



Published in final edited form as:

Int J Rob Res. 2008 ; 27(11-12): 1258–1273. doi:10.1177/0278364908097583.

Nonparametric Second-Order Theory of Error Propagation on Motion Groups

Yunfeng Wang¹ and Gregory S. Chirikjian²

¹Department of Mechanical Engineering, The College of New Jersey, Ewing, NJ 08628.
jwang@tcnj.edu

²Department of Mechanical Engineering, Johns Hopkins University, Baltimore, MD 21218.
gregc@jhu.edu

Abstract

Error propagation on the Euclidean motion group arises in a number of areas such as in dead reckoning errors in mobile robot navigation and joint errors that accumulate from the base to the distal end of kinematic chains such as manipulators and biological macromolecules. We address error propagation in rigid-body poses in a coordinate-free way. In this paper we show how errors propagated by convolution on the Euclidean motion group, $SE(3)$, can be approximated to second order using the theory of Lie algebras and Lie groups. We then show how errors that are small (but not so small that linearization is valid) can be propagated by a recursive formula derived here. This formula takes into account errors to second-order, whereas prior efforts only considered the first-order case. Our formulation is nonparametric in the sense that it will work for probability density functions of any form (not only Gaussians). Numerical tests demonstrate the accuracy of this second-order theory in the context of a manipulator arm and a flexible needle with bevel tip.

Keywords

Recursive error propagation; Euclidean group; spatial uncertainty

1 Introduction

Error propagation on the Euclidean motion group arises in a surprising number of different areas. For example, consider a robotic manipulator for which each joint angle has some backlash. If we describe this backlash as a distribution of possible angles around the nominal one, how will these joint errors add up to produce pose errors at the end effector? Similar problems arise in the study of chainlike biological macromolecules that undergo thermal fluctuations in solution. See, for example, [Zhou and Chirikjian 2006] and [Kim and Chirikjian 2005]. As another example, consider a nonholonomic mobile robot that executes an open loop trajectory. Uncertainties in pose will add up along the path, and if many trials are performed, what will the distribution of terminal poses be? Many such problems in ‘probabilistic robotics’ can be imagined with the recent popularity of SLAM [Thrun, Burgard and Fox 2005].

If the errors are small, Jacobian-based methods or first-order error propagation theories can be used. But what if the errors are very large? Here we address the propagation of large errors in rigid-body poses in a coordinate-free way. In this paper we show how errors propagated by convolution on the Euclidean motion group, $SE(3)$, can be approximated to second order using the theory of Lie algebras and Lie groups. We then show how errors of moderate size (but not

so small that linearization is valid) can be propagated by a recursive formula derived here. This formula takes into account errors to second-order, whereas prior efforts only considered the first-order case. Our formulation is nonparametric in the sense that it will work for probability density functions of any form (not only Gaussians).

In the remainder of this section we review the literature on error propagation, and review the terminology and notation used throughout the paper. In what follows, bold lower case letters denote vectors. N and n are positive integers. G denotes either the groups $SO(3)$ or $SE(3)$. All upper case letters (Roman or Greek) (except for N and G) denote matrices. Lower case letters denote scalars and group elements. A lower case letter followed by parenthesis denotes a scalar-valued function.

In Section 2, important definitions from the basic theory of Lie groups and probability and statistics are reviewed. In Section 3, several new theorems are proved. This forms the core of our paper. In Section 4, sampling is discussed and the theory is adapted for the case when a whole pdf is not available. Then numerical tests demonstrate the accuracy of this recursive second-order propagation formula relative to baseline truth generated by brute force. In Section 5 our conclusions are presented. Three appendices provide more detailed background material that is important for understanding the definitions and proofs presented in the main body of the paper. The remainder of the current section reviews the literature and basic definitions and notation used throughout the paper.

1.1 Literature Review

The Lie-group-theoretic notation and terminology which has now become standard vocabulary in the robotics community is presented in [Murray, Li and Sastry 1994], and Selig [Selig 1996]. In [Chirikjian 2001] many problems in robot kinematics and motion planning were formulated as the convolution of functions on the Euclidean group. The representation and estimation of spatial uncertainty has also received attention in the robotics and vision literature. Two classic works in this area are [Smith and Cheeseman 1986] and [Su and Lee 1992]. Recent work on error propagation describes the concatenation of Gaussian random variables on groups and applies this formalism to mobile robot navigation [Smith, Drummond and Roussopoulos 2003]. In all three of these works, errors are assumed to be small enough that co-variances can be propagated by the formula [Wang and Chirikjian 2006a, Wang and Chirikjian 2006b]

$$\Sigma_{1*2} = Ad(g_2^{-1})\Sigma_1 Ad^T(g_2^{-1}) + \Sigma_2, \quad (1)$$

where Ad is the adjoint operator for $SE(3)$ (See Appendix for a review of terminology). This equation essentially says that given two ‘noisy’ frames of reference $g_1, g_2 \in SE(3)$, each of which is a Gaussian random variable with 6×6 covariance matrices³ Σ_1 and Σ_2 , respectively, the covariance of $g_1 \circ g_2$ will be Σ_{1*2} . This approximation is very good when errors are very small. We extend this linearized approximation to the quadratic terms in the expansion of the matrix exponential parameterization of $SE(3)$. The origin of (1) will become clear for the special case of small errors in our more general nonparametric derivation.

1.2 Review of Rigid-Body Motions

The Euclidean motion group, $SE(3)$, is the semi-direct product of IR^3 with the special orthogonal group, $SO(3)$. We represent elements of $SE(3)$ using 4×4 homogeneous transformation matrices

³Exactly what is meant by a covariance for a Lie group is quantified later in the paper.

$$g = \begin{pmatrix} R & \mathbf{t} \\ 0^T & 1 \end{pmatrix},$$

and identify the group law with matrix multiplication. The inverse of any group element is written as

$$g^{-1} = \begin{pmatrix} R^T & -R^T \mathbf{t} \\ 0^T & 1 \end{pmatrix}.$$

For small translational (rotational) displacements from the identity along (about) the i^{th} coordinate axis, the homogeneous transforms representing infinitesimal motions look like

$$g_i(\epsilon) \triangleq \exp(\epsilon \tilde{E}_i) \approx I_4 + \epsilon \tilde{E}_i$$

where I_4 is the 4×4 identity matrix and

$$\begin{aligned} \tilde{E}_1 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}; & \tilde{E}_2 &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}; & \tilde{E}_3 &= \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}; \\ \tilde{E}_4 &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}; & \tilde{E}_5 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}; & \tilde{E}_6 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

These are related to the basis elements $\{E_i\}$ for $so(3)$ (the Lie algebra corresponding to the rotation group, $SO(3)$) as

$$\tilde{E}_i = \begin{pmatrix} E_i & 0 \\ 0^T & 0 \end{pmatrix}$$

when $i = 1, 2, 3$. Each \tilde{E}_i has a corresponding natural unit basis vector $\mathbf{e}_i \in \mathbb{R}^6$. For example, $\mathbf{e}_1 = [1, 0, 0, 0, 0, 0]^T$, $\mathbf{e}_2 = [0, 1, 0, 0, 0, 0]^T$, etc.

Large motions are also obtained by exponentiating these matrices. For example,

$$\exp(t \tilde{E}_3) = \begin{pmatrix} \cos t & -\sin t & 0 & 0 \\ \sin t & \cos t & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ and } \exp(t \tilde{E}_6) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & t \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

More generally, it can be shown that every element in the neighborhood of the identity of a matrix Lie group G can be described with the exponential parameterization

$$g = g(x_1, x_2, \dots, x_n) = \exp\left(\sum_{i=1}^n x_i E_i\right) \quad (2)$$

where n is the dimension of the group. For $SO(3)$ and $SE(3)$, $n = 3$ and 6 , respectively, and the exponential parameterization extends over the whole group.

One defines the ‘vee’ operator, \vee , such that

$$\left(\sum_{i=1}^n x_i E_i\right)^\vee = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix}$$

The vector, $\mathbf{x} \in \mathbb{R}^n$, can be obtained from $g \in G$ from the formula

$$\mathbf{x} = (\log g)^\vee. \quad (3)$$

For $SO(3)$ and $SE(3)$ this is defined except on a set of measure zero, which for all intents and purposes in the probability and statistics problems that we will consider means that the exponential and logarithm maps are ‘effectively’ bijective. See the Appendix for details.

When integrating a function over $SO(3)$ or $SE(3)$, a weight $w(\mathbf{x})$ is defined such that

$$\int_G f(g) dg = \int_{\mathbb{R}^n} f(g(\mathbf{x})) w(\mathbf{x}) d\mathbf{x}.$$

The exact form of the weighting function is

$$w(\mathbf{x}) = |\det J_r(\mathbf{x})| \quad \text{where} \quad J_r(\mathbf{x}) = \left[\left(g^{-1} \frac{\partial g}{\partial x_1}\right)^\vee, \dots, \left(g^{-1} \frac{\partial g}{\partial x_n}\right)^\vee \right]. \quad (4)$$

This is derived for $SO(3)$ and $SE(3)$ in the Appendix. The weighting function is even in the sense that $w(\mathbf{x}) = w(-\mathbf{x})$.

1.3 The Baker-Campbell-Hausdorff Formula

Given any two elements of a Lie algebra, X and Y , the Lie bracket is defined as $[X, Y] = XY - YX$. An important relationship called the *Baker-Campbell-Hausdorff formula* exists between the Lie bracket and matrix exponential (see [Baker 1904, Campbell 1897, Hausdorff 1906]). Namely, the logarithm of the product of two Lie group elements written as exponentials of Lie algebra elements can be expressed as

$$Z(X, Y) = \log(e^X e^Y)$$

where

$$\begin{aligned}
 Z(X, Y) = & X+Y && \text{zeroth order} \\
 & +\frac{1}{2} [X, Y] && \text{first order} \\
 & +\frac{1}{12} ([X, [X, Y]] + [Y, [Y, X]]) && \text{second order} \\
 & +\frac{1}{48} ([Y, [X, [Y, X]]) + [X, [Y, [Y, X]]) && \text{third order} \\
 & + \dots &&
 \end{aligned} \tag{5}$$

This expression is verified by expanding e^X and e^Y in Taylor series of the form in (37), and then substituting the result into (38) with $g = e^X e^Y$. If the \vee operation is applied (see Appendix for a review), (5) can be written as

$$\begin{aligned}
 \mathbf{z} = & \mathbf{x} + \mathbf{y} + \frac{1}{2} \text{ad}(X) \mathbf{y} + \frac{1}{12} (\text{ad}(X) \text{ad}(X) \mathbf{y} + \text{ad}(Y) \text{ad}(Y) \mathbf{x}) \\
 & + \frac{1}{48} (\text{ad}(Y) \text{ad}(X) \text{ad}(Y) \mathbf{x} + \text{ad}(X) \text{ad}(Y) \text{ad}(Y) \mathbf{x}) + \dots
 \end{aligned}$$

1.4 Probability and Statistics in \mathbb{R}^n : Multivariate Analysis

In \mathbb{R}^n , a probability density function (or pdf) is defined by the conditions

$$f(\mathbf{x}) \geq 0 \quad \forall \mathbf{x} \in \mathbb{R}^n \quad \text{and} \quad \int_{\mathbb{R}^n} f(\mathbf{x}) \, d\mathbf{x} = 1$$

where $d\mathbf{x} = dx_1 dx_2 \dots dx_n$ is the usual Lebesgue integration measure. The mean of a pdf, $f(\mathbf{x})$, is defined as

$$\boldsymbol{\mu} = \int_{\mathbb{R}^n} \mathbf{x} f(\mathbf{x}) \, d\mathbf{x} \quad \text{or} \quad \int_{\mathbb{R}^n} (\mathbf{x} - \boldsymbol{\mu}) f(\mathbf{x}) \, d\mathbf{x} = 0. \tag{6}$$

Note that $\boldsymbol{\mu}$ minimizes the cost function

$$c(\mathbf{x}) = \int_{\mathbb{R}^n} \|\mathbf{x} - \mathbf{y}\|^2 f(\mathbf{y}) \, d\mathbf{y} \tag{7}$$

where $\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$ is the 2-norm in \mathbb{R}^n .

The covariance of the same pdf about the mean is defined as

$$\boldsymbol{\Sigma} = \int_{\mathbb{R}^n} (\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T f(\mathbf{x}) \, d\mathbf{x}. \tag{8}$$

It follows that

$$C = \int_{\mathbb{R}^n} \mathbf{x}\mathbf{x}^T f(\mathbf{x}) \, d\mathbf{x} = \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T \tag{9}$$

where C is the *covariance about the origin* and $\boldsymbol{\Sigma}$ is the *covariance about the mean*.

Pdfs are often used to describe distributions of errors. If these errors are concatenated, they ‘add’ by convolution:

$$(f_1 * f_2)(\mathbf{x}) = \int_{\mathbb{R}^n} f_1(\boldsymbol{\xi}) f_2(\mathbf{x} - \boldsymbol{\xi}) \, d\boldsymbol{\xi}. \tag{10}$$

The mean and covariance of convolved distributions are found as

$$\mu_{1*2} = \mu_1 + \mu_2 \quad \text{and} \quad \Sigma_{1*2} = \Sigma_1 + \Sigma_2 \tag{11}$$

In other words, these quantities can be propagated without explicitly performing the convolution computation, or even knowing the full pdfs. This is independent of the parametric form of the pdf. Often one does not have access to the full pdf, but only samples from a process with an underlying pdf. In this case, the unbiased sample mean and covariance are defined as [Anderson 2005]:

$$\mu^{(N)} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i \quad \text{and} \quad \Sigma^{(N)} = \frac{1}{N-1} \sum_{i=1}^N (\mathbf{x}_i - \mu^{(N)}) (\mathbf{x}_i - \mu^{(N)})^T.$$

The reason for division by $N - 1$ rather than N is explained in texts on multivariate analysis such as [Anderson 2005]. As the sample size becomes large, the difference between N and $N - 1$ becomes negligible and these sampled quantities converge to those corresponding to the underlying pdf.

Our main purpose in this paper is to develop equations analogous to (11) to describe the propagation of error on the motion group $SE(3)$. In the process we will also do so for the rotation group $SO(3)$.

It is often convenient to use the Gaussian (or normal) distribution: to model errors in \mathbb{R}^n . This parametric distribution is completely defined by its mean and covariance. We will have no need to assume that densities are Gaussian. Our results are nonparametric, and therefore more general.

2 Definitions and Properties of Mean and Covariance on $SE(3)$

In this section we provide definitions of the mean and covariance of Lie-group-valued functions and illustrate some of their properties. We note in passing that a pdf that is a symmetric function, $\rho(g) = \rho(g^{-1})$, always satisfies the condition

$$\int_G (\log g)^\vee \rho(g) dg = 0, \tag{12}$$

for $G = SO(3)$ or $G = SE(3)$. This is easy to see if we let $\rho_0(\mathbf{x}) = \rho(e^{\mathbf{X}})$. Then $\rho_0(\mathbf{x}) = \rho_0(-\mathbf{x})$. This is an even function in the exponential coordinates, and so the odd function $\mathbf{x}\rho_0(\mathbf{x})$ integrates to zero over a symmetric domain of integration in the space of exponential parameters that maps to G . See appendix for discussion of integration measures. In our case this domain is the ball of radius π (for $SO(3)$), or the Cartesian product of this ball with \mathbb{R}^3 . Both of which are symmetric. Hence the integral in (12) vanishes. More generally, if $\rho(g)$ is a symmetric function on $G = SO(3)$ or $G = SE(3)$ then for $\sum_{i=1}^n n_i$ odd,

$$\int_G \prod_{i=1}^n [(\log g)^\vee \cdot \mathbf{e}_i]^{n_i} \rho(g) dg = 0. \tag{13}$$

This is because the integrand is an odd function of the components of \mathbf{x} . For example,

$$\int_G (\log g)^{2k+1} \rho(g) dg = 0_n.$$

DEFINITION 1: If a unique value $\mu \in G$ exists for which

$$\int_G [\log(\mu^{-1} \circ g)]^\vee f(g) dg = 0, \tag{14}$$

μ will be called the *mean* of a pdf $f(g)$ on G , which is a straightforward extension of (6). Furthermore, the *covariance about the mean* will be computed as

$$\Sigma = \int_G \log(\mu^{-1} \circ g)^\vee [\log(\mu^{-1} \circ g)^\vee]^T f(g) dg. \tag{15}$$

Note that while in the case of Euclidean space (6) and minimization of (7) both give the same value of the mean, the minimization of a functional of the form

$$c(h) = \int_G \|\log(h^{-1} \circ g)\|^\vee{}^2 f(g) dg$$

does not generally return a value h_{min} that is equal to μ . However, in the special case when $f(g)$ is unimodal and very concentrated, $h_{min} \approx \mu$.

The equality (12) can be thought of as a statement of when the mean is at the identity. If $\rho(g)$ has mean at the identity, then $f(g) = \rho(a^{-1} \circ g)$ has mean at a . We will use $\rho(g)$ to denote pdfs with mean at the identity, and $f(g)$ to denote pdfs that can have the mean at some other group element.

THEOREM 1: If $f(g)$ has mean μ and covariance Σ , then to second order

$$\mathbf{m} = [I + F_1(\Sigma)]\mu \tag{16}$$

where the following shorthand is used:

$$\mathbf{m} = \int_G (\log(g))^\vee f(g) dg \quad \text{and} \quad \mu = (\log(\mu))^\vee, \tag{17}$$

and the matrix-valued function $F_1(\Sigma)$ is defined as

$$F_1(\Sigma) = \frac{1}{12} \sum_{i,j=1}^6 \sigma_{ij} ad(\tilde{E}_i) ad(\tilde{E}_j) \tag{18}$$

and

$$\begin{aligned} C &= \int_G (\log(g))^\vee ((\log(g))^\vee)^T f(g) dg \\ &= \Sigma + (\log \mu)^\vee ((\log \mu)^\vee)^T + \frac{1}{2} (\Sigma ad^T(\log \mu) + ad(\log \mu) \Sigma). \end{aligned} \tag{19}$$

Here C is the *covariance about the identity*, which is defined in an analogy with the concept of covariance about the origin in the context of probability and statistics in IR^n .

Proof: Let $f(g) = \rho(\mu^{-1} \circ g)$ where $\rho(g)$ has mean at the identity. Then

$$\int_G (\log g)^\vee f(g) dg = \int_G (\log g)^\vee \rho(\mu^{-1} \circ g) dg = \int_G (\log(\mu \circ g))^\vee \rho(g) dg.$$

Expanding using the BCH formula (5) with $\mu = \exp X$ and $g = \exp Y$, and using the linearity of the Lie bracket, we find that since $\rho(g)$ is a pdf with mean at the identity,

$$\mathbf{x} + \sum_{i,j=1}^6 \sigma_{ij} \left\{ \frac{1}{12} \left[\tilde{E}_i, \left[\tilde{E}_j, X \right] \right] + \frac{1}{48} \left(\left[\tilde{E}_i, \left[X, \left[\tilde{E}_j, X \right] \right] \right] + \left[X, \left[\tilde{E}_i, \left[\tilde{E}_j, X \right] \right] \right] \right) \right\}^\vee + \dots$$

The first expression in the statement of the theorem results from the definition of the adjoint. Likewise,

$$\int_G [\log(\mu \circ g)]^\vee ([\log(\mu \circ g)]^\vee)^T \rho(g) dg = \int_G \left[\mathbf{x} + \mathbf{y} + \frac{1}{2} \text{ad}(X) \mathbf{y} + \dots \right] \left[\mathbf{x} + \mathbf{y} + \frac{1}{2} \text{ad}(X) \mathbf{y} + \dots \right]^T \rho(g) dg$$

Expanding out the product and eliminating terms linear in \mathbf{y} results in the second statement of the theorem.

3 Propagation of the Mean and Covariance of PDFs on SE(3)

Let $\mu_1, \mu_2 \in SE(3)$ be two precise reference frames. Then $\mu_{10} \circ \mu_2$ is the frame resulting from stacking one relative to the other. Now suppose that each has some uncertainty. Let $\{h_i\}$ and $\{k_j\}$ be two sets of frames of reference that are distributed around the identity. Let the first have N_1 elements, and the second have N_2 . What will the covariance of the set of $N_1 \cdot N_2$ frames $\{(\mu_1 \circ \mu_2)^{-1} \circ \mu_1 \circ h_i \circ \mu_2 \circ k_j\}$ (which are also distributed around the identity) look like ?

Let $\rho_i(g)$ be a unimodal pdf with mean at the identity and which has a preponderance of its mass concentrated in a unit ball around the identity (where distance from the identity is measured as $\|(\log g)^\vee\|$). Then $\rho_i(\mu_i^{-1} \circ g)$ will be a distribution with the same shape centered at μ_i . In general, the convolution of two pdfs is defined as

$$(f_1 * f_2)(g) = \int_G f_1(h) f_2(h^{-1} \circ g) dh,$$

and in particular if we make the change of variables $k = \mu_1^{-1} \circ h$, then

$$\rho(\mu_1^{-1} \circ g) * \rho_2(\mu_2^{-1} \circ g) = \int_G \rho_1(k) \rho_2(\mu_2^{-1} \circ k^{-1} \circ \mu_1^{-1} \circ g) dk.$$

Making the change of variables $g = \mu_1 \circ \mu_2 \circ q$, where q is a relatively small displacement measured from the identity, the above can be written as

$$\rho_{1*2}(\mu_1 \circ \mu_2 \circ q) = \int_G \rho_1(k) \rho_2(\mu_2^{-1} \circ k^{-1} \circ \mu_2 \circ q) dk. \tag{20}$$

The essence of this paper is the efficient approximation of covariances about the mean of ρ_{1*2} in (20) when the covariances about the means of ρ_1 and ρ_2 are known. In cases when $\mu_{1*2} = \mu_1 \circ \mu_2$, the problem reduces to the efficient approximation of

$$\begin{aligned} \Sigma_{1*2} &= \int_G \int_G \log(q)^\vee [\log(q)^\vee]^T \rho_1(k) \rho_2(\mu_2^{-1} \circ k^{-1} \circ \mu_2 \circ q) dk dq \\ &= \int_G \int_G \log(\mu_2^{-1} \circ k \circ \mu_2 \circ q')^\vee [\log(\mu_2^{-1} \circ k \circ \mu_2 \circ q')^\vee]^T \rho_1(k) \rho_2(q') dk dq'. \end{aligned} \tag{21}$$

LEMMA 1: The convolution of pdfs with mean at the identity results (to second order) in a pdf with mean at the identity. Furthermore, if $\rho_1 * \rho_2 = \rho_2 * \rho_1$ and $\rho_i(g) = \rho_i(g^{-1})$, then this result becomes exact.

Proof:

$$\begin{aligned} \int_G (\log g) (\rho_1 * \rho_2)(g) dg &= \int_G \int_G (\log g) \rho_1(h) \rho_2(h^{-1} \circ g) dh dg \\ &= \int_G \int_G \log(h \circ k) \rho_1(h) \rho_2(k) dh dk. \end{aligned}$$

To second order, all terms in the BCH expansion of $\log(h \circ k)$ are linear in either $\log h$ or $\log k$ (or both), and therefore at least one of the above integrals integrates to zero.

If $\rho_1 * \rho_2 = \rho_2 * \rho_1$ and $\rho_i(g) = \rho_i(g^{-1})$ then it is easy to show that $(\rho_1 * \rho_2)(g) = (\rho_1 * \rho_2)(g^{-1})$, which automatically means that the function $(\rho_1 * \rho_2)(g)$ has mean at the identity.

THEOREM 2: If $f_i(g)$ has mean μ_i and covariance Σ_i for $i = 1, 2$, then to second order, the mean and covariance of $(f_1 * f_2)(g)$ are respectively

$$\mu_{1*2} = \mu_1 \circ \mu_2 \tag{22}$$

and

$$\Sigma_{1*2} = A + B + F(A, B) \tag{23}$$

where

$$F(A, B) = \frac{1}{4} C(A, B) + \frac{1}{12} \left[A'' B + (A'' B)^T + B'' A + (B'' A)^T \right],$$

$A = Ad(\mu_2^{-1}) \Sigma_1 Ad^T(\mu_2^{-1})$, $B = \Sigma_2$, and $C(A, B)$ and A'' are computed as follows:

$$A'' = \begin{pmatrix} A_{11} - \text{tr}(A_{11}) I_3 & 0_3 \\ A_{12} + A_{12}^T - 2\text{tr}(A_{12}) I_3 & A_{11} - \text{tr}(A_{11}) I_3 \end{pmatrix}.$$

B'' is defined in the same way with B replacing A everywhere in the expression. The blocks of C are computed as

$$\begin{aligned} C_{11} &= -D_{11,11} \\ C_{12} &= -(D_{21,11})^T - D_{11,12} = C_{21} \\ C_{22} &= -D_{22,11} - D_{21,21} - (D_{21,12})^T - D_{11,22} \end{aligned}$$

where $D_{ij,kl} = D(A_{ij}, B_{kl})$, and the matrix-valued function $D(A', B')$ is defined relative to the entries in the 3×3 blocks A' and B' as

$$\begin{aligned} d_{11} &= -a'_{33}b'_{22} + a'_{31}b'_{32} + a'_{23}b'_{23} - a'_{22}b'_{33} \\ d_{12} &= a'_{33}b'_{21} - a'_{32}b'_{31} - a'_{13}b'_{23} + a'_{21}b'_{33} \\ d_{13} &= -a'_{23}b'_{21} + a'_{22}b'_{31} + a'_{13}b'_{22} - a'_{12}b'_{32} \\ d_{21} &= a'_{33}b'_{12} - a'_{31}b'_{32} - a'_{21}b'_{13} + a'_{21}b'_{33} \\ d_{22} &= -a'_{33}b'_{11} + a'_{31}b'_{31} + a'_{13}b'_{13} - a'_{11}b'_{33} \\ d_{23} &= a'_{23}b'_{11} - a'_{21}b'_{31} - a'_{13}b'_{12} + a'_{11}b'_{32} \\ d_{31} &= -a'_{32}b'_{12} + a'_{31}b'_{22} + a'_{22}b'_{13} - a'_{21}b'_{23} \\ d_{32} &= a'_{32}b'_{11} - a'_{31}b'_{21} - a'_{12}b'_{13} + a'_{11}b'_{23} \\ d_{33} &= -a'_{22}b'_{11} + a'_{21}b'_{21} + a'_{12}b'_{12} - a'_{11}b'_{22} \end{aligned}$$

Proof: The approximation in (22) follows directly from Lemma 1. Next, let

$X = \log(\mu_2^{-1} \circ k \circ \mu_2)$ where $k = \exp K$, and let $Y = \log q'$. Using the Baker-Campbell-Hausdorff formula (5) to evaluate the log terms in the definition of covariance, and retaining all even terms to second order (since first order terms will integrate to zero), we get

$$\begin{aligned} \left\{ (Z(X, Y))^{\vee} \left[(Z(X, Y))^{\vee} \right]^T \right\}_{\text{even}} &= \mathbf{x}\mathbf{x}^T + \mathbf{y}\mathbf{y}^T + \frac{1}{4} [X, Y]^{\vee} ([X, Y]^{\vee})^T \\ &+ \frac{1}{12} \mathbf{y} ([X, [X, Y]]^{\vee})^T + \frac{1}{12} \mathbf{x} ([Y, [Y, X]]^{\vee})^T \\ &+ \frac{1}{12} [X, [X, Y]]^{\vee} \mathbf{y}^T + \frac{1}{12} [Y, [Y, X]]^{\vee} \mathbf{x}^T \end{aligned} \tag{24} \tag{25}$$

Each of these terms can be expanded using the adjoint concept. For example,

$$[X, Y]^{\vee} ([X, Y]^{\vee})^T = \text{ad}(X) \mathbf{y}\mathbf{y}^T \text{ad}^T(X) \quad \text{and} \quad [X, [X, Y]]^{\vee} \mathbf{y}^T = \text{ad}(X) \text{ad}(X) \mathbf{y}\mathbf{y}^T. \tag{26}$$

In our formulation, $X = \mu_2^{-1} K \mu_2$ (where $k = e^K$ and so $\mu_2^{-1} \circ k \circ \mu_2 = \exp(\mu_2^{-1} K \mu_2) = e^X$). Defining the vector $\mathbf{x} = \text{Ad}(\mu_2^{-1}) \mathbf{k}$, then

$$\begin{aligned} A &= \int_G \mathbf{x}\mathbf{x}^T \rho_1(k) dk \\ &= \text{Ad}(\mu_2^{-1}) \left[\int_G (\log k)^{\vee} [(\log k)^{\vee}]^T \rho_1(k) dk \right] \text{Ad}^T(\mu_2^{-1}) \\ &= \text{Ad}(\mu_2^{-1}) \Sigma_1 \text{Ad}^T(\mu_2^{-1}) \end{aligned}$$

and since $q' = e^Y$,

$$B = \int_G \mathbf{y} \mathbf{y}^T \rho_2(q') dq' = \int_G (\log q')^\vee [(\log q')^\vee]^T \rho_2(q') dq' = \Sigma_2.$$

The following complicated looking integral (which is nothing more than (21) written in exponential coordinates)

$$\begin{aligned} \Sigma_{1*2} &= \int_{q' \in G} \int_{k \in G} \left\{ (Z(X, Y))^\vee [(Z(X, Y))^\vee]^T \right\} \rho_1(k) \rho_2(q') dk dq' \\ &= \int_{q' \in G} \int_{k \in G} \left\{ (Z(X, Y))^\vee [(Z(X, Y))^\vee]^T \right\}_{\text{even}} \rho_1(k) \rho_2(q') dk dq' \end{aligned}$$

can be simplified. This is because

$$x_i x_j = \mathbf{e}_i^T Ad(\mu_2^{-1}) (\log k)^\vee [(\log k)^\vee]^T Ad^T(\mu_2^{-1}) \mathbf{e}_j$$

and

$$y_k y_l = \mathbf{e}_k^T (\log q')^\vee [(\log q')^\vee]^T \mathbf{e}_l,$$

and since all terms in $\{(Z(X, Y))^\vee [(Z(X, Y))^\vee]^T\}_{\text{even}}$ can be expressed as weighted sums of such products, it follows that after integration we get

$$\Sigma_{1*2} = A + B + F(A, B). \tag{27}$$

For the $SE(3)$ case

$$Ad(g) = \begin{pmatrix} R & 0_3 \\ TR & R \end{pmatrix} \in IR^{6 \times 6} \quad \text{and} \quad ad(X) = \begin{pmatrix} \Omega & 0_3 \\ V & \Omega \end{pmatrix} \in IR^{6 \times 6}$$

where $T^\vee = \mathbf{t}$, $V^\vee = \mathbf{v}$ and $\Omega^\vee = \boldsymbol{\omega}$. Then (26) becomes:

$$\begin{aligned} [X, Y]^\vee [(X, Y)^\vee]^T &= \begin{pmatrix} \Omega_x & 0_3 \\ V_x & \Omega \end{pmatrix} \begin{pmatrix} \omega_y \\ \mathbf{v}_y \end{pmatrix} [\omega_y^T, \mathbf{v}_y^T] \begin{pmatrix} -\Omega_x & -V_x \\ 0_3 & -\Omega \end{pmatrix} \\ &= \begin{pmatrix} \Omega_x & 0_3 \\ V_x & \Omega \end{pmatrix} \begin{pmatrix} \omega_y \omega_y^T & \omega_y \mathbf{v}_y^T \\ \mathbf{v}_y \omega_y^T & \mathbf{v}_y \mathbf{v}_y^T \end{pmatrix} \begin{pmatrix} -\Omega_x & -V_x \\ 0_3 & -\Omega \end{pmatrix} \end{aligned}$$

and

$$\begin{aligned} [X, [X, Y]]^\vee \mathbf{y}^T &= \begin{pmatrix} \Omega_x & 0_3 \\ V_x & \Omega \end{pmatrix} \begin{pmatrix} \Omega_x & 0_3 \\ V_x & \Omega \end{pmatrix} \begin{pmatrix} \omega_y \\ \mathbf{v}_y \end{pmatrix} [\omega_y^T, \mathbf{v}_y^T] \\ &= \begin{pmatrix} \Omega_x & 0_3 \\ V_x & \Omega \end{pmatrix} \begin{pmatrix} \Omega_x & 0_3 \\ V_x & \Omega \end{pmatrix} \begin{pmatrix} \omega_y \omega_y^T & \omega_y \mathbf{v}_y^T \\ \mathbf{v}_y \omega_y^T & \mathbf{v}_y \mathbf{v}_y^T \end{pmatrix} \end{aligned}$$

If we divide the 6×6 symmetric matrices $A = Ad(\mu_2^{-1})\Sigma_1 Ad^T(\mu_2^{-1})$ and $B = \Sigma_2$ into 3×3 blocks as

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^T & A_{22} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} B_{11} & B_{12} \\ B_{12}^T & B_{22} \end{pmatrix},$$

then using the specific form of $ad(X)$ and integrating over q' we get

$$\int_G [X, [X, Y]]^\vee \mathbf{y}^T \rho_2(q') dq' = \begin{pmatrix} \Omega_x^2 & 0_3 \\ V_x \Omega_x + \Omega_x V_x & \Omega_x^2 \end{pmatrix} B$$

and

$$\int_G [X, Y]^\vee ([X, Y]^\vee)^T \rho_2(q') dq' = \begin{pmatrix} -\Omega_x B_{11} \Omega_x & -(V_x B_{11} \Omega_x)^T - \Omega_x B_{12} \Omega_x \\ -V_x B_{11} \Omega_x - (\Omega_x B_{12} \Omega_x)^T & -V_x B_{11} \Omega_x - V_x B_{12} \Omega_x - (V_x B_{12} \Omega_x)^T - \Omega_x B_{12} \Omega_x \end{pmatrix}$$

Then integrating over $k \in G$ gives

$$\int_G \int_G [X, [X, Y]]^\vee \mathbf{y}^T \rho_1(k) \rho_2(q') dk dq' = A'' B$$

$$\int_G \int_G [X, Y]^\vee ([X, Y]^\vee)^T \rho_1(k) \rho_2(q') dk dq' = C(A, B) = \begin{pmatrix} C_{11} & C_{12} \\ C_{12}^T & C_{22} \end{pmatrix}.$$

The BCH formula yields several such terms, each of which can be obtained by either transposing those given above or switching the roles of B and A .

4 Sampled Distributions and Numerical Examples

Evaluating the robustness of the first-order (1) and the second-order (23) covariance propagation formula over a wide range of kinematic errors is essential to understand effectiveness of these formulas. In this section, we test these two covariance propagation formulas with concrete numerical examples.

In many practical situations, discrete data are sampled from ρ_1 and ρ_2 rather than having complete knowledge of the distributions themselves. Therefore, sampled covariances can be computed by making the following substitutions:

$$\rho_1(g) = \sum_{i=1}^{N_1} \alpha_i \Delta(h_i^{-1} \circ g) \quad (28)$$

and

$$\rho_2(g) = \sum_{j=1}^{N_2} \beta_j \Delta(k_j^{-1} \circ g) \quad (29)$$

where

$$\sum_{i=1}^{N_1} \alpha_i = \sum_{j=1}^{N_2} \beta_j = 1.$$

Here $\Delta(g)$ is the Dirac delta function for the group G , which has the properties

$$\int_G f(g) \Delta(h^{-1} \circ g) dg = f(h) \quad \text{and} \quad \Delta(h^{-1} \circ g) = \Delta(g^{-1} \circ h).$$

Using these properties, if we substitute (28) and (29) into (21), the result is

$$\Sigma_{1*2} = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \alpha_i \beta_j \log(\mu_2^{-1} \circ h_i \circ \mu_2 \circ k_j) \left[\log(\mu_2^{-1} \circ h_i \circ \mu_2 \circ k_j) \right]^{\vee T}. \quad (30)$$

While this equation is exact, it has the drawback of requiring $O(N_1 \cdot N_2)$ arithmetic operations. In the first-order theory of error propagation, we made the approximation

$$\log(k^{-1} \circ q) = X - Y,$$

or equivalently

$$\left[\log(k^{-1} \circ q) \right]^{\vee} = \mathbf{x} - \mathbf{y},$$

where $k = \exp Y$ and $q = \exp X$ are elements of the Lie group $SE(3)$. This decouples the summations and makes the computation $O(N_1 + N_2)$. However, the first-order theory breaks down for large errors. Therefore, we explore here the numerical accuracy of the second-order theory developed in the previous section.

4.1 Error Propagation in a PUMA Manipulator

Consider a spatial serial manipulator, PUMA 560. The link-frame assignments of PUMA 560 for D-H parameters is the same as those given in [Craig 2005]. Table 1 lists the D-H parameters of PUMA 560, where $a_2 = 431.8$ mm, $a_3 = 20.32$ mm, $d_3 = 124.46$ mm, and $d_4 = 431.8$ mm. The solution of forward kinematics is the homogeneous transformations of the relative displacements from one D-H frame to another multiplied sequentially.

In order to test these covariance propagation formulas, we first need to create some kinematic errors. Since joint angles are the only variables of the PUMA 560, we assume that errors exist only in these joint angles. We generated errors by deviating each joint angle from its ideal value with uniform random absolute errors of $\pm\epsilon$. Therefore, each joint angle was sampled at three values: $\theta_i - \epsilon$, θ_i , $\theta_i + \epsilon$. This generates $N = 3^6$ different frames of references $\{g_{ee}^i\}$ that are clustered around desired g_{ee} . Here g_{ee} denotes the position and orientation of the distal end of the manipulator relative to the base in the form of homogeneous transformation matrix.

It is important to note that while the cloud of frames $\{g_{ee}^i\}$ is clustered around g_{ee} , it may not be the case that g_{ee} is actually the mean of this cloud. In the first-order theory, the cloud is assumed to be so tightly focused around g_{ee} that the approximation $\mu_{ee} \approx g_{ee}$ can be made

without causing significant errors. However, in the second order theory, one needs to be more precise. We can update our estimate of the mean as:

$$\mu_{ee} = g_{ee} \circ \exp \left[\frac{1}{N} \sum_{i=1}^N \log \left(g_{ee}^{-1} \circ g_{ee}^i \right) \right]. \quad (31)$$

In practice, for errors of moderate magnitude, only one such update is required to obtain the exact mean. For very large errors this formula can be iterated with the output, μ_{ee} , from one iteration serving as the input, g_{ee} , for the next iteration. A similar update to obtain μ_1 and μ_2 from the frame clouds around the frames g_1 and g_2 (the relative frames from base to mid point and mid point to distal end of the manipulator such that $g_1 \circ g_2 = g_{ee}$) should also be performed.

Three different methods for computing the same error covariances for the whole manipulator are computed. The first is to apply brute force enumeration, which gives the actual covariance of the whole manipulator:

$$\Sigma = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i \mathbf{x}_i^T \quad (32)$$

where $\mathbf{x}_i = \left[\log \left(\mu_{ee}^{-1} \circ g_{ee}^i \right) \right]^V$, and (32) is used for all of the 3^6 different frames of references $\{g_{ee}^i\}$. The second method is to apply the first-order propagation formula (1). The third is to apply the second-order propagation formula (23). For the covariance propagation methods, we only need to find the mean and covariance of each individual link. Then the covariance of the whole manipulator can be recursively calculated using the corresponding propagation formula.

In order to quantify the robustness of the two covariance approximation methods, we define a measure of deviation of results between the first/second order formula and the actual covariance using the Hilbert-Schmidt (Frobenius) norm as

$$\text{deviation} = \frac{\|\Sigma_{prop} - \Sigma_{actual}\|}{\|\Sigma_{actual}\|}, \quad (33)$$

where Σ_{prop} is the covariance of the whole manipulator calculated using either the first-order (1) or the second-order (23) propagation formula, Σ_{actual} is the actual covariance of the whole manipulator calculated using (32), and $\|\cdot\|$ denotes the Hilbert-Schmidt (Frobenius) norm.

With all the above information, we now can conduct the specific computation and analysis. The results of two different configurations of the manipulator are illustrated here. The ideal joint angles of one configuration from θ_1 to θ_6 were taken as $[0, \pi/2, -\pi/2, 0, 0, \pi/2]$. The ideal joint angles of the other configuration were taken as $[\pi/4, \pi/5, -\pi/4, \pi/10, \pi/8, \pi]$. As an initial test, the joint angle errors ϵ were taken from 0.1 rad to 0.6 rad, and the static DH parameters of the links were assumed to be error free. The covariances of the whole manipulator corresponding to these kinematic errors were then calculated through the three aforementioned methods. The resulting deviations between the covariance matrices computed directly using (32) and the first-order and second-order propagation formulas are plotted in Figures 1 and 2 with (33) on the y-axis for different amounts of noise on the x-axis.

Since physical manipulators cannot be manufactured with exact design parameters, and their real linkage parameters such as the static DH parameters (α_i, a_i, d_i) may have errors, the propagation theory is applied now to the case with both joint angle errors and linkage errors.

The same sets of calculations that were conducted for the case with only joint angles are now conducted for this case with the additional linkage errors. Our numerical simulations have shown that if the only static DH parameters that have errors are the translational parameters a_i and d_i , then they have essentially no effect the value of the deviation. In other words, both the first and second order propagation formulas capture the covariances resulting from these translational errors. However, the linkage errors in the angular DH parameters such as α_i create observable effects on the accuracy of the propagation formulas. In the given example, we assume that DH parameters α_0 , α_1 , and α_5 deviate from their ideal values with uniform random absolute errors of ± 0.2 rad. Therefore, they are sampled at three values: $\alpha_i - 0.2$, α_i , $\alpha_i + 0.2$. Together with the six joint angle errors, this generates $N = 3^9$ different frames of references $\{g_{ee}^i\}$ that are clustered around the baseline g_{ee} . The results of the first-order and second-order propagation formula of these cases were also graphed in Figures 1 and 2.

The numerical simulation results demonstrate that the propagation formula can efficiently deal with all kinematic errors including errors in joint angles and linkage parameters. It is also clear that the second-order propagation formula makes significant improvements in terms of accuracy when compared to the first-order formula. The second-order propagation theory is much more robust than the first-order formula over a wide range of kinematic errors. These two methods both work well for small errors, and deviate from the actual value more and more as the errors become large. However, the deviation of the first-order formula grows rapidly and breaks down while the second-order propagation method still retains a reasonable value.

To give the readers a sense of what these covariances look like, we listed the values of the covariance of the whole manipulator for the joint angle error $\epsilon = 0.3$ rad below.

The ideal pose of the end effector can be found easily via forward kinematics to be

$$g_{ee} = \begin{pmatrix} 0.0000 & -1.0000 & 0 & 0.0203 \\ -1.0000 & -0.0000 & 0 & 0.1245 \\ 0 & 0 & -1.0000 & -0.08636 \\ 0 & 0 & 0 & 1.0000 \end{pmatrix}.$$

The actual covariance of the whole manipulator calculated using equations (32) is

$$\Sigma_{actual} = \begin{pmatrix} 0.1748 & 0.0000 & 0.0000 & 0.0000 & -0.0755 & -0.0024 \\ 0.0000 & 0.0078 & 0.0000 & 0.0034 & -0.0000 & 0.0003 \\ 0.0000 & 0.0000 & 0.1747 & 0.0012 & -0.0072 & -0.0000 \\ 0.0000 & 0.0034 & 0.0012 & 0.0025 & -0.0001 & 0.0001 \\ -0.0755 & -0.0000 & -0.0072 & -0.0001 & 0.0546 & 0.0015 \\ -0.0024 & 0.0003 & -0.0000 & 0.0001 & 0.0015 & 0.0011 \end{pmatrix},$$

the covariance using the first-order propagation formula (1) is

$$\Sigma_{prop, 1st} = \begin{pmatrix} 0.1800 & 0.0000 & 0 & 0.0000 & -0.0777 & -0.0024 \\ 0.0000 & 0.0000 & 0 & 0.0000 & -0.0000 & -0.0000 \\ 0 & 0 & 0.1800 & 0.0012 & -0.0075 & 0 \\ 0.0000 & 0.0000 & 0.0012 & 0.0000 & -0.0002 & -0.0000 \\ -0.0777 & -0.0000 & -0.0075 & -0.0002 & 0.0569 & 0.0016 \\ -0.0024 & -0.0000 & 0 & -0.0000 & 0.0016 & 0.0000 \end{pmatrix},$$

and the covariance using the second-order propagation formula (23) is

$$\Sigma_{prop, 2nd} = \begin{pmatrix} 0.1748 & 0.0000 & 0.0000 & 0.0000 & -0.0743 & -0.0024 \\ 0.0000 & 0.0079 & 0.0000 & 0.0034 & -0.0000 & 0.0003 \\ 0.0000 & 0.0000 & 0.1747 & 0.0012 & -0.0072 & 0.0000 \\ 0.0000 & 0.0034 & 0.0012 & 0.0025 & -0.0001 & 0.0001 \\ -0.0743 & -0.0000 & -0.0072 & -0.0001 & 0.0546 & 0.0015 \\ -0.0024 & 0.0003 & 0.0000 & 0.0001 & 0.0015 & 0.0011 \end{pmatrix}.$$

4.2 Continuous-Time Covariance Propagation: The Stochastic Flexible Needle with Bevel Tip

The previous example in this paper illustrated how to obtain the mean and covariance of error pdfs resulting from convolutions of densities centered around discrete joints in a manipulator arm. In contrast, applications such as SLAM can be better described with a model in which the error accumulates continuously over time. This section addresses that problem. In particular, estimates of the mean and covariance of a process described by a stochastic differential equation can be obtained for small time intervals by numerical integration. The second-order propagation formulas derived earlier in the paper are then used to propagate these estimates for larger values of time. The example that is used to illustrate this technique is flexible needle steering.

Recently, a number of works have been concerned with the steering of flexible needles with bevel tips through soft tissue for minimally invasive medical treatments. See, for example, [Webster et al. 2006], [Park et al. 2005], [Alterovitz, Simeon and Goldberg 2007]. In this problem, a flexible needle is rotated with the angular speed $\omega(t)$ around its tangent while it is inserted with translational speed $v(t)$ in the tangential direction. Due to the bevel tip, the needle will not follow a straight line when $\omega(t) = 0$ and $v(t)$ is constant. Rather, in this case the tip of the needle will approximately follow a circular arc with curvature κ when the medium is very firm and the needle is very flexible. The specific value of the constant κ depends on parameters such as the angle of the bevel, how sharp the needle is, and properties of the tissue. In practice κ is fit to experimental observations of the needle bending in a particular medium during insertions with $\omega(t) = 0$ and $v(t)$ is constant. Using this as a baseline, and building in arbitrary $\omega(t)$ and $v(t)$, a nonholonomic kinematic model then predicts the time evolution of the position and orientation of the needle tip [Webster et al. 2006, Park et al. 2005].

In a reference frame attached to the needle tip with the local x_3 axis denoting the tangent to the “backbone curve” of the needle, and x_1 denoting the axis orthogonal to the direction of infinitesimal motion induced by the bevel (i.e., the needle bends in the $x_2 - x_3$ plane), the nonholonomic kinematic model for the evolution of the frame at the needle tip was developed in [Webster et al. 2006, Park et al. 2005] as:

$$\dot{\xi} = (g^{-1} \dot{g})^V = [\kappa \ 0 \ \omega(t) \ 0 \ 0 \ v(t)]^T. \quad (34)$$

If everything were certain, and if this model were exact, then $g(t)$ could be obtained by simply integrating the ordinary differential equation in (34). However, in practice a needle that is repeatedly inserted into a medium such as gelatin (which is used to simulate soft tissue) will demonstrate an ensemble of slightly different trajectories.

A simple stochastic model for the needle is obtained by letting [Park et al. 2005, Park et al. 2008]:

$$\omega(t) = \omega_0(t) + \lambda_1 w_1(t),$$

and

$$v(t) = v_0(t) + \lambda_2 w_2(t).$$

Here $\omega_0(t)$ and $v_0(t)$ are what the inputs would be in the ideal case, $w_1(t)$ and $w_2(t)$ are uncorrelated unit Gaussian white noises, and λ_i are constants.

Thus, a nonholonomic needle model with noise is

$$\begin{pmatrix} g^{-1} \dot{g} \end{pmatrix}^\vee dt = \begin{bmatrix} \kappa \\ 0 \\ \omega_0(t) \\ 0 \\ 0 \\ v_0(t) \end{bmatrix} dt + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \lambda_1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} dW_1 \\ dW_2 \end{bmatrix} \quad (35)$$

where $dW_i = W_i(t+dt) - W_i(t) = w_i(t)dt$ are the non-differentiable increments of a Wiener process $W_i(t)$. This noise model is a stochastic differential equation (SDE) on $SE(3)$. As shorthand, we write this as

$$\begin{pmatrix} g^{-1} \dot{g} \end{pmatrix}^\vee dt = \mathbf{h}(t) dt + H d\mathbf{W}(t).$$

In this subsection, the second-order covariance propagation formula is demonstrated by “pasting together” two ensembles of needle trajectories from $t = 0$ to $t = 1/2$ and $t = 1/2$ to $t = 1$ to get the mean and covariance of needle trajectories from $t = 0$ to $t = 1$. These needle trajectories are generated by integrating the SDE in (35) for these three time periods with $\Delta t = 0.01$ using a modified version of the Euler-Maruyama method for generating sample paths of SDEs [Higham 2001]. The mean and covariance resulted from the second-order propagation formula are then compared with those obtained by integrating the SDE from $t = 0$ to $t = 1$ as detailed below.

The reference frame $g(t)$ is generated from $\zeta(t) = (g^{-1}dg)^\vee$ by the product of exponentials formula at multiples of the small time step Δt as:

$$g(n\Delta t) = \exp \widehat{\xi}(\Delta t) \exp \widehat{\xi}(2\Delta t) \cdots \exp \widehat{\xi}(n\Delta t).$$

A cloud of frames $\{g^{(i)}(n\Delta t)\}$ for the trials $i = 1, \dots, 10,000$ are created with $\kappa = 0.05$ and a certain value of λ_1 and λ_2 . The actual $SE(3)$ means and covariances of the cloud of frames $\{g^{(i)}(n\Delta t)\}$ are then computed using brute force enumeration for the three time periods: $t_1 = [0, 1/2]$, $t_2 = [1/2, 1]$, and $t_3 = [0, 1]$ by applying (31) and (32) respectively. With the actual means and covariances of needle trajectories for $t_1 = [0, 1/2]$ and $t_2 = [1/2, 1]$, the estimated means and covariances of needle trajectories for $t_3 = [0, 1]$ are derived using the second-order propagation formula (22) and (23). These estimated means and covariances for period $t_3 = [0, 1]$ are compared with their corresponding actual values. The comparison results are quantitatively expressed through a definition on *deviation*. The measure of deviation on the covariance is defined as (33). Similarly, the measure of deviation on the mean is defined as

$$\text{deviation} = \frac{\|\mu_{prop} - \mu_{actual}\|}{\|\mu_{actual}\|}, \quad (36)$$

where μ_{prop} is the mean calculated using second-order propagation formula (22), μ_{actual} is the actual mean calculated using (31), and $\|\cdot\|$ denotes the Hilbert-Schmidt (Frobenius) norm.

A range of values of λ_1 and λ_2 are tested to verify the effectiveness of the second order propagation formula. These values are $\lambda_1^2 = \lambda_2^2 = 0.05, 0.1, 0.5, 1, 1.2, 1.5$. These comparison results are illustrated through the graphs of deviation versus Gaussian white noise constant λ^2 as shown in Figures 3 and 4. It can be observed that the deviation of the mean is less than 0.3% and the deviation of the covariance is less than 1% for $\lambda < 1$, where $\lambda = 1$ is a fairly large noise constant. These comparisons have shown that the mean and covariance computed from the second-order propagation formula are very good approximations to those obtained by integrating the SDE from $t = 0$ to $t = 1$.

5 Conclusions

In this paper, first-order kinematic error propagation formulas are modified to include second-order effects. This extends the usefulness of these formulas to errors that are not necessarily small. In fact, in the example to which the methodology is applied, errors in orientation can be as large as a radian or more and the second-order formula appears to capture the error well. The second-order propagation formula makes significant improvements in terms of accuracy than that of the first-order formula. The second-order propagation theory is much more robust than the first-order formula over a wide range of kinematic errors. This is demonstrated with the example of a PUMA manipulator arm with substantial errors in the joints, as well as stochastic trajectories of a nonholonomic kinematic model of a flexible needle.

Acknowledgments

This work was performed with partial support from the NIH Grants R01EB006435 "Steering Flexible Needles in Soft Tissue" and R01GM075310 "Group-Theoretic Methods in Protein Structure Determination." We thank Dr. Wooram Park for useful discussions on stochastic simulations of needles.

Appendix

A Matrix Lie Groups in General

A *matrix Lie group* is a Lie group where G is a set of square matrices and the group operation is matrix multiplication. In this work, only the groups $SO(3)$ and $SE(3)$ will be considered.

A.1 The Exponential and Logarithm Maps

Given a general matrix Lie group, elements sufficiently close to the identity are written as $g(t) = e^{tX}$ for some $X \in \mathcal{G}$ (the Lie algebra of G) and t near 0. Explicitly,

$$e^X = \sum_{k=0}^{\infty} \frac{X^k}{k!}. \quad (37)$$

The matrix logarithm is defined by the Taylor series about the identity matrix:

$$\log(g) = \log(I + (g - I)) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{(g - I)^k}{k}. \tag{38}$$

For matrix Lie groups, operations such as $g - I$ and division of g by a scalar are well defined. The exponential map takes an element of the Lie algebra and produces an element of the Lie group. This is written as:

$$\exp: \mathcal{G} \rightarrow G.$$

The logarithm map does just the opposite:

$$\log: G \rightarrow \mathcal{G}.$$

In other words, $\log(\exp X) = X$, and $\exp(\log(g)) = g$.

Given any smooth curve $g(t) \in G$, we can compute $g^{-1} \frac{dg}{dt}$ and $\frac{dg}{dt} g^{-1}$. These will be elements of \mathcal{G} .

A.2 The Lie Bracket and the Adjoint Matrices $Ad(g)$ and $ad(X)$

The *adjoint* operator is defined as

$$Ad(g_1)X = \left. \frac{d}{dt} (g_1 e^{tX} g_1^{-1}) \right|_{t=0} = g_1 X g_1^{-1}. \tag{39}$$

This gives a homomorphism $Ad: G \rightarrow GL(\mathcal{G})$ from the group into the set of all invertible linear transformations of \mathcal{G} onto itself. It is a homomorphism because

$$Ad(g_1)Ad(g_2)X = g_1 (g_2 X g_2^{-1}) g_1^{-1} = (g_1 g_2) X (g_1 g_2)^{-1} = Ad(g_1 g_2)X.$$

It is linear because

$$\begin{aligned} Ad(g)(c_1 X_1 + c_2 X_2) &= g(c_1 X_1 + c_2 X_2)g^{-1} = c_1 g X_1 g^{-1} + c_2 g X_2 g^{-1} \\ &= c_1 Ad(g)X_1 + c_2 Ad(g)X_2. \end{aligned}$$

In the special case of a 1-parameter subgroup when $g = g(t)$ is an element close to the identity.⁴ we can approximate $g(t) \approx I + tX$ for small t . Then we get $Ad(I + tX)Y = Y + t(XY - YX)$. The quantity

$$XY - YX = [X, Y] = \left. \frac{d}{dt} (Ad(g(t))Y) \right|_{t=0} \tag{40}$$

is called the *Lie bracket* of the elements $X, Y \in \mathcal{G}$.

⁴In the context of matrix Lie groups, one natural way to measure distance is as a matrix norm of the difference of two group elements.

It is clear from the definition in (40) that the Lie bracket is linear in each entry:

$$[c_1X_1+c_2X_2, Y]=c_1 [X_1, Y] +c_2 [X_2, Y]$$

and

$$[X, c_1Y_1+c_2Y_2]=c_1 [X, Y_1] +c_2 [X, Y_2].$$

Furthermore, the Lie bracket is anti-symmetric:

$$[X, Y] = - [Y, X], \tag{41}$$

and hence $[X, X] = 0$. Given a basis $\{E_1, \dots, E_n\}$ for the Lie algebra \mathcal{G} , any arbitrary element can be written as

$$X = \sum_{i=1}^n x_i E_i.$$

The Lie bracket of any two elements will result in a linear combination of all basis elements. This is written as

$$[E_i, E_j] = \sum_{k=1}^n C_{ij}^k E_k.$$

The constants C_{ij}^k are called the *structure constants* of the Lie algebra \mathcal{G} . Note that the structure constants are antisymmetric: $C_{ij}^k = -C_{ji}^k$.

It can be checked that for any three elements of the Lie algebra, the *Jacobi identity* is satisfied:

$$[X_1, [X_2, X_3]] + [X_2, [X_3, X_1]] + [X_3, [X_1, X_2]] = 0. \tag{42}$$

It is often convenient to write the independent entries of any $X \in \mathcal{G}$ as a column vector using the notation $\mathbf{x} = X^\vee$ where the rule $\mathbf{e}_i = E_i^\vee$ is used. The particular details of the \vee operator for the cases of $SO(3)$ and $SE(3)$ are given in the following appendices.

A matrix denoted as $ad(X)$ can then be defined such that for any $X, Y \in \mathcal{G}$

$$[X, Y] = ad(X) \mathbf{y}.$$

From (41), it follows that

$$ad(X) \mathbf{y} = - ad(Y) \mathbf{x}.$$

B The Rotation Group, $SO(3)$

The Lie algebra $so(3)$ consists of skew-symmetric matrices of the form

$$X = \begin{pmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{pmatrix} = \sum_{i=1}^3 x_i E_i. \tag{43}$$

The skew-symmetric matrices $\{E_i\}$ form a basis for the set of all such 3×3 skew-symmetric matrices, and the coefficients $\{x_i\}$ are all real. The \vee operation is defined to extract these coefficients from a skew symmetric matrix to form a column vector $[x_1, x_2, x_3]^T \in \mathbb{R}^3$ such that $X\mathbf{y} = \mathbf{x} \times \mathbf{y}$ for any $\mathbf{y} \in \mathbb{R}^3$, where \times is the usual vector cross product.

In this case, the adjoint matrices are

$$Ad(R) = R \quad \text{and} \quad ad(X) = X.$$

Furthermore,

$$[X, Y]^\vee = \mathbf{x} \times \mathbf{y}.$$

It is well known (see [Chirikjian 2001] for derivation and references) that

$$R(\mathbf{x}) = e^X = I + \frac{\sin\|\mathbf{x}\|}{\|\mathbf{x}\|} X + \frac{(1 - \cos\|\mathbf{x}\|)}{\|\mathbf{x}\|^2} X^2 \tag{44}$$

where $\|\mathbf{x}\| = (x_1^2 + x_2^2 + x_3^2)^{\frac{1}{2}}$. Clearly, since the instantaneous rotation axis is preserved under a rotation, $R(\mathbf{x})\mathbf{x} = \mathbf{x}$.

An interesting and useful fact is that except for a set of measure zero, all elements of $SO(3)$ can be captured with the parameters within the open ball defined by $\|\mathbf{x}\| < \pi$, and the matrix logarithm of any group element parameterized in this range is also well defined. It is convenient to know that the angle of the rotation, $\theta(R)$, is related to the exponential parameters as $|\theta(R)| = \|\mathbf{x}\|$. Furthermore,

$$\log(R) = \frac{1}{2} \frac{\theta(R)}{\sin\theta(R)} (R - R^T)$$

where

$$\theta(R) = \cos^{-1} \left(\frac{\text{trace}(R) - 1}{2} \right).$$

Invariant definitions of directional (Lie) derivatives and integration measure for $SO(3)$ can be defined. When computing these invariant quantities in coordinates (including exponential coordinates), a Jacobian matrix comes into play. There are two such Jacobian matrices:

$$J_I(\mathbf{x}) = \left[\left(\frac{\partial R}{\partial x_1} R^T \right)^\vee, \left(\frac{\partial R}{\partial x_2} R^T \right)^\vee, \left(\frac{\partial R}{\partial x_3} R^T \right)^\vee \right].$$

and

$$J_r(\mathbf{x}) = \left[\left(R^T \frac{\partial R}{\partial x_1} \right)^\vee, \left(R^T \frac{\partial R}{\partial x_2} \right)^\vee, \left(R^T \frac{\partial R}{\partial x_3} \right)^\vee \right].$$

The subscripts r and l denote the side where the partial derivative appears (right or left).

These two Jacobian matrices are related as

$$J_l = R J_r. \quad (45)$$

Relatively simple analytical expressions have been derived by Park [?] for the Jacobian J_l and its inverse when rotations are parameterized as in (44). These expressions are

$$J_l(\mathbf{x}) = I + \frac{1 - \cos \|\mathbf{x}\|}{\|\mathbf{x}\|^2} X + \frac{\|\mathbf{x}\| - \sin \|\mathbf{x}\|}{\|\mathbf{x}\|^3} X^2 \quad (46)$$

and

$$J_l^{-1}(\mathbf{x}) = I - \frac{1}{2} X + \left(\frac{1}{\|\mathbf{x}\|^2} - \frac{1 + \cos \|\mathbf{x}\|}{2\|\mathbf{x}\|\sin \|\mathbf{x}\|} \right) X^2.$$

The corresponding Jacobian J_r and its inverse are then calculated using (45) as [Chirikjian 2001]

$$J_r(\mathbf{x}) = I - \frac{1 - \cos \|\mathbf{x}\|}{\|\mathbf{x}\|^2} X + \frac{\|\mathbf{x}\| - \sin \|\mathbf{x}\|}{\|\mathbf{x}\|^3} X^2$$

and

$$J_r^{-1}(\mathbf{x}) = I - \frac{1}{2} X + \left(\frac{1}{\|\mathbf{x}\|^2} - \frac{1 + \cos \|\mathbf{x}\|}{2\|\mathbf{x}\|\sin \|\mathbf{x}\|} \right) X^2.$$

Note that

$$J_l = J_r^T.$$

The determinants are

$$|\det(J_l)| = |\det(J_r)| = \frac{2(1 - \cos \|\mathbf{x}\|)}{\|\mathbf{x}\|^2}.$$

Given a square-integrable function of rotation, $f(R) \in L^2(SO(3))$, the proper (invariant) way to integrate using exponential coordinates is

$$\int_{SO(3)} f(R) dR = \frac{1}{8\pi^2} \int_{\|\mathbf{x}\| < \pi} f(R\mathbf{x}) |\det J(\mathbf{x})| d\mathbf{x}$$

where $d\mathbf{x} = dx_1 dx_2 dx_3$ and J can denote either J_r or J_l . The normalization by $8\pi^2$ ensures that $\int_{SO(3)} 1 dR = 1$.

C The Special Euclidean Group, SE(3)

The Lie algebra $se(3)$ consists of “screw” matrices of the form

$$X = \begin{pmatrix} 0 & -x_3 & x_2 & x_4 \\ x_3 & 0 & -x_1 & x_5 \\ -x_2 & x_1 & 0 & x_6 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \sum_{i=1}^6 x_i \tilde{E}_i. \tag{47}$$

The matrices $\{\tilde{E}_i\}$ form a basis for the set of all such 4×4 screw matrices, and the coefficients $\{x_i\}$ are all real. The tilde is used to distinguish between these basis elements and those for $SO(3)$. The \vee operation is defined to extract these coefficients from a screw matrix to form a column vector $X^\vee = [x_1, x_2, x_3, x_4, x_5, x_6]^T \in \mathbb{R}^6$. The double use of \vee in the $so(3)$ and $se(3)$ cases will not cause confusion, since the object to which it is applied defines the sense in which it is used.

It will be convenient to define $\boldsymbol{\omega} = [x_1, x_2, x_3]^T$, and $\mathbf{v} = [x_4, x_5, x_6]^T$, so that

$$X^\vee = \mathbf{x} = \begin{pmatrix} \boldsymbol{\omega} \\ \mathbf{v} \end{pmatrix}.$$

It can be shown that [Chirikjian 2009]

$$g(X^\vee) = \exp X = \begin{pmatrix} R(\boldsymbol{\omega}) & J_l(\boldsymbol{\omega}) \mathbf{v} \\ 0^T & 1 \end{pmatrix}. \tag{48}$$

This follows from the expression for the matrix exponential given in [Murray, Li and Sastry 1994] and the definition of the $SO(3)$ Jacobian in (46). From the form of (48), it is clear that if g has rotational part R , and translational part \mathbf{t} , then the matrix logarithm can be written in closed form as

$$X = \log(g) = \begin{pmatrix} \log R & J_l^{-1}((\log R)^\vee) \mathbf{t} \\ 0^T & 0 \end{pmatrix},$$

and

$$X^\vee = \begin{pmatrix} (\log R)^\vee \\ J_l^{-1}((\log R)^\vee) \mathbf{t} \end{pmatrix}. \tag{49}$$

The adjoint matrices for $SE(3)$ are

$$Ad(g) = \begin{pmatrix} R & 0_3 \\ TR & R \end{pmatrix} \in IR^{6 \times 6}$$

where $T^V = \mathbf{t}$ and

$$ad(X) = \begin{pmatrix} \Omega & 0_3 \\ V & \Omega \end{pmatrix} \in IR^{6 \times 6}$$

where $V^V = \mathbf{v}$ and $\Omega^V = \boldsymbol{\omega}$.

The Jacobians for $SE(3)$ using exponential parameters are then

$$\mathcal{J}_l(\mathbf{x}) = \left[\left(\frac{\partial g}{\partial x_1} g^{-1} \right)^V, \left(\frac{\partial g}{\partial x_2} g^{-1} \right)^V, \dots, \left(\frac{\partial g}{\partial x_6} g^{-1} \right)^V \right].$$

and

$$\mathcal{J}_r(x) = \left[\left(g^{-1} \frac{\partial g}{\partial x_1} \right)^V, \left(g^{-1} \frac{\partial g}{\partial x_2} \right)^V, \dots, \left(g^{-1} \frac{\partial g}{\partial x_6} \right)^V \right].$$

The right Jacobian for $SE(3)$ in exponential coordinates can be computed from (48) as

$$\mathcal{J}_r(\mathbf{x}) = \begin{pmatrix} J_r \boldsymbol{\omega} & 0_3 \\ e^{-\Omega} \frac{\partial}{\partial \boldsymbol{\omega}} (J_l(\boldsymbol{\omega}) \mathbf{v}) & J_r \boldsymbol{\omega} \end{pmatrix}. \quad (50)$$

where 0_3 is the 3×3 zero matrix. It becomes immediately clear that

$$|\det(\mathcal{J}_r(\mathbf{x}))| = |\det(J_r(\boldsymbol{\omega}))|^2.$$

Given a square-integrable function of motion, $f(g) \in L^2(SE(3))$, the proper (invariant) way to integrate using exponential coordinates is

$$\int_{SE(3)} f(g) dg = \frac{1}{8\pi^2} \int_{\mathbf{v} \in IR^3} \int_{\|\boldsymbol{\omega}\| < \pi} f(e^X) |\det(J_r(\boldsymbol{\omega}))|^2 d\boldsymbol{\omega} d\mathbf{v}.$$

The normalization by $8\pi^2$ is an artifact of the $SO(3)$ case, which is retained since $SE(3)$ is not compact.

References

- Alterovitz, R.; Simeon, T.; Goldberg, K. Robotics Science and Systems. Atlanta, GA: Jun. 2007 The Stochastic Motion Roadmap: A Sampling Framework for Planning with Markov Motion Uncertainty.
- Anderson, TW. An Introduction to Multivariate Statistical Analysis. 3rd ed.. Wiley; New York: 2005.
- Baker HF. Alternants and Continuous Groups. Proc. London Mathematical Society 1904;3:24–47. Second Series

- Campbell JE. On a Law of Combination of Operators. Proc. London Mathematical Society 1897;29:14–32.
- Chirikjian, GS. Stochastic Models, Information, and Lie Groups. Birkhauser; New York: 2009.
- Chirikjian, GS.; Kyatkin, AB. Engineering Applications of Noncommutative Harmonic Analysis. CRC Press; Boca Raton, FL: 2001.
- Craig, JJ. Introduction to Robotics Mechanics and Control. 3rd edition. Prentice Hall: 2005.
- Higham DJ. An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations. SIAM Review 2001;43(3):525–546.
- Hausdorff, F. Berichte der Sachsichen Akademie der Wissenschaften. Vol. 58. Leipzig; 1906. Die symbolische Exponentialformel in der Gruppentheorie; p. 19-48.
- Kim J-S, Chirikjian GS. A unified approach to conformational statistics of classical polymer and polypeptide models. Polymer Nov 28;2005 46(25):11904–11917. [PubMed: 20165562]
- Murray, RM.; Li, Z.; Sastry, SS. A Mathematical Introduction to Robotic Manipulation. CRC Press; Boca Raton: 1994.
- Park, W.; Kim, JS.; Zhou, Y.; Cowan, NJ.; Okamura, AM.; Chirikjian, GS. Diffusion-based motion planning for a non-holonomic flexible needle model; IEEE International Conference on Robotics and Automation; 2005. p. 4600-4605.
- Park W, Liu Y, Moses M, Chirikjian GS. Kinematic State Estimation and Motion Planning for Stochastic Nonholonomic Systems Using the Exponential Map. Robotica. 2008 to appear.
- Selig, JM. Geometrical Methods in Robotics. Springer; New York: 1996.
- Smith, P.; Drummond, T.; Roussopoulos, K. Computing MAP trajectories by representing, propagating and combining PDFs over groups; Proceedings of the 9th IEEE International Conference on Computer Vision; Nice. 2003. p. 1275-1282.
- Smith RC, Cheeseman P. On the Representation and Estimation of Spatial Uncertainty. The International Journal of Robotics Research 1986;5(4):56–68.
- Su S, Lee CSG. Manipulation and Propagation of Uncertainty and Verification of Applicability of Actions in assembly Tasks. IEEE Transactions on Systems, Man, and Cybernetics 1992;22(6):1376–1389.
- Thrun, S.; Burgard, W.; Fox, D. Probabilistic Robotics. MIT Press; Cambridge, MA: 2005.
- Wang Y, Chirikjian GS. Error Propagation on the Euclidean Group with Applications to Manipulator Kinematics. IEEE Transactions on Robotics 2006;22(4):591–602.
- Wang, Y.; Chirikjian, GS. Error Propagation in Hybrid Serial-Parallel Manipulators; IEEE International Conference on Robotics and Automation; Orlando Florida. May. 2006 p. 1848-1853.
- Webster RJ III, Kim JS, Cowan NJ, Chirikjian GS, Okamura AM. Nonholonomic modeling of needle steering. International Journal of Robotics Research 2006;25:509–525.
- Zhou Y, Chirikjian GS. Conformational Statistics of Semi-Flexible Macromolecular Chains with Internal Joints. Macromolecules 2006;39(5):1950–1960.

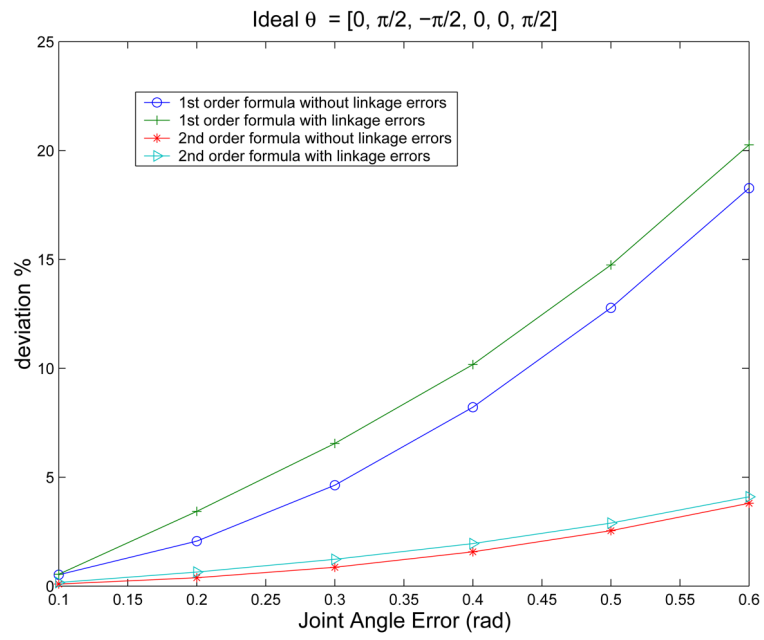


Figure 1.
The Deviation of the First and Second-order Propagation Methods for Configuration I

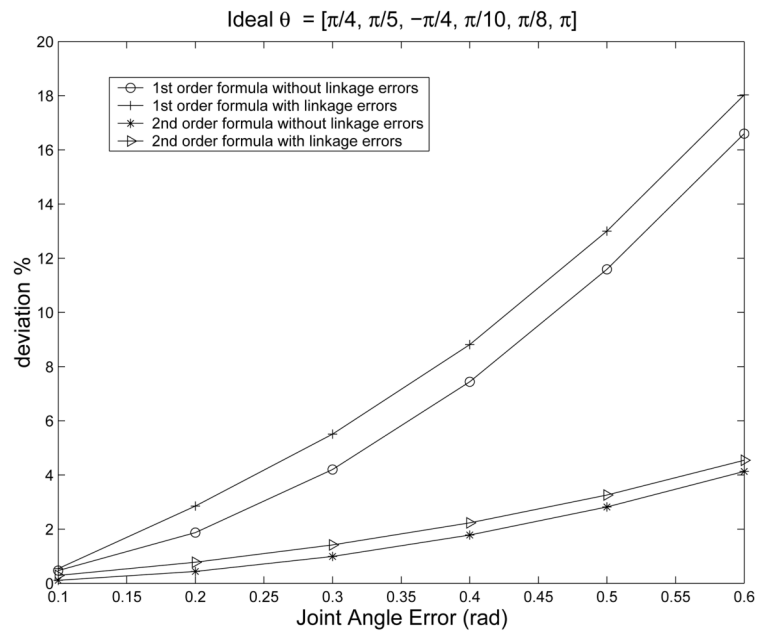


Figure 2.
The Deviation of the First and Second-order Propagation Methods for Configuration II

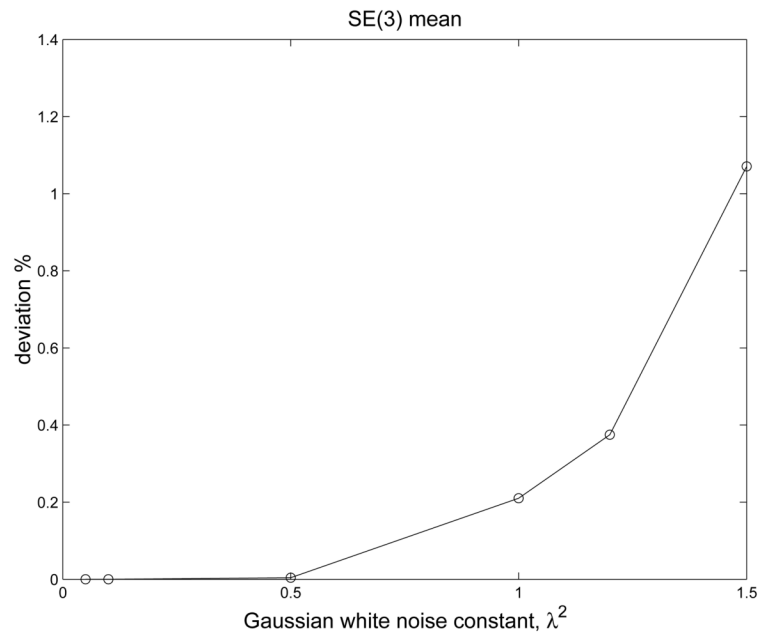


Figure 3.
The Deviation of the Mean

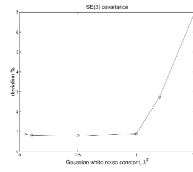


Figure 4.
The Deviation of the Covariance

Table 1

DH Parameters of the PUMA 560

i	α_{i-1}	a_{i-1}	d_i	θ_i
1	0	0	0	θ_1
2	-90°	0	0	θ_2
3	0	a_2	d_3	θ_3
4	-90°	a_3	d_4	θ_4
5	90°	0	0	θ_5
6	-90°	0	0	θ_6