrace{\beta |\det(\mathcal{J})| \exp \left(-\frac{1}{2} \xi^T \Sigma^{-1} \xi \right)}_{p(\xi)} d\xi \\ &= \int_{\mathbb{R}^6} f(\exp(\xi^\wedge)\bar{\mathbf{T}}) p(\xi) d\xi. \end{aligned} \quad (17)$$

We underscore that we do not require any approximation to write down this last expression, which does not contain $|\det(\mathcal{J})|$, and this is because of our choice to define $p(\xi)$ directly. However, $f(\cdot)$ is still a nonlinearity, and therefore, by expanding the matrix exponential, and performing a Taylor-series expansion in the ξ noise variables, we can easily compute approximations of the expectation to as many terms as we like.

In terms of the statistical moments of $p(\mathbf{T})$, our approach would seem to be the more natural choice, as we now show. One common (left) definition of the *mean* of $p(\mathbf{T})$ [25] is the unique $\mathbf{M} \in SE(3)$ that satisfies the following equation:

$$\int_{SE(3)} \ln (\mathbf{T}\mathbf{M}^{-1})^{\vee} p(\mathbf{T}) d\mathbf{T} = \mathbf{0}. \quad (18)$$

It is easy to see, under our version of $p(\mathbf{T})$, that in fact, $\mathbf{M} = \bar{\mathbf{T}}$ since

$$\int_{SE(3)} \ln (\mathbf{T}\bar{\mathbf{T}}^{-1})^{\vee} p(\mathbf{T}) d\mathbf{T} = \int_{\mathbb{R}^6} \xi p(\xi) d\xi = \mathbf{0}. \quad (19)$$

As pointed out by Wolfe *et al.* [27], using the other PDF definition, $p'(\mathbf{T})$, it is not necessarily true that $\mathbf{M}' = \bar{\mathbf{T}}$, where

$$\int_{SE(3)} \ln (\mathbf{T}\mathbf{M}'^{-1})^{\vee} p'(\mathbf{T}) d\mathbf{T} = \mathbf{0}. \quad (20)$$

Likewise, we see that our Σ matches one common (left) definition of *covariance* for $SE(3)$ [25]

$$\begin{aligned} \Sigma &:= \int_{SE(3)} \ln (\mathbf{T}\mathbf{M}^{-1})^{\vee} \ln (\mathbf{T}\mathbf{M}^{-1})^{\vee T} p(\mathbf{T}) d\mathbf{T} \\ &= \int_{\mathbb{R}^6} \xi \xi^T p(\xi) d\xi. \end{aligned} \quad (21)$$

Using the other PDF definition, $p'(\mathbf{T})$, we let

$$\Sigma' := \int_{SE(3)} \ln (\mathbf{T}\mathbf{M}'^{-1})^{\vee} \ln (\mathbf{T}\mathbf{M}'^{-1})^{\vee T} p'(\mathbf{T}) d\mathbf{T} \quad (22)$$

and then it is not necessarily true that $\Sigma' = \Sigma$. As discussed by Wolfe *et al.* [27], as Σ gets very small, we have that

$$\mathbf{M}' \approx \mathbf{M} = \bar{\mathbf{T}}, \quad \Sigma' \approx \Sigma \quad (23)$$

because in this case, $\mathcal{J} \approx \mathbf{1}$, and therefore, $\beta' \approx \beta$.

There is another reason why our approach to inject uncertainty into $SE(3)$ may be the more natural choice for robotics applications. Consider the equation for the continuous-time kinematics of a rigid body

$$\dot{\mathbf{T}} = \boldsymbol{\varpi}^\wedge \mathbf{T} \quad (24)$$

where $\boldsymbol{\varpi} := [\begin{smallmatrix} \nu \\ \omega \end{smallmatrix}]$ is the combined linear and angular velocity. If we assume that $\boldsymbol{\varpi}$ is constant from t_1 to t_2 , then the discrete-time update is

$$\mathbf{T}(t_2) = \exp(t\boldsymbol{\varpi}^\wedge) \mathbf{T}(t_1) \quad (25)$$

where $t := t_2 - t_1$. If we are dead-reckoning from the measured velocity, then in recursive estimation schemes such as the extended Kalman filter, we typically assume that $\boldsymbol{\varpi}$ is corrupted by Gaussian noise. This is precisely what our perturbation scheme in (3) assumes.

Finally, a deterministic change of coordinates, via transformation $\mathbf{T}_d \in SE(3)$, is straightforward under our parameterization [12], since

$$\mathbf{T}_d \mathbf{T} = \mathbf{T}_d \exp(\xi^\wedge) \bar{\mathbf{T}} = \exp((\mathcal{T}_d \xi)^\wedge) \mathbf{T}_d \bar{\mathbf{T}}, \quad (26)$$

where $\mathcal{T}_d = \text{Ad}(\mathbf{T}_d)$. The new mean is therefore $\mathbf{T}_d \bar{\mathbf{T}}$ and the new covariance is $\mathcal{T}_d \Sigma \mathcal{T}_d^T$. Thus, we do not require any approximation to apply a deterministic change of coordinates to our chosen PDF. We will further discuss the compounding of two uncertain transformations in Section III.

B. Baker–Campbell–Hausdorff Formula

One of the main tools that we will use to manipulate uncertainties on $SE(3)$ is the *Baker–Campbell–Hausdorff* (BCH) formula, which can be used to compound two matrix

exponentials

$$\begin{aligned} \ln(\exp(\mathbf{A})\exp(\mathbf{B})) &= \mathbf{A} + \mathbf{B} + \frac{1}{2}[\mathbf{A}, \mathbf{B}] + \frac{1}{12}[\mathbf{A}, [\mathbf{A}, \mathbf{B}]] \\ &\quad + \frac{1}{12}[\mathbf{B}, [\mathbf{B}, \mathbf{A}]] - \frac{1}{24}[\mathbf{B}, [\mathbf{A}, [\mathbf{A}, \mathbf{B}]]] + \dots \end{aligned} \quad (27)$$

where the *Lie bracket* is given by

$$[\mathbf{A}, \mathbf{B}] := \mathbf{AB} - \mathbf{BA}. \quad (28)$$

Note, the BCH formula is an infinite series. If we keep only terms linear in \mathbf{A} , it becomes [28]

$$\ln(\exp(\mathbf{A})\exp(\mathbf{B})) \approx \mathbf{B} + \sum_{n=0}^{\infty} \frac{B_n}{n!} \underbrace{[\mathbf{B}, [\mathbf{B}, \dots [\mathbf{B}, \mathbf{A}]] \dots]]}_{n}. \quad (29)$$

The B_n are the *Bernoulli numbers*

$$B_0 = 1, B_1 = -\frac{1}{2}, B_2 = \frac{1}{6}, B_3 = 0, B_4 = -\frac{1}{30}, B_5 = 0, \dots, \quad (30)$$

which appear frequently in number theory.

In the particular case of $SE(3)$, we can use the fact that

$$[\xi_a^\wedge, \xi_b^\wedge] = \xi_a^\wedge \xi_b^\wedge - \xi_b^\wedge \xi_a^\wedge = (\xi_a^\wedge \xi_b^\wedge)^\wedge \quad (31)$$

to show that

$$\begin{aligned} \ln(\mathbf{T}_a \mathbf{T}_b)^\vee &= \ln(\exp(\xi_a^\wedge) \exp(\xi_b^\wedge))^\vee = \xi_a + \xi_b + \frac{1}{2}\xi_a^\wedge \xi_b \\ &\quad + \frac{1}{12}\xi_a^\wedge \xi_a^\wedge \xi_b + \frac{1}{12}\xi_b^\wedge \xi_b^\wedge \xi_a - \frac{1}{24}\xi_b^\wedge \xi_a^\wedge \xi_a^\wedge \xi_b + \dots. \end{aligned} \quad (32)$$

Alternatively, if we assume that ξ_a is small, then using the approximate formula stated previously, we can show that

$$\ln(\mathbf{T}_a \mathbf{T}_b)^\vee = \ln(\exp(\xi_a^\wedge) \exp(\xi_b^\wedge))^\vee \approx \xi_b + \mathcal{J}_b^{-1} \xi_a \quad (33)$$

where

$$\mathcal{J}_b^{-1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} (\xi_b^\wedge)^n. \quad (34)$$

Inverting, we find

$$\mathcal{J}_b = \sum_{n=0}^{\infty} \frac{1}{(n+1)!} (\xi_b^\wedge)^n = \int_0^1 \exp(\xi_b^\wedge)^\alpha d\alpha. \quad (35)$$

In Lie group theory, \mathcal{J}_b is referred to as the left⁴ *Jacobian* of $SE(3)$. We saw this quantity in action in the previous section already; however, it is useful to understand that it comes directly from the BCH formula under the assumption that ξ_a is small. The appendix contains a closed-form expression to compute \mathcal{J}_b from ξ_b .

III. COMPOUNDING POSES

In this section, we investigate the problem of compounding two poses, each with associated uncertainty, as depicted in Fig. 1.

⁴There is also a right Jacobian if we instead assume ξ_b is small.

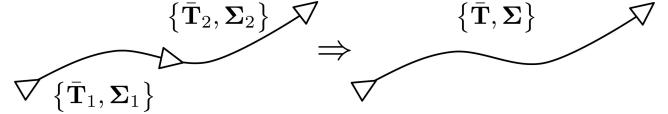


Fig. 1. Combining a chain of two poses into a single compound pose.

A. Theory

Consider two noisy poses \mathbf{T}_1 and \mathbf{T}_2 ; we keep track of their nominal values and associated uncertainties

$$\{\bar{\mathbf{T}}_1, \Sigma_1\}, \quad \{\bar{\mathbf{T}}_2, \Sigma_2\}.$$

Now, suppose, we let

$$\mathbf{T} = \mathbf{T}_1 \mathbf{T}_2$$

as depicted in Fig. 1. What is $\{\bar{\mathbf{T}}, \Sigma\}$? Under our perturbation scheme, we have

$$\exp(\xi^\wedge) \bar{\mathbf{T}} = \exp(\xi_1^\wedge) \bar{\mathbf{T}}_1 \exp(\xi_2^\wedge) \bar{\mathbf{T}}_2. \quad (36)$$

Moving all the uncertain factors to the left side, we have

$$\exp(\xi^\wedge) \bar{\mathbf{T}} = \exp(\xi_1^\wedge) \exp((\bar{\mathbf{T}}_1 \xi_2)^\wedge) \bar{\mathbf{T}}_1 \bar{\mathbf{T}}_2 \quad (37)$$

where $\bar{\mathbf{T}}_1 = \text{Ad}(\bar{\mathbf{T}}_1)$. If we let

$$\bar{\mathbf{T}} = \bar{\mathbf{T}}_1 \bar{\mathbf{T}}_2 \quad (38)$$

we are left with

$$\exp(\xi^\wedge) = \exp(\xi_1^\wedge) \exp((\bar{\mathbf{T}}_1 \xi_2)^\wedge). \quad (39)$$

Letting $\xi'_2 := \bar{\mathbf{T}}_1 \xi_2$, we can apply the BCH formula to find

$$\begin{aligned} \xi &= \xi_1 + \xi'_2 + \frac{1}{2}\xi_1^\wedge \xi'_2 + \frac{1}{12}\xi_1^\wedge \xi_1^\wedge \xi'_2 + \frac{1}{12}\xi_2^\wedge \xi'_2^\wedge \xi_1 \\ &\quad - \frac{1}{24}\xi'_2^\wedge \xi_1^\wedge \xi_1^\wedge \xi'_2 + \dots. \end{aligned} \quad (40)$$

For our approach to hold, we require that $E[\xi] = 0$. Assuming that $\xi_1 \sim \mathcal{N}(\mathbf{0}, \Sigma_1)$ and $\xi'_2 \sim \mathcal{N}(\mathbf{0}, \Sigma'_2)$ are uncorrelated with one another, we have

$$E[\xi] = -\frac{1}{24}E[\xi'_2^\wedge \xi_1^\wedge \xi_1^\wedge \xi'_2] + O(\xi^6) \quad (41)$$

since everything except the fourth-order term has zero mean. Thus, to third order, we can safely assume that $E[\xi] = \mathbf{0}$, and thus, (38) seems to be a reasonable way to compound the mean transformations.

It is also possible to show that the fourth-order term has zero mean, $E[\xi'_2^\wedge \xi_1^\wedge \xi_1^\wedge \xi'_2] = \mathbf{0}$, if Σ_1 is of the special form

$$\Sigma_1 = E[\xi_1 \xi_1^T] = \begin{bmatrix} \Sigma_{1,\rho\rho} & \mathbf{0} \\ \mathbf{0} & \sigma_{1,\phi\phi}^2 \mathbf{1} \end{bmatrix} \quad (42)$$

where $\xi_1 = [\rho_1 \phi_1]$. This is a common situation for Σ_1 when we are, for example, propagating uncertainty on velocity, as in (25). In this case, we are justified to assume $E[\xi] = \mathbf{0}$ all the way out to fifth order (and possibly further).

The next task is to compute $\Sigma = E [\xi \xi^T]$. Multiplying out to fourth order, we have

$$\begin{aligned} E [\xi \xi^T] &\approx E \left[\xi_1 \xi_1^T + \xi'_2 \xi_2^T + \frac{1}{12} \left((\xi_1^\wedge \xi_1^\wedge) (\xi'_2 \xi_2'^T) \right. \right. \\ &+ \left(\xi'_2 \xi_2'^T \right) (\xi_1^\wedge \xi_1^\wedge)^T + \left(\xi_2'^\wedge \xi_2^\wedge \right) (\xi_1 \xi_1^T) \\ &\left. \left. + \left(\xi_2'^\wedge \xi_2^\wedge \right) (\xi_1 \xi_1^T) \right) + \frac{1}{4} \xi_1^\wedge \left(\xi'_2 \xi_2'^T \right) \xi_1^{\wedge T} \right] \quad (43) \end{aligned}$$

where we have omitted showing any terms that have an odd power in either ξ_1 or ξ'_2 since these will by definition have expectation zero. This expression may look daunting, but we can take it term by term. To save space, we define and make use of the following two linear operators:

$$\langle\langle \mathbf{A} \rangle\rangle := -\text{tr}(\mathbf{A}) \mathbf{1} + \mathbf{A} \quad (44)$$

$$\langle\langle \mathbf{A}, \mathbf{B} \rangle\rangle := \langle\langle \mathbf{A} \rangle\rangle \langle\langle \mathbf{B} \rangle\rangle + \langle\langle \mathbf{B} \mathbf{A} \rangle\rangle \quad (45)$$

with $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$. These provide the useful identity

$$-\mathbf{u}^\wedge \mathbf{A} \mathbf{v}^\wedge \equiv \langle\langle \mathbf{v} \mathbf{u}^T, \mathbf{A}^T \rangle\rangle \quad (46)$$

where $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ and $\mathbf{A} \in \mathbb{R}^{3 \times 3}$. Making use of this repeatedly, we have out to fourth order

$$E [\xi_1 \xi_1^T] = \Sigma_1 = \begin{bmatrix} \Sigma_{1,\rho\rho} & \Sigma_{1,\rho\phi} \\ \Sigma_{1,\rho\phi}^T & \Sigma_{1,\phi\phi} \end{bmatrix} \quad (47)$$

$$E [\xi'_2 \xi_2'^T] = \Sigma'_2 = \begin{bmatrix} \Sigma'_{2,\rho\rho} & \Sigma'_{2,\rho\phi} \\ \Sigma'_{2,\rho\phi}^T & \Sigma'_{2,\phi\phi} \end{bmatrix} = \bar{\mathbf{T}}_1 \Sigma_2 \bar{\mathbf{T}}_1^T \quad (48)$$

$$E [\xi_1^\wedge \xi_1^\wedge] = \mathcal{A}_1 = \begin{bmatrix} \langle\langle \Sigma_{1,\phi\phi} \rangle\rangle & \langle\langle \Sigma_{1,\rho\phi} + \Sigma_{1,\rho\phi}^T \rangle\rangle \\ \mathbf{0} & \langle\langle \Sigma_{1,\phi\phi} \rangle\rangle \end{bmatrix} \quad (49)$$

$$E [\xi_2'^\wedge \xi_2^\wedge] = \mathcal{A}'_2 = \begin{bmatrix} \langle\langle \Sigma'_{2,\phi\phi} \rangle\rangle & \langle\langle \Sigma'_{2,\rho\phi} + \Sigma'_{2,\rho\phi}^T \rangle\rangle \\ \mathbf{0} & \langle\langle \Sigma'_{2,\phi\phi} \rangle\rangle \end{bmatrix} \quad (50)$$

$$E [\xi_1^\wedge (\xi'_2 \xi_2'^T) \xi_1^{\wedge T}] = \mathcal{B} = \begin{bmatrix} \mathbf{B}_{\rho\rho} & \mathbf{B}_{\rho\phi} \\ \mathbf{B}_{\rho\phi}^T & \mathbf{B}_{\phi\phi} \end{bmatrix} \quad (51)$$

where

$$\begin{aligned} \mathbf{B}_{\rho\rho} &= \langle\langle \Sigma_{1,\phi\phi}, \Sigma'_{2,\rho\phi} \rangle\rangle + \langle\langle \Sigma_{1,\rho\phi}^T, \Sigma'_{2,\rho\phi} \rangle\rangle \\ &+ \langle\langle \Sigma_{1,\rho\phi}, \Sigma'_{2,\rho\phi}^T \rangle\rangle + \langle\langle \Sigma_{1,\rho\phi}, \Sigma'_{2,\phi\phi} \rangle\rangle \quad (52) \end{aligned}$$

$$\mathbf{B}_{\rho\phi} = \langle\langle \Sigma_{1,\phi\phi}, \Sigma'_{2,\rho\phi}^T \rangle\rangle + \langle\langle \Sigma_{1,\rho\phi}^T, \Sigma'_{2,\phi\phi} \rangle\rangle \quad (53)$$

$$\mathbf{B}_{\phi\phi} = \langle\langle \Sigma_{1,\phi\phi}, \Sigma'_{2,\phi\phi} \rangle\rangle. \quad (54)$$

The resulting covariance is then

$$\begin{aligned} \Sigma_{4\text{th}} &\approx \underbrace{\Sigma_1 + \Sigma'_2}_{\Sigma_{2\text{nd}}} \\ &+ \underbrace{\frac{1}{12} \left(\mathcal{A}_1 \Sigma'_2 + \Sigma'_2 \mathcal{A}_1^T + \mathcal{A}'_2 \Sigma_1 + \Sigma_1 \mathcal{A}'_2^T \right)}_{\text{additional fourth-order terms}} + \frac{1}{4} \mathcal{B} \quad (55) \end{aligned}$$

correct to fourth order.⁵ This result is essentially the same as that of Wang and Chirikjian [26] but worked out for our slightly

⁵The sixth-order terms require a lot more work, but it is possible to compute them using Isserlis' theorem.

different PDF, as discussed earlier; it is important to note that, while our method is fourth order in the perturbation variables, it is only second order in the covariance, which is the same as [26]. In summary, to compound two poses, we propagate the mean using (38) and the covariance using (55).

B. Sigma-point Method

We can also make use of the popular *sigmapoint transformation* [29] to pass uncertainty through the compound pose change. In this section, we tailor this to our specific type of $SE(3)$ perturbation. Our approach to handling sigmapoints is quite similar to that taken by Hertzberg *et al.* [15] as well as Brookshire and Teller [30]. In our case, we begin by approximating the joint input Gaussian using a finite number of samples, $\{\mathbf{T}_{1,\ell}, \mathbf{T}_{2,\ell}\}$

$$\mathbf{S}\mathbf{S}^T := \text{diag}\{\Sigma_1, \Sigma_2\}, \quad (\text{Cholesky decomposition})$$

$$\boldsymbol{\psi}_\ell := \sqrt{\lambda} \text{col}_\ell \mathbf{S}, \quad \ell = 1 \dots L$$

$$\boldsymbol{\psi}_{\ell+L} := -\sqrt{\lambda} \text{col}_\ell \mathbf{S}, \quad \ell = 1 \dots L$$

$$\begin{bmatrix} \xi_{1,\ell} \\ \xi_{2,\ell} \end{bmatrix} := \boldsymbol{\psi}_\ell, \quad \ell = 1 \dots 2L$$

$$\mathbf{T}_{1,\ell} := \exp(\xi_{1,\ell}^\wedge) \bar{\mathbf{T}}_1, \quad \ell = 1 \dots 2L$$

$$\mathbf{T}_{2,\ell} := \exp(\xi_{2,\ell}^\wedge) \bar{\mathbf{T}}_2, \quad \ell = 1 \dots 2L$$

where λ is a user-definable scaling constant,⁶ and $L = 12$. We then pass each of these samples through the compound pose change and compute the difference from the mean

$$\boldsymbol{\xi}_\ell = \ln(\mathbf{T}_{1,\ell} \mathbf{T}_{2,\ell} \bar{\mathbf{T}}^{-1}), \quad \ell = 1 \dots 2L. \quad (56)$$

These are combined to create the output covariance according to

$$\Sigma_{\text{sp}} = \frac{1}{2\lambda} \sum_{\ell=1}^{2L} \boldsymbol{\xi}_\ell \boldsymbol{\xi}_\ell^T. \quad (57)$$

Note that we have assumed that the output sigmapoint samples have zero mean in this formula, to be consistent with our mean propagation. Interestingly, this turns out to be algebraically equivalent to the second-order method (from the previous section) for this particular nonlinearity, since the noise sources on \mathbf{T}_1 and \mathbf{T}_2 are assumed to be uncorrelated.

C. Simple Compound Example

In this section, we present a simple qualitative example of pose compounding and in Section III-D, we carry out a more quantitative study on a different setup. To see the qualitative difference between the second- and fourth-order methods, let us consider the case of compounding transformations many times in a row

$$\exp(\xi_K^\wedge) \bar{\mathbf{T}}_K = \left(\prod_{k=1}^K \exp(\xi^\wedge) \bar{\mathbf{T}} \right) \exp(\xi_0^\wedge) \bar{\mathbf{T}}_0. \quad (58)$$

⁶For all experiments in this paper, we used $\lambda = 1$; we need to ensure that the sigmapoints associated with the rotational degrees of freedom have length less than π to avoid numerical problems.

As discussed earlier, this can be viewed as a discrete-time integration of the $SE(3)$ kinematic equations, as in (25). To keep things simple, we make the following assumptions:

$$\bar{\mathbf{T}}_0 = \mathbf{1}, \quad \xi_0 \sim \mathcal{N}(\mathbf{0}, \mathbf{0}) \quad (59)$$

$$\bar{\mathbf{T}} = \begin{bmatrix} \bar{\mathbf{C}} & \bar{\mathbf{r}} \\ \mathbf{0}^T & 1 \end{bmatrix}, \quad \xi \sim \mathcal{N}(\mathbf{0}, \Sigma) \quad (60)$$

$$\bar{\mathbf{C}} = \mathbf{1}, \quad \bar{\mathbf{r}} = \begin{bmatrix} r \\ 0 \\ 0 \end{bmatrix}, \quad \Sigma = \text{diag}(0, 0, 0, 0, 0, \sigma^2). \quad (61)$$

Although this example uses our 3-D tools, it is confined to a plane for the purpose of illustration and ease of plotting; it corresponds to a rigid body moving along the x -axis but with some uncertainty only on the rotational velocity about the z -axis. This could model a unicycle robot driving in the plane with a constant translational speed and slightly uncertain rotational speed (centered about zero). We are interested in how the covariance matrix fills in over time.

According to the second-order scheme, we have

$$\bar{\mathbf{T}}_K = \begin{bmatrix} 1 & 0 & 0 & Kr \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (62)$$

$$\Sigma_K = \begin{bmatrix} 0 & \frac{K(K-1)(2K-1)}{6}r^2\sigma^2 & 0 & 0 & 0 & -\frac{K(K-1)}{2}r\sigma^2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{K(K-1)}{2}r\sigma^2 & 0 & 0 & 0 & K\sigma^2 \end{bmatrix} \quad (63)$$

where we see that the top-left entry of Σ_K , corresponding to uncertainty in the x -direction, does not have any growth of uncertainty. However, in the fourth-order scheme, the fill-in pattern is such that the top-left entry is nonzero. This happens for several reasons, but mainly through the $\mathbf{B}_{\rho\rho}$ submatrix of \mathbf{B} . This leaking of uncertainty into an additional degree of freedom cannot be captured by keeping only the second-order terms. Fig. 2 provides a numerical example of this effect. It shows that both the second- and fourth-order schemes do a good job of representing the ‘banana’-like distribution over poses, as discussed by Long *et al.* [31]. However, the fourth-order scheme has some finite uncertainty in the straight-ahead direction (as do the sampled trajectories), while the second-order scheme does not.

D. Compound Experiment

To quantitatively evaluate the pose-compounding techniques, we ran a second numerical experiment in which we compounded two poses including their associated covariance matrices

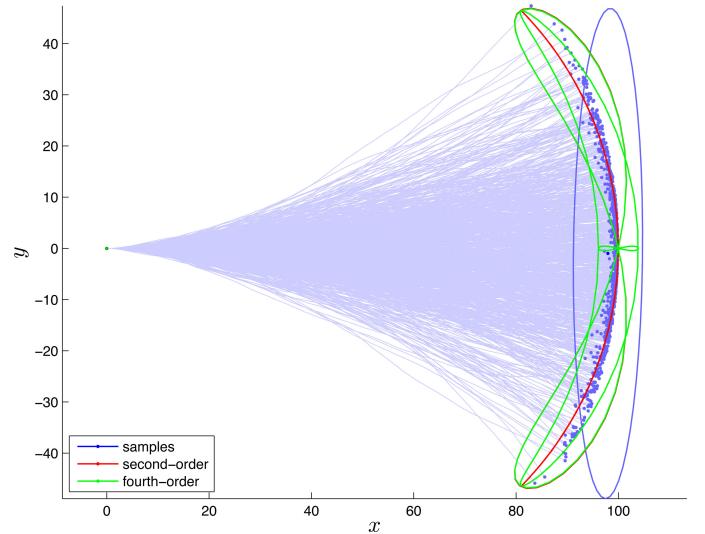


Fig. 2. Example of compounding $K = 100$ uncertain transformations (see Section III-C). The light blue lines and blue dots show 1000 individual sampled trajectories starting from $(0, 0)$ and moving nominally to the right at constant translational speed but with some uncertainty on the rotational velocity. The blue 1-sigma covariance ellipse is simply fitted to the blue samples to show what keeping xy -covariance relative to the start looks like. The red (second order) and green (fourth order) lines are the principal great circles of the 1-sigma covariance ellipsoid, given by Σ_K , mapped to the xy plane. Looking at the area $(95, 0)$, corresponding to straight ahead, the fourth-order scheme has some nonzero uncertainty (as do the samples), whereas the second-order scheme does not. We used $r = 1$ and $\sigma = 0.03$.

$$\begin{aligned} \bar{\mathbf{T}}_1 &:= \exp(\bar{\xi}_1^\wedge), \quad \bar{\xi}_1 := [0 \ 2 \ 0 \ \pi/6 \ 0 \ 0]^T \\ \Sigma_1 &:= \alpha \times \text{diag}\left\{10, 5, 5, \frac{1}{2}, 1, \frac{1}{2}\right\} \\ \bar{\mathbf{T}}_2 &:= \exp(\bar{\xi}_2^\wedge), \quad \bar{\xi}_2 := [0 \ 0 \ 1 \ 0 \ \pi/4 \ 0]^T \\ \Sigma_2 &:= \alpha \times \text{diag}\left\{5, 10, 5, \frac{1}{2}, \frac{1}{2}, 1\right\} \end{aligned}$$

where $\alpha \in [0, 1]$ is a scaling parameter that increases the magnitude of the input covariances parametrically.

We compounded these two poses according to (36), which results in a mean of $\bar{\mathbf{T}} = \bar{\mathbf{T}}_1 \bar{\mathbf{T}}_2$. The covariance Σ was computed using four methods.

- 1) *Monte Carlo*: We drew a large number, $M = 1\,000\,000$, of random samples $(\xi_{m_1}^\wedge, \xi_{m_2}^\wedge)$ from the input covariance matrices, compounded the resulting transformations, and computed the covariance as $\Sigma_{mc} = \frac{1}{M} \sum_{m=1}^M \xi_{m_1} \xi_{m_2}^T$ with $\bar{\mathbf{T}}_m = \exp(\xi_{m_1}^\wedge) \bar{\mathbf{T}}_1 \exp(\xi_{m_2}^\wedge) \bar{\mathbf{T}}_2$, and $\xi_m = \ln(\bar{\mathbf{T}}_m \bar{\mathbf{T}}^{-1})^\vee$. This slow-but-accurate approach served as our benchmark to which the other three much faster methods were compared.
- 2) *Second Order*: We used the second-order method described previously to compute Σ_{2nd} .
- 3) *Fourth Order*: We used the fourth-order method described previously to compute Σ_{4th} .
- 4) *Sigmapoint*: We used the *sigmapoint* transformation described previously [29] to compute Σ_{sp} .

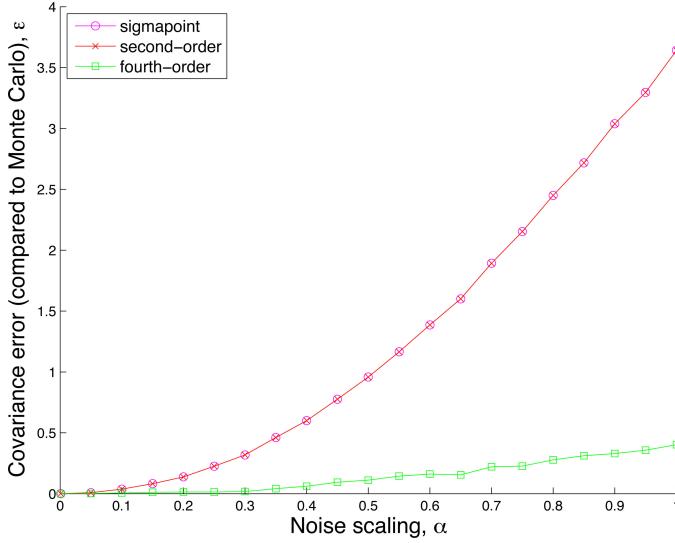


Fig. 3. Results from compound experiment (see Section III-D): Error ϵ in computing covariance associated with compounding two poses using three methods, as compared with Monte Carlo. The sigmapoint and second-order methods are algebraically equivalent for this problem and, thus, appear the same on the plot. The input covariances were gradually scaled up via the parameter α , highlighting the improved performance of the fourth-order method.

We compared each of the last three covariance matrices with the Monte Carlo one, using the *Frobenius norm*

$$\epsilon := \sqrt{\text{tr} \left((\Sigma - \Sigma_{\text{mc}})^T (\Sigma - \Sigma_{\text{mc}}) \right)}.$$

Fig. 3 shows that for small input covariance matrices (i.e., α small), there is very little difference between the various methods and the errors are all low compared with our benchmark. However, as we increase the magnitude of the input covariances, all the methods get worse, with the fourth-order method faring the best by about a factor of seven based on our error metric. Note, since α is scaling the covariance, the applied noise is increasing quadratically.

The second-order method and the sigmapoint method have indistinguishable performance, as they are algebraically equivalent. The fourth-order method goes beyond both of these by considering higher-order terms in the input covariance matrices. We did not compare the computational costs of the various methods as they are all extremely efficient, as compared with Monte Carlo.

It is also worth noting that our ability to correctly keep track of uncertainties on $SE(3)$ decreases with increasing uncertainty. This can be seen directly in Fig. 3, as error increases with increasing uncertainty. This suggests that it may be wise to use only relative pose variables in order to keep uncertainties small; this idea was suggested early on by Brooks [5] and demonstrated recently by Sibley *et al.* [32]. If uncertainty must be tracked globally on $SE(3)$, it may be necessary to move to an approach more along the lines of Lee *et al.* [24].

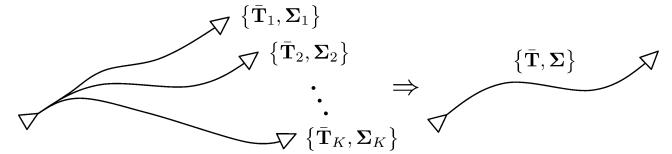


Fig. 4. Combining K pose estimates into a single fused estimate.

IV. FUSING POSES

This section will investigate a different type of nonlinearity, i.e., the fusing of several pose estimates, as depicted in Fig. 4.

A. Theory

Suppose that we have K estimates of a pose and associated uncertainties

$$\{\bar{T}_1, \Sigma_1\}, \{\bar{T}_2, \Sigma_2\}, \dots, \{\bar{T}_K, \Sigma_K\}. \quad (64)$$

If we think of these as uncertain measurements of the true pose, T_{true} , how can we optimally combine these into a single estimate $\{\bar{T}, \Sigma\}$?

The vectorspace solution to fusion is straightforward and can be found exactly in closed form

$$\bar{x} = \Sigma \sum_{k=1}^K \Sigma_k^{-1} \bar{x}_k, \quad \Sigma = \left(\sum_{k=1}^K \Sigma_k^{-1} \right)^{-1}. \quad (65)$$

The situation is somewhat more complicated when dealing with $SE(3)$, and we shall resort to an iterative scheme.

We define the error (that we will seek to minimize) as $\epsilon_k \sim \mathcal{N}(\mathbf{0}, \Sigma_k)$, which occurs between the individual estimate and the optimal estimate \bar{T}^* so that

$$\begin{aligned} \epsilon_k &:= \ln (\bar{T}^* \bar{T}_k^{-1})^\vee = \ln \left(\exp (\xi^\wedge) \underbrace{\bar{T} \bar{T}_k^{-1}}_{\text{small}} \right)^\vee \\ &= \ln (\exp (\xi^\wedge) \exp (\xi_k^\wedge))^\vee \end{aligned} \quad (66)$$

where \bar{T} is our best guess so far, ξ is a small (unknown but optimal) perturbation between our best guess and the optimum, and

$$\xi_k := \ln (\bar{T} \bar{T}_k^{-1})^\vee \quad (67)$$

is the difference between our best guess so far and each individual estimate.

Applying the version of the BCH formula in (29), we have

$$\epsilon_k \approx \xi_k + \mathcal{J}_k^{-1} \xi \quad (68)$$

correct to first order in ξ with

$$\mathcal{J}_k^{-1} := \sum_{n=0}^{\infty} \frac{B_n}{n!} (\xi_k^\wedge)^n. \quad (69)$$

Since ξ_k is fairly small, this series will converge rapidly and we can get away with keeping just a few terms. With our iterative scheme, ξ will (hopefully) converge to zero, and hence, we are justified to keep only terms linear in this quantity. Note that it is also possible to keep all the terms in the previous expansion by

utilizing the expression in (100) to compute \mathcal{J}_k analytically and then invert. The ξ_k will be the residual errors after convergence.

We define the cost function that we want to minimize as

$$\begin{aligned} V &:= \frac{1}{2} \sum_{k=1}^K \boldsymbol{\epsilon}_k^T \boldsymbol{\Sigma}_k^{-1} \boldsymbol{\epsilon}_k \\ &\approx \frac{1}{2} \sum_{k=1}^K (\boldsymbol{\xi}_k + \mathcal{J}_k^{-1} \boldsymbol{\xi})^T \boldsymbol{\Sigma}_k^{-1} (\boldsymbol{\xi}_k + \mathcal{J}_k^{-1} \boldsymbol{\xi}) \end{aligned}$$

which is already quadratic in $\boldsymbol{\xi}$. It is in fact a *Mahalanobis distance* [33], since we have chosen the weighting matrices to be the inverse covariance matrices; thus, minimizing V with respect to $\boldsymbol{\xi}$ is equivalent to maximizing the joint likelihood of the individual estimates. It is worth noting that because we are using a constraint-sensitive perturbation scheme, we do not need to worry about enforcing any constraints on our state variables during the optimization procedure. Taking the derivative with respect to $\boldsymbol{\xi}$ and setting to zero results in the following system of linear equations for the optimal value of $\boldsymbol{\xi}$:

$$\left(\sum_{k=1}^K \mathcal{J}_k^{-T} \boldsymbol{\Sigma}_k^{-1} \mathcal{J}_k^{-1} \right) \boldsymbol{\xi} = - \sum_{k=1}^K \mathcal{J}_k^{-T} \boldsymbol{\Sigma}_k^{-1} \boldsymbol{\xi}_k.$$

While this may look strange compared with (65), the Jacobian terms appear because our choice of error definition is in fact nonlinear owing to presence of the matrix exponentials. We then apply this optimal perturbation to our best guess so far

$$\bar{\mathbf{T}} \leftarrow \exp(\boldsymbol{\xi}^\wedge) \bar{\mathbf{T}}$$

which ensures $\bar{\mathbf{T}}$ remains in $SE(3)$, and iterate to convergence. This optimization strategy was also used by Strasdat *et al.* [34], who argued that it avoids singularities in the representation of the mean, $\bar{\mathbf{T}}$, but retains a minimal parameterization during the optimization step $\boldsymbol{\xi}$. At the last iteration, we take

$$\boldsymbol{\Sigma} = \left(\sum_{k=1}^K \mathcal{J}_k^{-T} \boldsymbol{\Sigma}_k^{-1} \mathcal{J}_k^{-1} \right)^{-1}$$

for the covariance matrix. This approach has the form of a Gauss–Newton method.

This fusion problem is similar to the one investigated by Smith *et al.* [14], but they only discuss the $K = 2$ case. Our study is closer to that of Long *et al.* [31], who discuss the $N = 2$ case and derive closed-form expressions for the fused mean and covariance for an arbitrary number of individual measurements, K ; however, they do not iterate their solution and they are tracking a slightly different PDF, as discussed earlier. Wolfe *et al.* [27] also discuss fusion at length, albeit again using a slightly different PDF than us. They discuss noniterative methods of fusion for arbitrary K and show numerical results for $K = 2$. We believe our approach generalizes all of these previous works by 1) allowing the number of individual estimates K to be arbitrary, 2) keeping an arbitrary number of terms in the approximation of the inverse Jacobian N , and 3) iterating to convergence via a Gauss–Newton style optimization method. Our approach may also be simpler to implement than some of these previous methods.

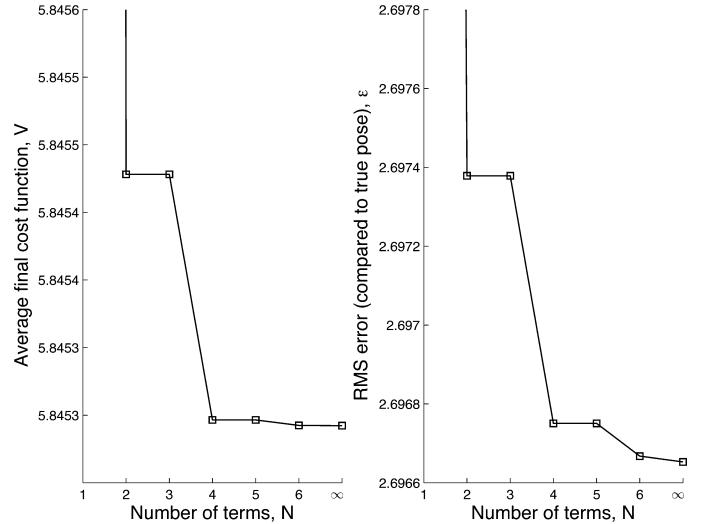


Fig. 5. Results from the fusion experiment (see Section IV-B). (Left) Average final cost function V as a function of the number of terms N , kept in \mathcal{J}_k^{-1} . (Right) Same for the root-mean-squared pose error with respect to the true pose. Both plots show there is benefit in keeping more than one term in \mathcal{J}_k^{-1} . The datapoint that is denoted by ‘ ∞ ’ uses the analytical expression in (100) to keep all the terms in the expansion.

B. Fusion Experiment

To validate the pose fusion method from the previous section, we used a true pose given by

$$\mathbf{T}_{\text{true}} := \exp(\boldsymbol{\xi}_{\text{true}}^\wedge), \quad \boldsymbol{\xi}_{\text{true}} := [1 \ 0 \ 0 \ 0 \ 0 \ \pi/6]^T$$

and then generated three random pose measurements

$$\begin{aligned} \bar{\mathbf{T}}_1 &:= \exp(\boldsymbol{\xi}_1^\wedge) \mathbf{T}_{\text{true}}, & \bar{\mathbf{T}}_2 &:= \exp(\boldsymbol{\xi}_2^\wedge) \mathbf{T}_{\text{true}} \\ \bar{\mathbf{T}}_3 &:= \exp(\boldsymbol{\xi}_3^\wedge) \mathbf{T}_{\text{true}} \end{aligned} \quad (70)$$

where

$$\begin{aligned} \boldsymbol{\xi}_1 &\sim \mathcal{N}\left(\mathbf{0}, \text{diag}\left\{10, 5, 5, \frac{1}{2}, 1, \frac{1}{2}\right\}\right) \\ \boldsymbol{\xi}_2 &\sim \mathcal{N}\left(\mathbf{0}, \text{diag}\left\{5, 15, 5, \frac{1}{2}, \frac{1}{2}, 1\right\}\right) \\ \boldsymbol{\xi}_3 &\sim \mathcal{N}\left(\mathbf{0}, \text{diag}\left\{5, 5, 25, 1, \frac{1}{2}, \frac{1}{2}\right\}\right). \end{aligned} \quad (71)$$

We then solved for the pose using our Gauss–Newton technique (iterating until convergence), using the initial condition $\bar{\mathbf{T}} = \mathbf{1}$. We repeated this for $N = 1 \dots 6$, which is the number of terms kept in \mathcal{J}_k^{-1} in (103), shown later. We also used the expression given in (100), shown later, to compute \mathcal{J}_k analytically (and then inverted numerically) and this is denoted by ‘ $N = \infty$ ’.

Fig. 5 plots two performance metrics. First, it plots the final converged value of the cost function, V_m , averaged over $M = 1000$ random trials $V := \frac{1}{M} \sum_{m=1}^M V_m$. Second, it plots the root-mean-squared pose error (with respect to the true pose) of our estimate $\bar{\mathbf{T}}_m$, again averaged over the same M random trials

$$\varepsilon := \sqrt{\frac{1}{M} \sum_{m=1}^M \boldsymbol{\epsilon}_m^T \boldsymbol{\epsilon}_m}, \quad \boldsymbol{\epsilon}_m := \ln(\mathbf{T}_{\text{true}} \bar{\mathbf{T}}_m^{-1})^\vee.$$

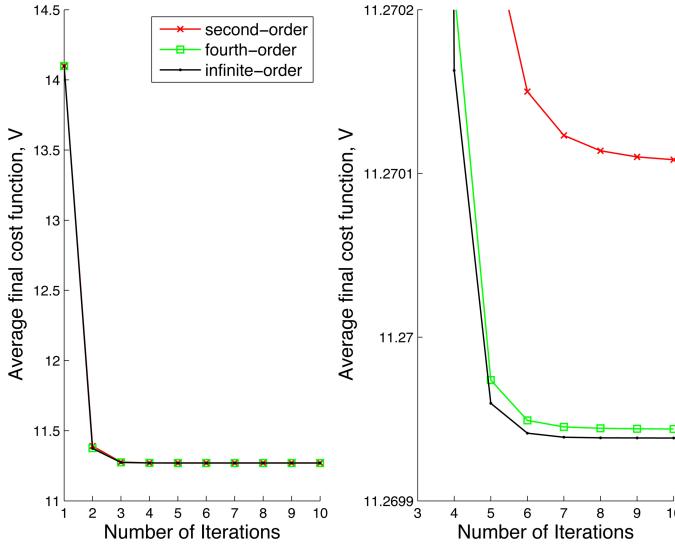


Fig. 6. Results from the fusion experiment. (Left) Convergence of the cost function V with successive Gauss–Newton iterations. This is for just one of the M trials that is used to generate Fig. 5. (Right) Same as left but zoomed in to show that the $N = 2, 4, \infty$ solutions converge to progressively lower costs.

The plots show that both measures of error are monotonically reduced with increasing N . Moreover, we see that for this example, almost all of the benefit is gained with just four terms (or possibly even two). The results for $N = 2, 3$ are identical as are those for $N = 4, 5$. This is because in the Bernoulli number sequence, $B_3 = 0$ and $B_5 = 0$; therefore, these terms make no additional contribution to \mathcal{J}_k^{-1} in (103). It is also worth stating that if we make the rotational part of the covariances in (71) any bigger, we end up with a lot of samples that have rotated by more than angle π , and this can be problematic for the performance metrics we are using.

Fig. 6 shows the convergence history of the cost, V , for a single random trial. The left side shows the strong benefit of iterating over the solution, while the right side shows that the cost converges to a lower value by keeping more terms in the approximation of \mathcal{J}_k^{-1} (cases $N = 2, 4, \infty$ shown). It would seem that taking $N = 4$ for about seven iterations gains most of the benefit, for this example.

V. PROPAGATING UNCERTAINTY THROUGH A NONLINEAR CAMERA MODEL

In estimation problems, we are often faced with passing uncertain quantities through nonlinear measurement models to produce expected measurements. Typically, this is carried out via linearization [35]. Sibley [36] shows how to carry out a second-order propagation for a stereo camera model accounting for landmark uncertainty but not pose uncertainty. Here, we derive the full second-order expression for the mean (and covariance) and compare this with Monte Carlo, the sigma-point transformation, and linearization. We begin by discussing our representation of points and then present the Taylor-series expansion of the measurement (camera) model followed by an experiment.

A. Homogeneous Points

Points in \mathbb{R}^3 can be represented using 4×1 homogeneous coordinates [3] as follows:

$$\mathbf{p} = \begin{bmatrix} sx \\ sy \\ sz \\ s \end{bmatrix} = \begin{bmatrix} \boldsymbol{\varepsilon} \\ \eta \end{bmatrix}$$

where s is some real, non-negative scalar, $\boldsymbol{\varepsilon} \in \mathbb{R}^3$, and η is scalar. When s is zero, it is not possible to convert back to \mathbb{R}^3 , as this case represents points that are infinitely far away. Thus, homogeneous coordinates can be used to describe near and distant landmarks with no singularities or scaling issues [37]. They are also a natural representation in that points may then be transformed from one frame to another very easily using transformation matrices (e.g., $\mathbf{p}_2 = \mathbf{T}_{21} \mathbf{p}_1$).

We will also make use of the following two operators⁷ to manipulate 4×1 columns

$$\begin{bmatrix} \boldsymbol{\varepsilon} \\ \eta \end{bmatrix}^\odot := \begin{bmatrix} \eta \mathbf{1} & -\boldsymbol{\varepsilon}^\wedge \\ \mathbf{0}^T & \mathbf{0}^T \end{bmatrix}, \quad \begin{bmatrix} \boldsymbol{\varepsilon} \\ \eta \end{bmatrix}^\odot := \begin{bmatrix} \mathbf{0} & \boldsymbol{\varepsilon} \\ -\boldsymbol{\varepsilon}^\wedge & \mathbf{0} \end{bmatrix} \quad (72)$$

which result in a 4×6 and 6×4 , respectively. With these definitions, we have the following useful identities:

$$\boldsymbol{\xi}^\wedge \mathbf{p} \equiv \mathbf{p}^\odot \boldsymbol{\xi}, \quad \mathbf{p}^T \boldsymbol{\xi}^\wedge \equiv \boldsymbol{\xi}^T \mathbf{p}^\odot \quad (73)$$

where $\boldsymbol{\xi} \in \mathbb{R}^6$ and $\mathbf{p} \in \mathbb{R}^4$, which will prove useful when manipulating expressions involving points and poses together. We also have the identity

$$(\mathbf{T}\mathbf{p})^\odot \equiv \mathbf{T}\mathbf{p}^\odot \mathbf{T}^{-1} \quad (74)$$

which is similar to (108) and (109), shown later. To the best of our knowledge, these operators and identities have not previously appeared in the literature.

To perturb points in homogeneous coordinates, we will operate directly on the xyz components by writing

$$\mathbf{p} = \bar{\mathbf{p}} + \mathbf{D} \boldsymbol{\zeta} \quad (75)$$

where $\boldsymbol{\zeta} \in \mathbb{R}^3$ is the perturbation and \mathbf{D} is a dilation matrix given by

$$\mathbf{D} := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}. \quad (76)$$

We thus have that $E[\mathbf{p}] = \bar{\mathbf{p}}$ and that

$$E[(\mathbf{p} - \bar{\mathbf{p}})(\mathbf{p} - \bar{\mathbf{p}})^T] = \mathbf{D} E[\boldsymbol{\zeta} \boldsymbol{\zeta}^T] \mathbf{D}^T \quad (77)$$

with no approximation.

B. Taylor-Series Expansion of a Camera Model

It is common to linearize a nonlinear observation model for use in pose estimation. In this section, we show how to do a

⁷The \odot operator for 4×1 columns is similar to the \boxdot operator defined by Furgale [17], which did not have the negative sign.

more general Taylor-series expansion of such a model and work out the second-order case in detail. Our camera model will be

$$\mathbf{y} = \mathbf{f}(\mathbf{T}\mathbf{p}) \quad (78)$$

where \mathbf{T} is the pose of the camera, and \mathbf{p} is the position of a landmark (as a homogeneous point). The nonlinear function $\mathbf{f}(\cdot)$ maps homogeneous points in the camera frame $\mathbf{T}\mathbf{p}$ to measurements \mathbf{y} . Our task will be to pass a Gaussian representation of the pose and landmark, given by $\{\mathbf{T}, \mathbf{p}, \boldsymbol{\Xi}\}$, where $\boldsymbol{\Xi}$ is a 9×9 covariance for both quantities, through the camera model to produce a mean and covariance for the measurement $\{\mathbf{y}, \mathbf{R}\}$.

We can think of this as the composition of two nonlinearities, one to transfer the landmark into the camera frame and one to produce the observations. We will treat each one in turn. If we change the pose of the camera and/or the position of the landmark a little bit, we have

$$\begin{aligned} \mathbf{T}\mathbf{p} &= \exp(\boldsymbol{\xi}^\wedge) \bar{\mathbf{T}} (\bar{\mathbf{p}} + \mathbf{D}\boldsymbol{\zeta}) \\ &\approx \left(1 + \boldsymbol{\xi}^\wedge + \frac{1}{2}\boldsymbol{\xi}^\wedge\boldsymbol{\xi}^\wedge\right) \bar{\mathbf{T}} (\bar{\mathbf{p}} + \mathbf{D}\boldsymbol{\zeta}) \end{aligned} \quad (79)$$

where we have kept the first two terms in the Taylor series for the pose perturbation. If we multiply out and continue to keep only those terms that are second order or lower in $\boldsymbol{\xi}$ and $\boldsymbol{\zeta}$, we have

$$\mathbf{T}\mathbf{p} \approx \mathbf{g} + \mathbf{G}\boldsymbol{\theta} + \frac{1}{2} \sum_{i=1}^4 \underbrace{\boldsymbol{\theta}^T \mathcal{G}_i \boldsymbol{\theta}}_{\text{scalar}} \mathbf{1}_i \quad (80)$$

where $\mathbf{1}_i$ is the i th column of the 4×4 identity matrix, and

$$\mathbf{g} := \bar{\mathbf{T}}\bar{\mathbf{p}} \quad (81)$$

$$\mathbf{G} := [\bar{\mathbf{T}}\bar{\mathbf{p}}^\top \quad \bar{\mathbf{T}}\mathbf{D}] \quad (82)$$

$$\mathcal{G}_i := \begin{bmatrix} \mathbf{1}_i^\top (\bar{\mathbf{T}}\bar{\mathbf{p}})^\top & \mathbf{1}_i^\top \bar{\mathbf{T}}\mathbf{D} \\ (\mathbf{1}_i^\top \bar{\mathbf{T}}\mathbf{D})^\top & \mathbf{0} \end{bmatrix} \quad (83)$$

$$\boldsymbol{\theta} := \begin{bmatrix} \boldsymbol{\xi} \\ \boldsymbol{\zeta} \end{bmatrix}. \quad (84)$$

Arriving at these expressions requires repeated application of the identities in the previous section.

To then apply the nonlinear camera model, we use the chain rule (for first and second derivatives) so that

$$\mathbf{f}(\mathbf{T}\mathbf{p}) \approx \mathbf{h} + \mathbf{H}\boldsymbol{\theta} + \frac{1}{2} \sum_{j=1}^J \underbrace{\boldsymbol{\theta}^T \mathcal{H}_j \boldsymbol{\theta}}_{\text{scalar}} \mathbf{1}_j \quad (85)$$

correct to second order in $\boldsymbol{\theta}$, where

$$\mathbf{h} := \mathbf{f}(\mathbf{g}) \quad (86)$$

$$\mathbf{H} := \mathbf{F}\mathbf{G}, \quad \mathbf{F} := \left. \frac{\partial \mathbf{f}}{\partial \mathbf{g}} \right|_{\mathbf{g}} \quad (87)$$

$$\mathcal{H}_j := \mathbf{G}^T \mathcal{F}_j \mathbf{G} + \sum_{i=1}^4 \underbrace{\mathbf{1}_j^T \mathbf{F} \mathbf{1}_i}_{\text{scalar}} \mathcal{G}_i \quad (88)$$

$$\mathcal{F}_j := \left. \frac{\partial^2 f_j}{\partial \mathbf{g} \partial \mathbf{g}^T} \right|_{\mathbf{g}} \quad (89)$$

j is an index over the J rows of $\mathbf{f}(\cdot)$, and $\mathbf{1}_j$ is the j th column of the $J \times J$ identity matrix. The Jacobian of $\mathbf{f}(\cdot)$ is \mathbf{F} , and the Hessian of the j th row $f_j(\cdot)$ is \mathcal{F}_j .

If we only care about the first-order perturbation, we simply have

$$\mathbf{f}(\mathbf{T}\mathbf{p}) = \mathbf{h} + \mathbf{H}\boldsymbol{\theta} \quad (90)$$

where \mathbf{h} and \mathbf{H} are unchanged.

These perturbed measurement equations can then be used within any estimation scheme we like; in the next section, we will use these with a stereo camera model to show the benefit of the second-order terms.

C. Propagating Gaussian Uncertainty Through the Camera

Suppose that the input uncertainties, embodied by $\boldsymbol{\theta}$, are zero-mean, Gaussian

$$\boldsymbol{\theta} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Xi}) \quad (91)$$

where we note that in general there could be correlations between the pose \mathbf{T} and the landmark \mathbf{p} .

Then, to first order, our measurement is given by

$$\mathbf{y}_{1\text{st}} = \mathbf{h} + \mathbf{H}\boldsymbol{\theta} \quad (92)$$

and $\bar{\mathbf{y}}_{1\text{st}} = E[\mathbf{y}_{1\text{st}}] = \mathbf{h}$, since $E[\boldsymbol{\theta}] = \mathbf{0}$ by assumption. The (second-order) covariance associated with the first-order camera model is given by

$$\mathbf{R}_{2\text{nd}} = E[(\mathbf{y}_{1\text{st}} - \bar{\mathbf{y}}_{1\text{st}})(\mathbf{y}_{1\text{st}} - \bar{\mathbf{y}}_{1\text{st}})^T] = \mathbf{H}\boldsymbol{\Xi}\mathbf{H}^T. \quad (93)$$

For the second-order camera model, we have

$$\mathbf{y}_{2\text{nd}} = \mathbf{h} + \mathbf{H}\boldsymbol{\theta} + \frac{1}{2} \sum_{j=1}^J \boldsymbol{\theta}^T \mathcal{H}_j \boldsymbol{\theta} \mathbf{1}_j \quad (94)$$

and consequently

$$\bar{\mathbf{y}}_{2\text{nd}} = E[\mathbf{y}_{2\text{nd}}] = \mathbf{h} + \frac{1}{2} \sum_{j=1}^J \text{tr}(\mathcal{H}_j \boldsymbol{\Xi}) \mathbf{1}_j \quad (95)$$

which has an extra nonzero term as compared with the first-order camera model. The larger the input covariance $\boldsymbol{\Xi}$ is, the larger this term can become, depending on the nonlinearity. For a linear camera model, $\mathcal{H}_j = \mathbf{0}$ and the second- and first-order camera model means are identical.

We will also compute a (fourth-order) covariance, but with just second-order terms in the camera model expansion. To do this properly, we should expand the camera model to third order as there is an additional fourth-order covariance term involving the product of first- and third-order camera-model terms; however, this would involve a complicated expression employing the third derivative of the camera model. As such, the approximate fourth-order covariance we will use is given by

$$\begin{aligned} \mathbf{R}_{4\text{th}} &\approx E[(\mathbf{y}_{2\text{nd}} - \bar{\mathbf{y}}_{2\text{nd}})(\mathbf{y}_{2\text{nd}} - \bar{\mathbf{y}}_{2\text{nd}})^T] \\ &= \mathbf{H}\boldsymbol{\Xi}\mathbf{H}^T - \frac{1}{4} \left(\sum_{i=1}^J \text{tr}(\mathcal{H}_i \boldsymbol{\Xi}) \mathbf{1}_i \right) \left(\sum_{j=1}^J \text{tr}(\mathcal{H}_j \boldsymbol{\Xi}) \mathbf{1}_j \right)^T \end{aligned}$$

$$+ \frac{1}{4} \sum_{i,j=1}^J \sum_{k,\ell,m,n=1}^9 \mathcal{H}_{ik\ell} \mathcal{H}_{jm\ell} (\Xi_{k\ell} \Xi_{mn} \\ + \Xi_{km} \Xi_{\ell n} + \Xi_{kn} \Xi_{\ell m})$$

where $\mathcal{H}_{ik\ell}$ is the (k, ℓ) th element of \mathcal{H}_i , and $\Xi_{k\ell}$ is the (k, ℓ) th element of Ξ . The first- and third-order terms in the covariance expansion are identically zero owing to the symmetry of the Gaussian density. The last term in the above makes use of *Isserlis' theorem* for Gaussian variables

$$E[\xi_k \xi_\ell \xi_m \xi_n] = E[\xi_k \xi_\ell] E[\xi_m \xi_n] + E[\xi_k \xi_m] E[\xi_\ell \xi_n] \\ + E[\xi_k \xi_n] E[\xi_\ell \xi_m].$$

D. Sigma point Method

Finally, we can also make use of the popular *sigmapoint transformation* [29] to pass uncertainty through the nonlinear camera model. As in the pose compounding problem, we tailor this to our specific type of $SE(3)$ perturbation. We begin by approximating the input Gaussian using a finite number of samples $\{\mathbf{T}_\ell, \mathbf{p}_\ell\}$

$$\mathbf{S}\mathbf{S}^T := \Xi, \quad (\text{Cholesky decomposition})$$

$$\boldsymbol{\theta}_\ell := \mathbf{0}$$

$$\boldsymbol{\theta}_\ell := \sqrt{L + \kappa} \text{col}_\ell \mathbf{S}, \quad \ell = 1 \dots L$$

$$\boldsymbol{\theta}_{\ell+L} := -\sqrt{L + \kappa} \text{col}_\ell \mathbf{S}, \quad \ell = 1 \dots L$$

$$\begin{bmatrix} \boldsymbol{\xi}_\ell \\ \boldsymbol{\zeta}_\ell \end{bmatrix} := \boldsymbol{\theta}_\ell$$

$$\mathbf{T}_\ell := \exp(\boldsymbol{\xi}_\ell^\wedge) \bar{\mathbf{T}}$$

$$\mathbf{p}_\ell := \bar{\mathbf{p}} + \mathbf{D} \boldsymbol{\zeta}_\ell$$

where κ is a user-definable constant,⁸ and $L = 9$. We then pass each of these samples through the nonlinear camera model

$$\mathbf{y}_\ell = \mathbf{f}(\mathbf{T}_\ell \mathbf{p}_\ell), \quad \ell = 0 \dots 2L.$$

These are combined to create the output mean and covariance according to

$$\bar{\mathbf{y}}_{\text{sp}} = \frac{1}{L + \kappa} \left(\kappa \mathbf{y}_0 + \frac{1}{2} \sum_{\ell=1}^{2L} \mathbf{y}_\ell \right)$$

$$\mathbf{R}_{\text{sp}} = \frac{1}{L + \kappa} \left(\kappa (\mathbf{y}_0 - \bar{\mathbf{y}}_{\text{sp}})(\mathbf{y}_0 - \bar{\mathbf{y}}_{\text{sp}})^T + \frac{1}{2} \sum_{\ell=1}^{2L} (\mathbf{y}_\ell - \bar{\mathbf{y}}_{\text{sp}})(\mathbf{y}_\ell - \bar{\mathbf{y}}_{\text{sp}})^T \right).$$

The next section will provide the details for a specific nonlinear camera model $\mathbf{f}(\cdot)$ representing a stereo camera.

E. Stereo Camera Model

To demonstrate the propagation of uncertainty through a nonlinear measurement model, $\mathbf{f}(\cdot)$, we will employ a stereo camera

⁸For all experiments in this paper, we used $\kappa = 0$.

given by

$$\mathbf{f}(\mathbf{p}) = \mathbf{M} \frac{1}{p_3} \mathbf{p}$$

where

$$\mathbf{f} := \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix}, \quad \mathbf{p} := \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix}, \quad \mathbf{M} := \begin{bmatrix} f_u & 0 & c_u & f_u \frac{b}{2} \\ 0 & f_v & c_v & 0 \\ f_u & 0 & c_u & -f_u \frac{b}{2} \\ 0 & f_v & c_v & 0 \end{bmatrix}$$

and f_u, f_v are the horizontal, vertical focal lengths (in pixels), (c_u, c_v) is the optical center of the images (in pixels), and b is the separation between the cameras (in meters). The optical axis of the camera is along the z -, or p_3 -, direction.

The Jacobian of this measurement model is given by

$$\frac{\partial \mathbf{f}}{\partial \mathbf{p}} = \mathbf{M} \frac{1}{p_3} \begin{bmatrix} 1 & 0 & -\frac{p_1}{p_3} & 0 \\ 0 & 1 & -\frac{p_2}{p_3} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{p_4}{p_3} & 1 \end{bmatrix}$$

and the Hessian is given by

$$\frac{\partial^2 f_1}{\partial \mathbf{p} \partial \mathbf{p}^T} = \frac{f_u}{p_3^2} \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & \frac{2p_1 + bp_4}{p_3} & -\frac{b}{2} \\ 0 & 0 & \frac{b}{2} & 0 \end{bmatrix}$$

$$\frac{\partial^2 f_2}{\partial \mathbf{p} \partial \mathbf{p}^T} = \frac{\partial^2 f_4}{\partial \mathbf{p} \partial \mathbf{p}^T} = \frac{f_v}{p_3^2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & \frac{2p_2}{p_3} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\frac{\partial^2 f_3}{\partial \mathbf{p} \partial \mathbf{p}^T} = \frac{f_u}{p_3^2} \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & \frac{2p_1 - bp_4}{p_3} & \frac{b}{2} \\ 0 & 0 & \frac{b}{2} & 0 \end{bmatrix}$$

where we have shown each component separately.

F. Camera Experiment

We used the following methods to pass a Gaussian uncertainty on camera pose and landmark position through the nonlinear stereo camera model.

- 1) *Monte Carlo*: We drew a large number $M = 1000\,000$ of random samples from the input distribution, passed these through the camera model, and then computed the mean $\bar{\mathbf{y}}_{\text{mc}}$ and covariance \mathbf{R}_{mc} . This slow-but-accurate approach served as our benchmark to which the other three much faster methods were compared.
- 2) *First/Second Order*: We used the first-order camera model to compute $\bar{\mathbf{y}}_{1\text{st}}$ and $\mathbf{R}_{2\text{nd}}$, as described previously.

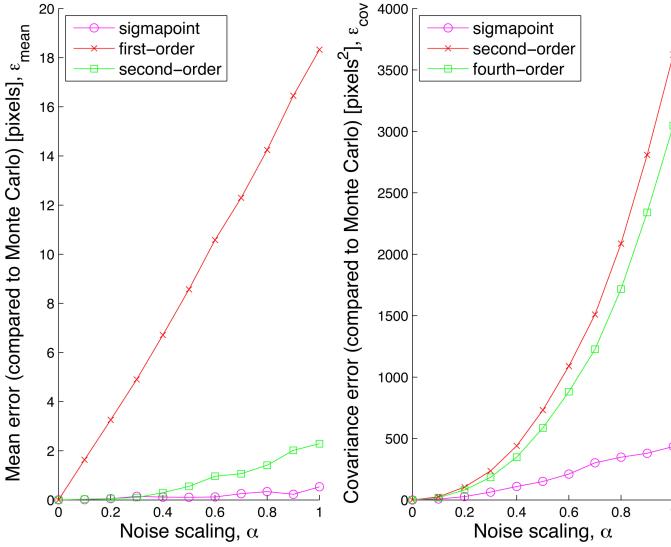


Fig. 7. Results from the stereo camera experiment (see Section V-F). (Left) Mean and (right) covariance errors $\varepsilon_{\text{mean}}$ and ε_{cov} for three methods of passing a Gaussian uncertainty through a nonlinear stereo camera model, as compared with Monte Carlo. The parameter α scales the magnitude of the input covariance matrix.

- 3) *Second/Fourth Order*: We used the second-order camera model to compute $\bar{\mathbf{y}}_{2\text{nd}}$ and $\mathbf{R}_{4\text{th}}$, as described previously.
- 4) *Sigmapoint*: We used the sigmapoint method that has been described previously to compute $\bar{\mathbf{y}}_{\text{sp}}$ and \mathbf{R}_{sp} .

The camera parameters were

$$b = 0.25 \text{ m}, f_u = f_v = 200 \text{ pixels}, c_u = c_v = 0 \text{ pixels}.$$

We used the camera pose $\mathbf{T} := \mathbf{1}$ and let the landmark be located at $\mathbf{p} := [10 \ 10 \ 10 \ 1]^T$. For the combined pose/landmark uncertainty, we used an input covariance of

$$\mathbf{\Xi} := \alpha \times \text{diag} \left\{ \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{100}, \frac{1}{100}, \frac{1}{100}, 1, 1, 1 \right\}$$

where $\alpha \in [0, 1]$ is a scaling parameter that allowed us to parametrically increase the magnitude of the uncertainty.

To gauge the performance, we evaluated both the mean and covariance of each method by comparing the results to those of the Monte Carlo simulation according to the following metrics:

$$\begin{aligned} \varepsilon_{\text{mean}} &:= \sqrt{(\bar{\mathbf{y}} - \bar{\mathbf{y}}_{\text{mc}})^T (\bar{\mathbf{y}} - \bar{\mathbf{y}}_{\text{mc}})} \\ \varepsilon_{\text{cov}} &:= \sqrt{\text{tr}((\mathbf{R} - \mathbf{R}_{\text{mc}})^T (\mathbf{R} - \mathbf{R}_{\text{mc}}))} \end{aligned}$$

where the latter is the *Frobenius norm*.

Fig. 7 shows the two performance metrics, $\varepsilon_{\text{mean}}$ and ε_{cov} , for each of the three techniques over a wide range of noise scalings α . We see that the sigmapoint technique does the best on both mean and covariance. The second-order technique does reasonably well on the mean, but the corresponding fourth-order technique does poorly on the covariance (because of our inability to compute a fully fourth-order-accurate covariance, as explained earlier).

Fig. 8 provides a snapshot of a portion of the left image of the stereo camera with the mean and one-standard-deviation

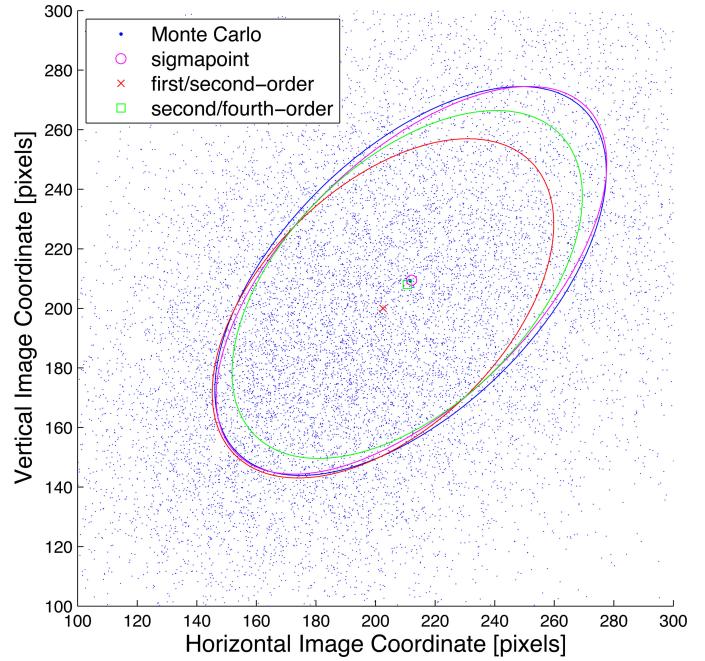


Fig. 8. Results from the stereo camera experiment (see Section V-F). A portion of the left image of a stereo camera showing the mean and covariance (as a one-standard-deviation ellipse) for four methods of imaging a landmark with Gaussian uncertainty on the camera's pose and the landmark's position. This case corresponds to the $\alpha = 1$ datapoint in Fig. 7.

covariance ellipses shown for all techniques. We see that the sigmapoint technique does an excellent job on both the mean and the covariance, while the others do not fare as well.

We do not believe there has been another comparison of these methods of propagating uncertainty through a nonlinear camera model with uncertainty on the camera pose and landmark position, the closest being the work of Sibley [36].

VI. CONCLUSION

We have presented some generic techniques (and provided accompanying MATLAB scripts⁹) to associate uncertainty with 3-D pose (and landmark) variables and shown how to use these in three worked examples:

- 1) compounding two poses with associated uncertainties (to fourth order in our noise variables or second order in the associated covariances);
- 2) fusing multiple uncertain estimates of a pose (iteratively, to arbitrary order);
- 3) propagating uncertainty through a nonlinear camera model (to fourth order in our noise variables or second order in the associated covariances).

The general outcome is that depending on the application, different methods have advantages, with the higher-order methods always doing better than the lower ones. The choice of whether to use a sigmapoint method to propagate uncertainty depends on

⁹An accompanying set of scripts to implement many of the key equations and regenerate all of the plots in the paper is downloadable from <http://asrl.utias.utoronto.ca/code/>

the application: *no* for pose compounding but *yes* to propagate uncertainty through a nonlinear camera model.

The contributions of this paper are both the specific examples as well as the detailed notation and identities to manipulate poses and associated uncertainties. Our use of the BCH formula in these manipulations, particularly our use of (29) and the tangible connection to the $SE(3)$ Jacobian, is novel; this is important because it allows us to avoid unnecessarily introducing the concept of *Lie derivatives*, and instead, we can simply manipulate perturbations representing noise variables. We believe our comparison of the sigmapoint transformation to the Taylor-expansion methods to be new as well. Finally, the appendix contains a few previously unpublished tidbits including a closed-form expression for the $SE(3)$ Jacobian in terms of all of its constituent blocks.

We have been using these methods in several bundle-adjustment and pose-graph optimization problems for the past several years with good success. Looking forward, we believe these techniques could find application in other problems requiring detailed bookkeeping of pose uncertainties and, thus, hope others will find them useful.

APPENDIX

This Appendix contains additional notation and useful expressions for implementation of key expressions in this paper.

A. Useful Closed-Form Expressions

The matrix exponential to build a transformation can be evaluated as

$$\mathbf{T} = \exp(\boldsymbol{\xi}^\wedge) = \sum_{n=0}^{\infty} \frac{1}{n!} (\boldsymbol{\xi}^\wedge)^n = \begin{bmatrix} \mathbf{C} & \mathbf{J}\boldsymbol{\rho} \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (96)$$

where the rightmost expression can be computed in closed form. To do so, we use the components of $\boldsymbol{\xi} = [\begin{smallmatrix} \boldsymbol{\rho} \\ \phi \end{smallmatrix}]$ to compute the rotation matrix $\mathbf{C} \in SO(3)$

$$\begin{aligned} \mathbf{C} := \exp(\boldsymbol{\phi}^\wedge) &= \sum_{n=0}^{\infty} \frac{1}{n!} (\boldsymbol{\phi}^\wedge)^n \\ &= \cos \phi \mathbf{1} + (1 - \cos \phi) \mathbf{a} \mathbf{a}^T + \sin \phi \mathbf{a}^\wedge \end{aligned} \quad (97)$$

where $\phi = \|\boldsymbol{\phi}\|$ is the angle of rotation, and $\mathbf{a} = \boldsymbol{\phi}/\phi$ is the axis of rotation. We also compute

$$\begin{aligned} \mathbf{J} := \int_0^1 \mathbf{C}^\alpha d\alpha &= \sum_{n=0}^{\infty} \frac{1}{(n+1)!} (\boldsymbol{\phi}^\wedge)^n \\ &= \frac{\sin \phi}{\phi} \mathbf{1} + \left(1 - \frac{\sin \phi}{\phi}\right) \mathbf{a} \mathbf{a}^T + \frac{1 - \cos \phi}{\phi} \mathbf{a}^\wedge \end{aligned} \quad (98)$$

which is the (left) *Jacobian* of $SO(3)$. The inverse of \mathbf{C} is simply $\mathbf{C}^{-1} = \mathbf{C}^T$, and the inverse of \mathbf{J} is given by

$$\begin{aligned} \mathbf{J}^{-1} &= \sum_{n=0}^{\infty} \frac{B_n}{n!} (\boldsymbol{\phi}^\wedge)^n \\ &= \frac{\phi}{2} \cot \frac{\phi}{2} \mathbf{1} + \left(1 - \frac{\phi}{2} \cot \frac{\phi}{2}\right) \mathbf{a} \mathbf{a}^T - \frac{\phi}{2} \mathbf{a}^\wedge. \end{aligned} \quad (99)$$

Similar closed-form expressions can be found in [21] and [38].

We can also compute the Jacobian for $SE(3)$ according to

$$\mathcal{J} := \int_0^1 \mathbf{T}^\alpha d\alpha \equiv \sum_{n=0}^{\infty} \frac{1}{(n+1)!} (\boldsymbol{\xi}^\wedge)^n \equiv \begin{bmatrix} \mathbf{J} & \mathbf{Q} \\ \mathbf{0} & \mathbf{J} \end{bmatrix} \quad (100)$$

where

$$\mathcal{T} = \text{Ad}(\mathbf{T}) = \exp(\text{ad}(\boldsymbol{\xi}^\wedge)) = \exp(\boldsymbol{\xi}^\wedge) = \begin{bmatrix} \mathbf{C} & (\mathbf{J}\boldsymbol{\rho})^\wedge \mathbf{C} \\ \mathbf{0} & \mathbf{C} \end{bmatrix}. \quad (101)$$

The closed-form expression on the right can be populated using \mathbf{J} above and

$$\begin{aligned} \mathbf{Q} &= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{1}{(n+m+2)!} (\boldsymbol{\phi}^\wedge)^n \boldsymbol{\rho}^\wedge (\boldsymbol{\phi}^\wedge)^m \\ &= \frac{1}{2} \boldsymbol{\rho}^\wedge + \frac{\phi - \sin \phi}{\phi^3} (\boldsymbol{\phi}^\wedge \boldsymbol{\rho}^\wedge + \boldsymbol{\rho}^\wedge \boldsymbol{\phi}^\wedge + \boldsymbol{\phi}^\wedge \boldsymbol{\rho}^\wedge \boldsymbol{\phi}^\wedge) \\ &\quad - \frac{1 - \frac{\phi^2}{2} - \cos \phi}{\phi^4} (\boldsymbol{\phi}^\wedge \boldsymbol{\phi}^\wedge \boldsymbol{\rho}^\wedge + \boldsymbol{\rho}^\wedge \boldsymbol{\phi}^\wedge \boldsymbol{\phi}^\wedge - 3\boldsymbol{\phi}^\wedge \boldsymbol{\rho}^\wedge \boldsymbol{\phi}^\wedge) \\ &\quad - \frac{1}{2} \left(\frac{1 - \frac{\phi^2}{2} - \cos \phi}{\phi^4} - 3 \frac{\phi - \sin \phi - \frac{\phi^3}{6}}{\phi^5} \right) \\ &\quad \times (\boldsymbol{\phi}^\wedge \boldsymbol{\rho}^\wedge \boldsymbol{\phi}^\wedge \boldsymbol{\phi}^\wedge + \boldsymbol{\phi}^\wedge \boldsymbol{\phi}^\wedge \boldsymbol{\rho}^\wedge \boldsymbol{\phi}^\wedge). \end{aligned} \quad (102)$$

To the best of our knowledge, this last expression does not appear in the previous literature. The inverse of \mathcal{J} is simply

$$\mathcal{J}^{-1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} (\boldsymbol{\xi}^\wedge)^n \equiv \begin{bmatrix} \mathbf{J}^{-1} & -\mathbf{J}^{-1} \mathbf{Q} \mathbf{J}^{-1} \\ \mathbf{0} & \mathbf{J}^{-1} \end{bmatrix}. \quad (103)$$

Bullo and Murray [39, eq. (2.21)] also discuss \mathcal{J}^{-1} , pointing out that the two diagonal blocks are indeed \mathbf{J}^{-1} but do not work out the details of top-right block, which amounts to knowing the details of \mathbf{Q} .

B. Converting Between $\boldsymbol{\xi}$ and \mathbf{T}

The detailed steps to convert (in closed form) between a $\boldsymbol{\xi} \in \mathbb{R}^6$ and a $\mathbf{T} \in SE(3)$ are provided as follows.

- 1) From $\boldsymbol{\xi}$ to \mathbf{T} :
 - a) Deconstruct $\boldsymbol{\xi} = [\begin{smallmatrix} \boldsymbol{\rho} \\ \phi \end{smallmatrix}]$ into $\boldsymbol{\rho}$ and ϕ .
 - b) Using ϕ , compute \mathbf{C} and \mathbf{J} according to (97) and (98).
 - c) Using \mathbf{C} , \mathbf{J} , and $\boldsymbol{\rho}$, construct $\mathbf{T} = [\begin{smallmatrix} \mathbf{C} & \mathbf{J}\boldsymbol{\rho} \\ \mathbf{0}^T & 1 \end{smallmatrix}]$.
- We have exactly that $\mathbf{T} = \exp(\boldsymbol{\xi}^\wedge)$.
- 2) From \mathbf{T} to $\boldsymbol{\xi}$:
 - a) Deconstruct $\mathbf{T} = [\begin{smallmatrix} \mathbf{C} & \mathbf{r} \\ \mathbf{0}^T & 1 \end{smallmatrix}]$ into \mathbf{C} and \mathbf{r} .
 - b) Using \mathbf{C} , compute ϕ . This can be done by
 - i) computing the rotation axis \mathbf{a} as the eigenvector corresponding to a unit eigenvalue of \mathbf{C} (since $\mathbf{C}\mathbf{a} = \mathbf{a}$);

- ii) solving for the rotation angle ϕ in $\text{tr}(\mathbf{C}) = 2 \cos \phi + 1$ [check for the correct sign by plugging ϕ and \mathbf{a} back into (97)];
- iii) letting $\phi = \phi\mathbf{a}$.
- c) Using ϕ , compute \mathbf{J}^{-1} according to (99).
- d) Using \mathbf{J}^{-1} and \mathbf{r} , compute $\boldsymbol{\rho} = \mathbf{J}^{-1}\mathbf{r}$.
- e) Using $\boldsymbol{\rho}$ and ϕ , construct $\boldsymbol{\xi} = [\begin{smallmatrix} \boldsymbol{\rho} \\ \phi \end{smallmatrix}]$.

We have exactly that $\boldsymbol{\xi} = \ln(\mathbf{T})^\vee$.

C. Proof that $\text{Ad}(\exp(\boldsymbol{\xi}^\wedge)) \equiv \exp(\text{ad}(\boldsymbol{\xi}^\wedge))$

We were unable to find this direct proof in the literature and therefore provide it. Starting from the right-hand side, we have

$$\begin{aligned} \exp(\text{ad}(\boldsymbol{\xi}^\wedge)) &= \exp(\boldsymbol{\xi}^\wedge) = \sum_{n=0}^{\infty} \frac{1}{n!} (\boldsymbol{\xi}^\wedge)^n \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \begin{bmatrix} \boldsymbol{\phi}^\wedge & \boldsymbol{\rho}^\wedge \\ \mathbf{0} & \boldsymbol{\phi}^\wedge \end{bmatrix}^n = \begin{bmatrix} \mathbf{C} & \mathbf{K} \\ \mathbf{0} & \mathbf{C} \end{bmatrix} \quad (104) \end{aligned}$$

where \mathbf{C} is the usual expression in (97), and

$$\mathbf{K} := \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{1}{(n+m+1)!} (\boldsymbol{\phi}^\wedge)^n \boldsymbol{\rho}^\wedge (\boldsymbol{\phi}^\wedge)^m$$

which can be found through careful manipulation. Starting from the left-hand side, we have

$$\text{Ad}(\exp(\boldsymbol{\xi}^\wedge)) = \text{Ad}\left(\begin{bmatrix} \mathbf{C} & \mathbf{J}\boldsymbol{\rho} \\ \mathbf{0}^T & 1 \end{bmatrix}\right) = \begin{bmatrix} \mathbf{C} & (\mathbf{J}\boldsymbol{\rho})^\wedge \mathbf{C} \\ \mathbf{0} & \mathbf{C} \end{bmatrix} \quad (105)$$

where \mathbf{J} is given in (98). Comparing (104) and (105), what remains to be shown is that $\mathbf{K} = (\mathbf{J}\boldsymbol{\rho})^\wedge \mathbf{C}$. To see this, we use the following sequence of manipulations:

$$\begin{aligned} (\mathbf{J}\boldsymbol{\rho})^\wedge \mathbf{C} &= \left(\int_0^1 \mathbf{C}^\alpha d\alpha \boldsymbol{\rho} \right)^\wedge \mathbf{C} = \int_0^1 (\mathbf{C}^\alpha \boldsymbol{\rho})^\wedge \mathbf{C} d\alpha \\ &= \int_0^1 \mathbf{C}^\alpha \boldsymbol{\rho}^\wedge \mathbf{C}^{1-\alpha} d\alpha = \int_0^1 \exp(\alpha \boldsymbol{\phi}^\wedge) \boldsymbol{\rho}^\wedge \exp((1-\alpha) \boldsymbol{\phi}^\wedge) d\alpha \\ &= \int_0^1 \left(\sum_{n=0}^{\infty} \frac{1}{n!} (\alpha \boldsymbol{\phi}^\wedge)^n \right) \boldsymbol{\rho}^\wedge \left(\sum_{m=0}^{\infty} \frac{1}{m!} ((1-\alpha) \boldsymbol{\phi}^\wedge)^m \right) d\alpha \\ &= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{1}{n! m!} \left(\int_0^1 \alpha^n (1-\alpha)^m d\alpha \right) (\boldsymbol{\phi}^\wedge)^n \boldsymbol{\rho}^\wedge (\boldsymbol{\phi}^\wedge)^m \end{aligned}$$

where we have used that \wedge is linear and that $(\mathbf{Cu})^\wedge = \mathbf{Cu}^\wedge \mathbf{C}^T$. We can show that $\int_0^1 \alpha^n (1-\alpha)^m d\alpha = \frac{n! m!}{(n+m+1)!}$ (after several integrations by parts), and therefore, $\mathbf{K} \equiv (\mathbf{J}\boldsymbol{\rho})^\wedge \mathbf{C}$, concluding the proof.

D. Additional Useful Identities

Some more useful identities for $SE(3)$ and $se(3)$ are the following:

$$\boldsymbol{\xi}_a^\wedge \boldsymbol{\xi}_a \equiv \mathbf{0} \quad (106)$$

$$\boldsymbol{\xi}_a^\wedge \boldsymbol{\xi}_b \equiv -\boldsymbol{\xi}_b^\wedge \boldsymbol{\xi}_a \quad (107)$$

$$(\mathcal{T}_a \boldsymbol{\xi}_b)^\wedge \equiv \mathbf{T}_a \boldsymbol{\xi}_b^\wedge \mathbf{T}_a^{-1} \quad (108)$$

$$(\mathcal{T}_a \boldsymbol{\xi}_b)^\wedge \equiv \mathcal{T}_a \boldsymbol{\xi}_b^\wedge \mathcal{T}_a^{-1} \quad (109)$$

$$\exp((\mathcal{T}_a \boldsymbol{\xi}_b)^\wedge) \equiv \mathbf{T}_a \exp(\boldsymbol{\xi}_b^\wedge) \mathbf{T}_a^{-1} \quad (110)$$

$$\exp((\mathcal{T}_a \boldsymbol{\xi}_b)^\wedge) \equiv \mathcal{T}_a \exp(\boldsymbol{\xi}_b^\wedge) \mathcal{T}_a^{-1} \quad (111)$$

$$\text{Ad}(\boldsymbol{\xi}_b^\wedge \mathbf{T}_a) \equiv \boldsymbol{\xi}_b^\wedge \mathcal{T}_a \quad (112)$$

$$[\boldsymbol{\xi}_a^\wedge, \boldsymbol{\xi}_b^\wedge] = \boldsymbol{\xi}_a^\wedge \boldsymbol{\xi}_b^\wedge - \boldsymbol{\xi}_b^\wedge, \boldsymbol{\xi}_a^\wedge \equiv (\boldsymbol{\xi}_a^\wedge \boldsymbol{\xi}_b)^\wedge \quad (113)$$

$$[\boldsymbol{\xi}_a^\wedge, \boldsymbol{\xi}_b^\wedge] = \boldsymbol{\xi}_a^\wedge \boldsymbol{\xi}_b^\wedge - \boldsymbol{\xi}_b^\wedge, \boldsymbol{\xi}_a^\wedge \equiv (\boldsymbol{\xi}_a^\wedge \boldsymbol{\xi}_b)^\wedge \quad (114)$$

where $\boldsymbol{\xi}_a, \boldsymbol{\xi}_b \in \mathbb{R}^6$ and $\mathbf{T}_a = \exp(\boldsymbol{\xi}_a^\wedge), \mathcal{T}_a = \exp(\boldsymbol{\xi}_a^\wedge) \in SE(3)$.

ACKNOWLEDGMENT

The authors would like to thank Prof. F. Dellaert at Georgia Tech and H. Sommer at ETH for many useful discussions during the preparation of this paper.

REFERENCES

- [1] D. C. Brown, "A solution to the general problem of multiple station analytical stereotriangulation," Patrick Air Force Base, Florida, RCA-MTP Data Reduction Tech. Rep. 43 (or AFMTC TR 58-8), 1958.
- [2] R. M. Murray, Z. Li, and S. Sastry, *A Mathematical Introduction to Robotic Manipulation*. Boca Raton, FL, USA: CRC, 1994.
- [3] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, U.K.: Cambridge Univ. Press, 2000.
- [4] J. Stillwell, *Naive Lie Theory*. New York, NY, USA: Springer-Verlag, 2008.
- [5] R. Brooks, "Visual map making for a mobile robot," in *Proc. IEEE Int. Conf. Robot. Autom.*, Mar. 1985, vol. 2, pp. 824–829.
- [6] R. C. Smith and P. Cheeseman, "On the representation and estimation of spatial uncertainty," *Int. J. Robot. Res.*, vol. 5, no. 4, pp. 56–68, 1986.
- [7] H. F. Durrant-Whyte, "Uncertain geometry in robotics," *IEEE J. Robot. Autom.*, vol. 4, no. 1, pp. 23–31, Feb. 1988.
- [8] R. C. Smith, M. Self, and P. Cheeseman, "Estimating uncertain spatial relationships in robotics," in *Autonomous Robot Vehicles*, I. J. Cox and G. T. Wilfong, Eds. New York, NY, USA: Springer-Verlag, 1990, pp. 167–193.
- [9] H. Durrant-Whyte and T. Bailey, "Simultaneous localisation and mapping (SLAM): Part I the essential algorithms," *IEEE Robot. Autom. Mag.*, vol. 11, no. 3, pp. 99–110, Jun. 2006.
- [10] G. S. Chirikjian, *Stochastic Models, Information Theory, and Li Groups: Classical Results and Geometric Methods*. vol. 1, New York, NY, USA: Birkhauser, 2009.
- [11] G. S. Chirikjian, *Stochastic Models, Information Theory, and Li Groups: Analytic Methods and Modern Applications*. vol. 2, New York, NY, USA: Birkhauser, 2009.
- [12] S. Fu and C. S. G. Lee, "Uncertainty manipulation and propagation and verification of applicability of actions in assembly tasks," in *Proc. IEEE Int. Conf. Robot. Autom.*, 1991, vol. 3, pp. 2471–2476.
- [13] S. Fu and C. S. G. Lee, "Manipulation and propagation of uncertainty and verification of applicability of actions in assembly tasks," *IEEE Trans. Syst., Man Cybern.*, vol. 22, no. 6, pp. 1376–1389, Nov./Dec. 1992.
- [14] P. Smith, T. Drummond, and K. Roussopoulos, "Computing MAP trajectories by representing, propagating, and combining PDFs over groups," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2003, pp. 1275–1282.
- [15] C. Hertzberg, R. Wagner, U. Frese, and L. Schröder, "Integrating generic sensor fusion algorithms with sound state representations through encapsulation of manifolds," *Inf. Fusion*, vol. 14, no. 1, pp. 57–77, 2013.
- [16] S. Sastry, *Nonlinear Systems: Analysis, Stability, and Control*. New York, NY, USA: Springer-Verlag, 1999.
- [17] P. T. Furgale, "Extensions to the visual odometry pipeline for the exploration of planetary surfaces" Ph.D. dissertation, Dept. Aerosp. Eng., Univ. Toronto, Toronto, ON, Canada, 2011.
- [18] P. C. Hughes, *Spacecraft Attitude Dynamics*. New York, NY, USA: Wiley, 1986.
- [19] M. D. Shuster, "A survey of attitude representations," *J. Astronaut. Sci.*, vol. 41, no. 4, pp. 439–517, 1993.

- [20] H. McKean, *Stochastic Integrals*. New York, NY, USA: Academic, 1969.
- [21] F. C. Park, "The optimal kinematic design of mechanisms" Ph.D. dissertation, Harvard Univ., Cambridge, MA, USA, 1991.
- [22] G. M. T. D'Eleuterio, "Multibody dynamics for space station manipulators: Recursive dynamics of topological chains," Dynacon Enterprises Ltd., Mississauga, ON, Canada, Tech. Rep. SS-3, Jun. 1985
- [23] G. M. T. D'Eleuterio, "Dynamics of an elastic multibody chain: Part C—Recursive dynamics," *Dyn. Stabil. Syst.*, vol. 7, no. 2, pp. 61–89, 1992.
- [24] T. Lee, M. Leok, and N. H. McClamroch, "Global symplectic uncertainty propagation on $\text{SO}(3)$," in *Proc. 47th IEEE Conf. Decision Control*, 2008, pp. 61–66.
- [25] Y. Wang and G. S. Chirikjian, "Error propagation on the Euclidean group with applications to manipulator kinematics," *IEEE Trans. Robot.*, vol. 22, no. 4, pp. 591–602, Aug. 2006.
- [26] Y. Wang and G. S. Chirikjian, "Nonparametric second-order theory of error propagation on motion groups," *Int. J. Robot. Res.*, vol. 27, no. 11, pp. 1258–1273, 2008.
- [27] K. Wolfe, M. Mashner, and G. Chirikjian, "Bayesian fusion on Lie groups," *J. Algebr. Statist.*, vol. 2, no. 1, pp. 75–97, 2011.
- [28] S. Klarsfeld and J. A. Oteo, "The Baker-Campbell-Hausdorff formula and the convergence of the Magnus expansion," *J. Phys. A: Math. Gen.*, vol. 22, pp. 4565–4572, 1989.
- [29] S. Julier and J. Uhlmann, "A general method for approximating nonlinear transformations of probability distributions," Robotics Res. Group, Univ. Oxford, Oxford, U.K., Tech. Rep., 1996
- [30] J. Brookshire and S. Teller, "Extrinsic calibration from per-sensor egomotion," presented at the Robot.: Sci. Syst. Conf., Sydney, Australia, Jul. 2012.
- [31] A. W. Long, K. C. Wolfe, M. J. Mashner, and G. S. Chirikjian, "The banana distribution is Gaussian: A localization study with exponential coordinates," presented at the Robot.: Sci. Syst. Conf., Sydney, Australia, 2012.
- [32] G. Sibley, C. Mei, I. Reid, and P. Newman, "Vast scale outdoor navigation using adaptive relative bundle adjustment," *Int. J. Robot. Res.*, vol. 29, no. 8, pp. 958–980, Jul. 2010.
- [33] P. Mahalanobis, "On the generalized distance in statistics," in *Proc. Nat. Inst. Sci.*, 1936, vol. 2, pp. 49–55.
- [34] H. Strasdat, J. M. M. Montiel, and A. Davison, "Scale drift-aware large scale monocular SLAM," presented at the Robot.: Sci. Syst. Conf., Zaragoza, Spain, Jun. 2010.
- [35] L. Matthies and S. A. Shafer, "Error modeling in stereo navigation," *IEEE J. Robot. Autom.*, vol. RA-3, no. 3, pp. 239–248, Jun. 1987.
- [36] G. Sibley, "Long range stereo data-fusion from moving platforms" Ph.D. dissertation, Dept. Comput. Sci., Univ. Southern California, Los Angeles, CA, USA, 2007.
- [37] W. Triggs, P. McLauchlan, R. Hartley, and A. Fitzgibbon, "Bundle adjustment: A modern synthesis," in *Vision Algorithms: Theory and Practice* (LNCS series), W. Triggs, A. Zisserman, and R. Szeliski, Eds. New York, NY, USA: Springer-Verlag, 2000, pp. 298–375.
- [38] F. C. Park, J. E. Bobrow, and S. R. Ploen, "A Lie group formulation of robot dynamics," *Int. J. Robot. Res.*, vol. 14, pp. 609–618, 1995.
- [39] F. Bullo and R. M. Murray, "Proportional derivative control on the Euclidean group," in *Proc. Eur. Control Conf.*, 1995, pp. 1091–1097.



Timothy D. Barfoot (M'10) received the Ph.D. degree from University of Toronto, Toronto, ON, Canada, in 2002.

He currently leads the Autonomous Space Robotics Lab, University of Toronto, which develops methods to allow mobile robots to operate in large-scale, unstructured, 3-D environments, using rich on-board sensing (e.g., cameras and laser rangefinders) and computation. He is also an Associate Professor with the University of Toronto Institute for Aerospace Studies, where he holds the Canada Research Chair (Tier II) in Autonomous Space Robotics and works in the area of guidance, navigation, and control of mobile robots for space and terrestrial applications. Prior to his current position, he spent four years at MDA Space Missions, where he developed autonomous vehicle navigation technologies for both planetary rovers and terrestrial applications such as underground mining.

Dr. Barfoot is an Ontario Early Researcher Award holder and a licensed Professional Engineer in the Province of Ontario.



Paul T. Furgale (M'11) received the B.A.Sc. degree in computer science from the University of Manitoba, Winnipeg, MB, Canada, in 2006 and the Ph.D. degree from the University of Toronto Institute for Aerospace Studies, Toronto, ON, Canada, 2011.

He is currently the Deputy Director with the Autonomous Systems Lab, Swiss Federal Institute of Technology, Zurich, Switzerland, where he is the scientific coordinator of the European Project V-Charge. His research interests include state estimation, sensor fusion, and calibration for complicated real-world

robotic systems.