Error Propagation on the Euclidean Group With Applications to Manipulator Kinematics

Yunfeng Wang, *Member, IEEE*, and Gregory S. Chirikjian, *Member, IEEE*

***Abstract*—Error propagation on the Euclidean motion group arises in a number of areas such as errors that accumulate from the base to the distal end of manipulators. We address error propagation in rigid-body poses in a coordinate-free way, and explain how this differs from other approaches proposed in the literature. In this paper, we show that errors propagate by convolution on the Euclidean motion group, $SE(3)$. When local errors are small, they can be described well as distributions on the Lie algebra $se(3)$. We show how the concept of a highly concentrated Gaussian distribution on $SE(3)$ is equivalent to one on $se(3)$. We also develop closure relations for these distributions under convolution on $SE(3)$. Numerical examples illustrate how convolution is a valuable tool for computing the propagation of both small and large errors.**

***Index Terms*—Euclidean group, error propagation, manipulator kinematics, spatial uncertainty.**

I. Introduction

In THIS paper, we address how errors propagate on the Euclidean motion group. Applications include the accumulation of errors in serial linkages and the estimation of the state of a rigid body from noisy measurements. Our approach is to treat errors using probability densities on the Euclidean group. Whereas concepts such as integration and convolution of these densities follow in a natural way when considering the Lie group setting [4], standard concepts associated with the Gaussian distribution in $\mathbb{R}^N$ do not follow in a natural way to Lie groups. For example, a Gaussian distribution in $\mathbb{R}^N$ is the solution to a diffusion equation, it is the maximum entropy distribution; the family of Gaussians is closed under convolution and conditioning. In the Lie group setting, one can often satisfy one or several of these properties with specialized distributions, but not all, that is, at least not when discussing distributions with mass that is spread over a large region in the group. In contrast, concentrated distributions on Lie groups (which are often the most appropriate distributions to describe the sorts of small errors encountered in practice) can be constructed to have all of the properties associated with Gaussians.

In the following sections, the relevant literature is reviewed, and an overview of rigid-body motions is provided. In Section II, the motivating application of error accumulation in serial (and hybrid serial-parallel) manipulators is discussed. In Section III, the concept of highly concentrated Gaussian distributions is discussed, and several of their important properties are examined. In Section IV, closed-form expressions for the convolution of these densities are derived. Section V illustrates with numerical examples that both small and large serial errors propagate by convolution, and examines the range of values over which a distribution can be considered highly concentrated. Section VI presents our conclusions and discusses other potential applications of this formulation. The Appendix provides some background mathematics.

A. Literature Review

Several distinct research fields relate to the study presented in this paper. These include the theory of Lie groups, probability and statistics, robot kinematics, methods for describing spatial uncertainty, and state estimation. We review several of the most closely related works in each of these areas here.

Murray *et al.* [19] and Selig [23] presented Lie-group-theoretic notation and terminology to the robotics community, which has now become standard vocabulary. Park and Brockett [21] showed how dexterity measures can be viewed in a Lie-group setting, and how this coordinate-free approach can be used in robot design. Wang and Chirikjian [31] showed that the workspace densities of manipulators with many degrees of freedom can be generated by solving a diffusion equation on the Euclidean group. Blackmore and Leu [1] showed that problems in manufacturing associated with swept volumes can be cast within a Lie-group setting. Kyatkin and Chirikjian [4], [13] showed that many problems in robot kinematics and motion planning can be formulated as the convolution of functions on the Euclidean group.

Starting with the pioneering work of Brockett [2], the controls community has embraced group-theoretic problems for many years. This includes proportional-derivative (PD) control on the Euclidean group [3], [14], [33], tracking problems [8], [9], and estimation [15]. The representation and estimation of spatial uncertainty has also received attention in the robotics and vision literature [25]. Kinematic error propagation for use in assembly planning has also been studied [26]. Recent work on error propagation described by the concatenation of random variables on groups has also found promising applications in mobile robot navigation [24].

B. Review of Rigid-Body Motions

The Euclidean motion group $SE(3)$ is the semidirect product of $\mathbb{R}^3$ with the special orthogonal group $SO(3)$. We denote elements of $SE(3)$ as $g = (\mathbf{a}, A) \in SE(3)$, where $A \in SO(3)$ and $\mathbf{a} \in \mathbb{R}^3$. For any $g = (\mathbf{a}, A)$ and $h = (\mathbf{r}, R) \in SE(3)$, the

Manuscript received August 23, 2005; revised January 5, 2006. This paper was recommended for publication by Associate Editor J. Kosecka and Editor F. Park upon evaluation of the reviewers' comments. This work was supported by the National Science Foundation under Grant NSF-RHA 0098382 "Diffusion Processes in Motion Planning and Control."

Y. Wang is with the Department of Mechanical Engineering, The College of New Jersey, Ewing, NJ 08628 USA (e-mail: jwang@tcnj.edu).

G. S. Chirikjian is with the Department of Mechanical Engineering, Johns Hopkins University, Baltimore, MD 21218 USA (e-mail: gregc@jhu.edu).

Digital Object Identifier 10.1109/TRO.2006.878978

group law is written as $g \circ h = (\mathbf{a} + A\mathbf{r}, AR)$, and $g^{-1} = (-A^T\mathbf{a}, A^T)$. Alternately, one may represent any element of $SE(3)$ as a 4×4 homogeneous transformation matrix of the form

$$H(g) = \begin{pmatrix} A & \mathbf{a} \\ \mathbf{0}^T & 1 \end{pmatrix}$$

in which case, the group law is matrix multiplication.

For small translational (rotational) displacements from the identity along (about) the ith coordinate axis, the homogeneous transforms representing infinitesimal motions look like

$$H_i(\epsilon) \triangleq \exp(\epsilon \tilde{E}_i) \approx \mathbb{1}_{4\times 4} + \epsilon \tilde{E}_i$$

where $I_{4\times 4}$ is the 4×4 identity matrix, and

$$\tilde{E}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \tilde{E}_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\tilde{E}_3 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \tilde{E}_4 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\tilde{E}_5 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \tilde{E}_6 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

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

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

$$\exp(t\tilde{E}_6) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & t \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

In what follows, it will be convenient to describe elements of $SE(3)$ with the exponential parametrization

$$g = g(x_1, x_2, \ldots, x_6) = \exp\left(\sum_{i=1}^{6} x_i \tilde{E}_i\right). \quad (1)$$

This is common in the study of Lie groups and algebras [29].

One defines the "vee" operator $\vee$ such that

$$\left(\sum_{i=1}^{6} x_i \tilde{E}_i\right)^{\vee} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix}.$$

The total vector $\mathbf{x} \in \mathbb{R}^6$ can be obtained from $g \in SE(3)$ from the formula

$$\mathbf{x} = (\log g)^{\vee}. \quad (2)$$

II. Propagation of Error in Serial Linkages

Intuitively, if two rigid parts are manufactured with errors and those parts are bolted together at an interface, the errors will "add" in some way. Likewise, a manipulator that is constructed from several subunits, each with some manufacturing error and/or backlash, will have errors that accumulate as the length from base to end-effector is traversed. In this section, we quantify how errors accumulate in serial and hybrid serial-parallel devices. We formulate this as a convolution of highly concentrated error densities on $SE(3)$.

Suppose we are given a manipulator consisting of two concatenated units. These units could be Stewart–Gough platforms or serial links connected with revolute joints. One unit is stacked on top of the other one. The proximal unit will be able to reach each frame $h_1 \in SE(3)$ with some error when its proximal end is located at the identity $e \in SE(3)$. This error may be different for each different frame h_1. This is expressed mathematically as a real-valued function of $g_1 \in SE(3)$ which has a peak in the neighborhood of h_1, and decays rapidly away from h_1. If the unit could reach h_1 exactly, this function would be a delta function. Explicitly, the error may be described by one of many possible density functions depending on what error model is used. However, it will always be the case that it is of the form $\rho_1(h_1, g_1)$ for $h_1, g_1 \in SE(3)$, that is, the error will be a function of $g_1 \in SE(3)$ for each frame h_1 that the top of the module tries to attain relative to its base. Likewise, the second module will have an error function $\rho_2(h_2, g_2)$ for $h_2, g_2 \in SE(3)$ that describes the distribution of frames around h_2 that might be reached when h_2 is the expected end frame for module 2 relative to its base, and the base of module 2 is located at the identity $e \in SE(3)$.

The error distribution that results from the concatenation of two modules with errors $\rho_1(\cdot)$ and $\rho_2(\cdot)$ results from sweeping the error distribution of the second module by that of the first. This is written mathematically as

$$\begin{aligned} \rho(h_1 \circ h_2, g) &= (\rho_1 \otimes \rho_2)(h_1 \circ h_2, g) \\ &\triangleq \int_{SE(3)} \rho_1(h_1, g_1)\rho_2\left(h_2, g_1^{-1} \circ g\right) d(g_1). \end{aligned} \quad (3)$$

Here, $d(g)$ is the unique bi-invariant integration measure for $G = SE(3)$ evaluated at g [4]. Sometimes this is simply written as dg. In the case of no error, the multiplication of homogeneous transforms h_1 and h_2 as $h_1 \circ h_2$ represents the composite change in position and orientation from the base of the lower unit to the interface between units, and from the interface to the top of the upper unit. In the case of inexact kinematics, the error function for the upper unit is shifted by the lower unit $(\rho_2(h_2, g_1^{-1} \circ g))$, weighted by the error distribution of the lower unit $(\rho_1(h_1, g_1))$ and integrated over the support of the error distribution of the lower unit (which is the same as integrating over all of $SE(3)$, since outside of the support of the error distribution, the integral is zero). The result of this integration is, by definition, the error density function around the frame $h_1 \circ h_2$, and this is denoted

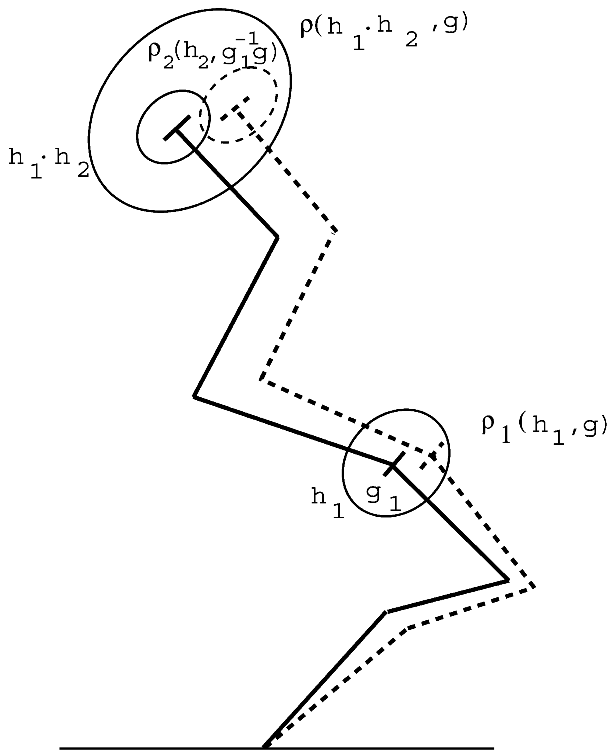

Fig. 1. Error propagation in serial linkages.

as $(\rho_1 \otimes \rho_2)(h_1 \circ h_2, g)$. We illustrated (3) in Fig. 1. Parametric distributions that can be used for this application are discussed in Section III, but it should be noted that (3) holds, regardless of the size of the errors or the form of the error density.

To test this formulation, consider the case of exact kinematics. In this case, the error distributions have a very special form: they are Dirac delta functions on $SE(3)$. In complete analogy with the usual Dirac delta function on the real line, we have the properties

$$\int_{SE(3)} \delta(g) d(g) = 1$$
$$\delta(h^{-1} \circ g) = \delta(g^{-1} \circ h)$$
$$\int_{SE(3)} f(h)\delta(h^{-1} \circ g) d(h) = f(g).$$

Using these properties, the error distributions for both units may be written as

$$\rho_i(h_i, g) = \delta\left(h_i^{-1} \circ g\right).$$

Then, in this special case, (3) reduces to

$$\begin{aligned}(\delta \otimes \delta)(h_1 \circ h_2, g) &= \int_{SE(3)} \delta\left(h_1^{-1} \circ g_1\right) \delta\left(h_2^{-1} \circ g_1^{-1} \circ g\right) d(g_1) \\ &= \delta\left(h_2^{-1} \circ h_1^{-1} \circ g\right) \\ &= \delta\left((h_1 \circ h_2)^{-1} \circ g\right). \end{aligned} \quad (4)$$

In other words, in the case of exact kinematics, we have exactly the result which is expected.

III. Special Features of Concentrated Probability Functions

Errors in manufactured parts, and in the assembly of those parts into larger structures, are typically small, but not so small as to be ignored. Therefore, having a way to describe small errors using concentrated probability density functions (pdfs) is useful. This section focuses on the properties of concentrated pdfs on the Euclidean group. In the proof that follows, a number detailed mathematical steps are skipped. These details can be obtained from the authors on request.

A. Probability Densities Concentrated at the Identity

Suppose that instead of a deterministic and exactly measured frame of reference $h \in SE(3)$, we have a distribution (or cloud) of frames of reference that are tightly clustered around h. How do we describe such things in a quantitative way? Let us first consider a cloud clustered around the identity $e \in SE(3)$. In order to quantify what is meant by a highly concentrated/tightly clustered density, a few definitions are required.

Definition 1: Compatibility: Let G be a real unimodular matrix Lie group[1] (of which $SE(n)$ and $SO(n)$ are examples), and let $\{E_i\}$ be an orthogonal basis for the associated Lie algebra $\mathcal{G}$, that is, $(E_i, E_j) = \lambda_i \delta_{ij}$, where $(\cdot,\cdot)$ is an inner product on $\mathcal{G}$, δ_{ij} is the Kronecker delta, and λ_i is a scale factor (which can be set to $\lambda_i = 1$ if each E_i is scaled appropriately). Let $X = \sum x_i E_i$ where $x_i \in \mathbb{R}$. Let ϵ be a small positive real number. Let $d : G \times G \to \mathbb{R}$ be a metric (distance function) for G, with the additional property that

$$\lim_{(X,X)^{\frac{1}{2}} \to \epsilon} \left| d\left(e, \exp(X)\right) - (X,X)^{\frac{1}{2}} \right| = O(\epsilon^2)$$

where e is the identity element of G. Let us say that when this condition holds, the metric $d(\cdot,\cdot)$ and inner product $(\cdot,\cdot)$ are *compatible with each other*. (Examples for $SE(3)$ can be found in the Appendix.)

Definition 2: Rapidly Decreasing Unimodal Distribution: A pdf on a real unimodular matrix Lie group is a function $\rho : G \to \mathbb{R}$, with the property that $\int_G \rho(g) d(g) = 1$, where $d(g)$ is the appropriately normalized bi-invariant integration measure for G. It is called unimodal and rapidly decreasing with mode at the identity $e \in G$ if, for all $X \in \mathcal{G}$, the following inequality holds:

$$\rho\left(\exp(X)\right) > \rho\left(\exp(\alpha X)\right) m(X, \alpha)$$

for all values of α for which $d(e, \exp(X)) < d(e, \exp(\alpha X))$, and some real-valued function $m(X, \alpha) \geq 1$ that increases exponentially for sufficiently large values of α. In other words, $\rho(\cdot)$ decreases monotonically and rapidly as it traverses any one-dimensional (1-D) subgroup away from the identity.

Definition 3: Tightness of a Distribution: Let G be a real unimodular matrix Lie group, and let $\rho : G \to \mathbb{R}$ be a smooth and rapidly decreasing unimodal pdf. $\rho(\cdot)$ is called *tightly focused*

[1]Recall that a Lie group is called unimodular if its integration measure $d(g)$ has the property that $d(g_0 \circ g) = d(g \circ g_0)$ for all $g_0 \in G$

or *highly concentrated* at the identity if, for compatible $d(\cdot,\cdot)$ and $(\cdot,\cdot)$, the following is true for a small number $\epsilon \in \mathbb{R}^+$:

$$1 - \int_{G|(X,X)^{\frac{1}{2}}<\epsilon} \rho\left(\exp(X)\right) d\left(\exp(X)\right) = O(\epsilon)$$

and, in addition, $m(X,\alpha)$ grows exponentially for values of $\alpha > \epsilon$. Stated intuitively, a distribution satisfies this condition if most of its mass is supported on a small region around the identity. Such a distribution can be moved elsewhere by the left-shift operation.

If two functions, each concentrated within a ball of radius ϵ about the identity are convolved, the result will be concentrated in a ball of radius 2ϵ. Therefore, if one performs n convolutions, in order for the result to be safely considered as tightly focused, each of the original functions should be concentrated within a ball of radius ϵ/n.

Tightly focused distributions are essentially pdfs on the Lie algebra $se(3)$, and we therefore can use any number of parametric distributions that are used in $\mathbb{R}^N$. For example, the Gaussian distribution

$$\rho(g) = c \exp\left(-\frac{1}{2}\mathbf{x}^T C \mathbf{x}\right) \tag{5}$$

can be used, where $g = g(x_1, x_2, \ldots, x_6)$ as in (1), and $\mathbf{x}$ is defined as in (2).

We note that while the exponential mapping from $se(3)$ to $SE(3)$ is not bijective, this is irrelevant for two reasons: 1) the set of measure zero for which bijectivitity fails has no effect on nonpathological pdfs; and 2) the small errors to which this mapping is applied are not located at the singularities of the mapping, which are far from the identity.

The normalization constant c is determined by setting

$$\int_G \rho(g) dg = 1$$

so as to make $\rho(g)$ a pdf. Here, dg is the unique bi-invariant integration measure for $SE(3)$. In exponential parameters

$$dg = w(x_1, x_2, \ldots, x_6) dx_1 dx_2 \cdots dx_6$$

near the identity $w \approx 1$. Therefore, when $\rho(g)$ is tightly concentrated around the identity, we have

$$\int_G \rho(g) dg \approx \int_{\mathbf{R}^6} \rho\left(g(x_1, x_2, \ldots, x_6)\right) dx_1 dx_2 \cdots dx_6.$$

This is true for exponential coordinates and a distribution highly concentrated at the identity. Therefore, the constant c in (5) can be set in the usual way that it is for Gaussian distributions. In particular, if Σ is the matrix of covariances with elements defined by

$$\sigma_{ij} = \int_{\mathbf{R}^6} x_i x_j \rho\left(g(x_1, x_2, \ldots, x_6)\right) dx_1 dx_2 \cdots dx_6 \tag{6}$$

then

$$C = \Sigma^{-1}$$
$$c = \left(8\pi^3 |\det \Sigma|^{\frac{1}{2}}\right)^{-1}.$$

Special properties of the distribution in (5) are proved in the Appendix, as is the issue of whether or not covariances should be defined as in (6).

Given two probability densities on $SE(3)$, their convolution is defined as

$$(\rho_1 * \rho_2)(g) = \int_G \rho_1(h) \rho_2(h^{-1} \circ g) dh. \tag{7}$$

This can be considered as a special case of (3), when the dependence on h_1 and h_2 either does not exist or is suppressed for notational convenience. If ρ_1 describes a distribution of frames of reference $\{h_1, \ldots, h_n\}$, and ρ_2 describes a distribution of frames of reference $\{g_1, \ldots, g_m\}$, then the convolution $(\rho_1 * \rho_2)$ is the distribution that describes the distribution of all pairs $\{h_i \circ g_j | (i,j) \in [1, 2, \ldots, n] \times [1, 2, \ldots, m]\}$. In general, since $h_i \circ g_j \neq g_j \circ h_i$, it follows that $(\rho_1 * \rho_2)(g) \neq (\rho_2 * \rho_1)(g)$. However, convolutions of two distributions centered tightly around the identity do commute.

In what follows, the functions $\rho_i(g)$ are interpreted as functions with the argument in $SE(3)$ described as 4×4 homogeneous transformations. These functions can be extended to have argument in $\mathbb{R}^{4\times 4}$ in a number of ways, e.g., by setting $\rho(k) = 0$ for all $k \in \mathbb{R}^{4\times 4} - SE(3)$, or by having $\rho(k)$ decay rapidly to zero as the distance between k and G increases. When such extensions are smooth, then expanding ρ in a Taylor series in $\mathbb{R}^{4\times 4}$ yields

$$\rho\left(I + X + O(X^2)\right) = \rho(I + X)\left(1 + O\left(\|x\|\right)\right) \tag{8}$$

where $\|x\| = (X,X)^{1/2}$ and $O(X^2)$ is defined in a natural way, i.e., $O(X^2)$ is a matrix with entries each of $O((X,X))$ with $(X,X) = (1/2) tr(XX^T)$.

Equation (8) is useful in evaluating expressions in the proof below. Note that equalities that are presented below are true to $O(\epsilon)$ in the sense that $a(g) = b(g)$ denotes $\int_G |a(g) - b(g)| d(g) = O(\epsilon)$. With this, we have the following.

Theorem: Convolution of two functions on $SE(3)$, each tightly focused at the identity, is the same as convolution on $\mathbb{R}^6$ using exponential parameters as coordinates.

Proof: Let $g = \exp(\sum_{i=1}^{6} x_i \tilde{E}_i)$ and $h = \exp(\sum_{i=1}^{6} \xi_i \tilde{E}_i)$. Let $\rho_i(g)$, for $i = 1, 2$, be functions tightly focused at the identity. Then[2]

$$\rho_1(h) = \rho_1 \left(I + \sum_{i=1}^{6} \xi_i \tilde{E}_i \right) \triangleq \tilde{\rho}_1(\boldsymbol{\xi}).$$

Let us define $\tilde{\rho}_2(\cdot)$ in an analogous way. Then, using (8) and retaining zeroth-order terms, we have

$$\begin{aligned}
\rho_2(h^{-1} \circ g) &= \rho_2 \left(\exp\left(-\sum_{i=1}^{6} \xi_i \tilde{E}_i \right) \exp\left(\sum_{i=1}^{6} x_i \tilde{E}_i \right) \right) \\
&= \rho_2 \left(\left(I - \sum_{i=1}^{6} \xi_i \tilde{E}_i \right) \left(I + \sum_{i=1}^{6} x_i \tilde{E}_i \right) \right) \\
&= \rho_2 \left(I + \sum_{i=1}^{6} (x_i - \xi_i) \tilde{E}_i \right) \\
&= \tilde{\rho}_2(\mathbf{x} - \boldsymbol{\xi}).
\end{aligned}$$

Then, the convolution (7) can be written in this special case as

$$(\rho_1 * \rho_2)(g) = \int_{\mathbb{R}^6} \tilde{\rho}_1(\boldsymbol{\xi}) \tilde{\rho}_2(\mathbf{x} - \boldsymbol{\xi}) w(\boldsymbol{\xi}) d\boldsymbol{\xi}$$

and, since $\boldsymbol{\xi} \approx \mathbf{0}$, $w(\boldsymbol{\xi}) \approx 1$. Therefore, we can write

$$(\widetilde{\rho_1 * \rho_2})(\mathbf{x}) = (\tilde{\rho}_1 \star \tilde{\rho}_2)(\mathbf{x}) \tag{9}$$

where $\star$ is the convolution of functions in $\mathbb{R}^6$.

B. Convolution of Probability Densities Shifted From the Identity

The issue of how to describe highly concentrated distributions around a frame of reference h is handled easily by left translating a distribution defined around the identity

$$\rho_h(g) = L(h) \{\rho(g)\} = \rho(h^{-1} g).$$

Given two shifted functions $f_a^1(g) = f^1(a^{-1} \circ g)$ and $f_b^2(g) = f^2(b^{-1} \circ g)$, the convolution is

$$\begin{aligned}
\left(f_a^1 * f_b^2 \right)(g) &= \int_G f_a^1(h) f_b^2(h^{-1} \circ g) dh \\
&= \int_G f^1(a^{-1} \circ h) f^2(b^{-1} \circ h^{-1} \circ g) dh.
\end{aligned}$$

If we define the new variable $k = a^{-1} \circ h$, then $h^{-1} = k^{-1} \circ a^{-1}$. Therefore

$$\left(f_a^1 * f_b^2 \right)(g) = \int_G f^1(k) f^2(b^{-1} \circ k^{-1} \circ a^{-1} \circ g) dk.$$

[2]Equation (8) and the tightness of the distribution are used here and in the manipulations that follow.

If we define q such that $g = a \circ b \circ q$, then

$$\left(f_a^1 * f_b^2 \right)(a \circ b \circ q) = \int_G f^1(k) f^2(b^{-1} \circ k^{-1} \circ b \circ q) dk.$$

Now, if f^1 and f^2 are highly concentrated in a small neighborhood of the identity, the only values of k that matter will be close to the identity. The inverse of these values of k also will be close to the identity. The automorphism $k' = b^{-1} \circ k^{-1} \circ b$ preserves closeness to the identity. Therefore, the fact that f^2 is concentrated in a small neighborhood of the identity, and the fact that k' is close to the identity, means that $f^2(k' \circ q)$ forces q to have importance only near the identity.

Since k is close to the identity, $k = I + \sum_{i=1}^{6} k_i \tilde{E}_i$. Then, by definition, we have

$$(k - I)^\vee = \mathbf{k} \in \mathbb{R}^6.$$

Likewise, it can be shown that

$$(b^{-1} \circ k^{-1} \circ b - I)^\vee = -Ad_{b^{-1}} \mathbf{k}$$

where Ad_g is defined by the expression $Ad_g \mathbf{k} = (g(\sum_{i=1}^{6} k_i \tilde{E}_i) g^{-1})^\vee$. See [4] and [19] for the explicit form of $Ad(g)$ as a 6×6 matrix.

Since k and q are both close to the identity, an extension of (9) can be applied to yield

$$\left(f_a^1 * f_b^2 \right)(a \circ b \circ q) = \int_{\mathbf{k} \in \mathbb{R}^6} \tilde{f}^1(\mathbf{k}) \tilde{f}^2(\mathbf{q} - Ad_{b^{-1}} \mathbf{k}) d\mathbf{k}. \tag{10}$$

Note that whereas q is close to the identity, g, in general, will not be, since a and b are not small motions. In order to compute $(f_a^1 * f_b^2)(g)$, one must substitute

$$\mathbf{q} = \left(\log \left[(a \circ b)^{-1} \circ g \right] \right)^\vee$$

into the above expression.

Finally, we note that

$$\left(f_a^1 * f_b^2 \right)(g) \neq \left(f_b^2 * f_a^1 \right)(g)$$

even though

$$(f^1 * f^2)(g) \approx (f^2 * f^1)(g).$$

IV. Form Closure for Convolution and Conditioning of Concentrated Gaussians on $SE(3)$

Let us assume that f^1 and f^2 are both concentrated $SE(3)$-Gaussian functions of the form in (5). This section computes (10) explicitly in closed form, and establishes how the mean and variance of each of the initial concentrated distributions "mix" to result in the mean and variance of their convolution.

We note that while form closure under convolution of Gaussian functions, as defined in (5), results trivially from (9), the case when both functions are shifted is more challenging. Fundamental to all of the calculations in this section is the identity [4]

$$\int_{\mathbb{R}^N} \exp\left(-\frac{1}{2}\mathbf{x}^T M\mathbf{x} - \mathbf{m}^T\mathbf{x}\right) d\mathbf{x} = (2\pi)^{N/2}|\det M|^{-\frac{1}{2}} \exp\left(\frac{1}{2}\mathbf{m}^T M^{-1}\mathbf{m}\right). \quad (11)$$

If $f^i(g)$ are taken to be of the form (5), then direct substitution into (10) and use of (11) with $N = 6$ produces the result

$$C_{1*2} = C_2 - C_2\left(Ad_b^T C_1 Ad_b + C_2\right)^{-1} C_2. \quad (12)$$

We note that since $(Ad_b)^{-1} = Ad_{b^{-1}}$, and that, in general, for invertible matrices of compatible dimensions $(X^TYX)^{-1} = X^{-1}Y^{-1}X^{-T}$, our result can be written in the alternate, slightly more complicated, form

$$C_{1*2} = C_2 - C_2 Ad_{b^{-1}}\left(C_1 + Ad_{b^{-1}}^T C_2 Ad_{b^{-1}}\right)^{-1} Ad_{b^{-1}}^T C_2.$$

Now, the following formula [10, eq. 2.22]:

$$(A + SBT^T)^{-1} = A^{-1} - A^{-1}S(B^{-1} + T^TA^{-1}S)^{-1}T^TA^{-1}$$

which holds when A, B, S, and T are of compatible dimensions and all of the indicated inversions are well defined, can then be used in reverse with $A = \Sigma_2 = C_2^{-1}$, $B = \Sigma_1 = C_1^{-1}$, and $S = T = Ad_{b^{-1}}$ to yield

$$C_{1*2}^{-1} = \Sigma_{1*2} = Ad_{b^{-1}}\Sigma_1 Ad_{b^{-1}}^T + \Sigma_2. \quad (13)$$

This result is one which also can be obtained from the theory of extended Kalman filtering [24], and has been obtained with other arguments [26].

This provides all that is required to propagate error densities in closed form, rather than numerically performing the convolution for the special case of highly concentrated distributions. Returning to error propagation and (3), the results of this section can be seen to be directly applicable by observing that one can define

$$\rho_i(h_i, g_i) = c_i(h_i)\exp\left(\left[\left(\log\left(h_i^{-1}g_i\right)\right)^\vee\right]^T C_i(h_i)\left(\log\left(h_i^{-1}g_i\right)\right)^\vee\right).$$

In other words, each error density is Gaussian shifted from the identity to h_i and, in addition, the covariance matrix ($\Sigma_i = C_i^{-1}$) and scalars c_i depend on the amount of shift.

Conditioning of highly concentrated densities at the identity follows in exactly the same way that it does in $\mathbb{R}^N$.

V. Numerical Examples

Here, we present two examples. In the first subsection, a cascade of two Stewart–Gough platforms, each with small errors in their leg lengths, is analyzed using the covariance propagation method presented earlier. This example is used to verify and validate the part of our formulation that was devoted to small errors. In the second example, the propagation of large backlash in a planar revolute manipulator is analyzed. This example illustrates the universality of the convolution formulation, even in a case when the errors are too large for covariance propagation to be applicable.

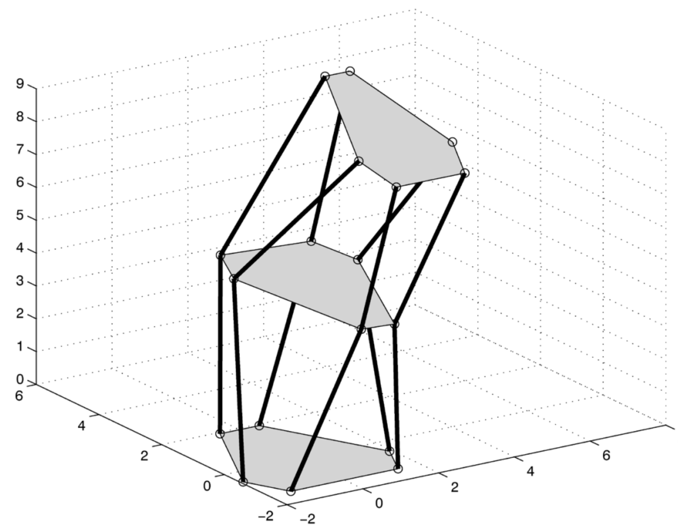

Fig. 2. Hybrid manipulator of two stacked 6-D Stewart platforms.

A. Propagation of Covariances in a Hybrid Serial-Parallel Manipulator

Consider a hybrid manipulator of two stacked 6-D Stewart platforms shown in Fig. 2. For this Stewart platform, the coordinates of the six connection points at the base and the platform are chosen as

$$\begin{pmatrix} 2\sin(2(i-1)\pi/3 \pm \pi/12) \\ 2\cos(2(i-1)\pi/3 \pm \pi/12) \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 2\sin(2(i-1)\pi/3 \pm 11\pi/12) \\ 2\cos(2(i-1)\pi/3 \pm 11\pi/12) \\ 0 \end{pmatrix}$$

for $i = 1, \ldots, 3$, respectively. The configurations of the first and second module are taken as

$$g_1 = \begin{pmatrix} 0.9755 & -0.1010 & 0.1953 & 2 \\ 0.1545 & 0.9469 & -0.2819 & 2 \\ -0.1564 & 0.3052 & 0.9393 & 4 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$g_2 = \begin{pmatrix} 0.9799 & -0.0972 & 0.1742 & 3 \\ 0.1552 & 0.9200 & -0.3599 & 2 \\ -0.1253 & 0.3797 & 0.9166 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The orientation parts of g_1 and g_2 are generated using the Z-X-Y Euler angles, i.e., $(\pi/10, \pi/20, \pi/20)$ for g_1 and

$(\pi/8, \pi/25, \pi/20)$ for g_2. Obviously, when two such platforms are stacked, the frame of reference at the end, g_{ee}, is then $g_1 \circ g_2$.

With given g_1 and g_2, the six leg lengths of the first module can be easily calculated as

$$\mathbf{L1} = [5.2692,\ 4.9044,\ 5.5869,\ 4.9083,\ 3.9383,\ 5.4231]$$

and those of the second module as

$$\mathbf{L2} = [4.8020,\ 5.7855,\ 4.3951,\ 4.2350,\ 5.4216,\ 4.1635].$$

In order to test the covariance formula derived in this paper, we generated small deviations of their leg lengths from the above ideal values by assuming that each leg length has a uniformly random error of $\pm 1\%$. Therefore, each leg length was sampled at three values, $L_i, 0.99L_i$, and $1.01Li$. This generates $n = 3^6$ different frames of reference $\{g_i\}$ that are clustered around g. While this distribution is not Gaussian, as will be seen, the derived covariance propagation method still works reasonably well. Here, g_i is obtained using the forward kinematics method developed in [32].

We compute

$$\mathbf{x}_i = \left[\log(g^{-1} \circ g_i)\right]^\vee$$

and then the "experimental" covariances as

$$\Sigma = \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T \qquad C = \Sigma^{-1}. \tag{14}$$

For leg lengths with $\pm 1\%$ error, the experimental results for the first and second module are computed, respectively, as C_1 and C_2, shown at the bottom of the page.

Using the covariance propagation formula (12), the inverse covariance of the whole manipulator is obtained as C_{1*2}, as shown at the bottom of the page.

To verify the proposed covariance propagation method, brute force enumeration is also used to get the covariance of the whole manipulator directly. In this case, the formula in (14) is used with the n^2 discrete poses obtained by concatenating every element of $\{g_i\}$ with every other, and one obtains C_{brute}, defined as shown at the bottom of the page.

$$C_1 = 10^3 \begin{pmatrix} 3.1164 & 0.3579 & -1.5907 & -0.2254 & -0.1555 & -0.1112 \\ 0.3579 & 2.1469 & -1.6126 & -0.0944 & -0.0159 & 0.1460 \\ -1.5907 & -1.6126 & 2.2213 & 0.2794 & 0.2849 & 0.2097 \\ -0.2254 & -0.0944 & 0.2794 & 0.2730 & 0.3890 & 0.4883 \\ -0.1555 & -0.0159 & 0.2849 & 0.3890 & 0.7561 & 0.8974 \\ -0.1112 & 0.1460 & 0.2097 & 0.4883 & 0.8974 & 1.1640 \end{pmatrix}$$

$$C_2 = 10^3 \begin{pmatrix} 1.3295 & -0.2835 & -0.9809 & -0.1399 & 0.0689 & 0.0446 \\ -0.2835 & 2.0282 & -1.4874 & 0.0446 & 0.0314 & 0.1703 \\ -0.9809 & -1.4874 & 2.8103 & 0.4053 & 0.1312 & 0.0501 \\ -0.1399 & 0.0446 & 0.4053 & 0.8184 & 0.6960 & 0.7052 \\ 0.0689 & 0.0314 & 0.1312 & 0.6960 & 0.7286 & 0.6369 \\ 0.0446 & 0.1703 & 0.0501 & 0.7052 & 0.6369 & 0.6857 \end{pmatrix}$$

$$C_{1*2} = 10^3 \begin{pmatrix} 0.8347 & 0.0347 & -0.6499 & -0.0030 & 0.1347 & 0.0714 \\ 0.0347 & 1.0027 & -0.9893 & -0.0199 & 0.0647 & 0.1127 \\ -0.6499 & -0.9893 & 1.7318 & 0.0424 & -0.1788 & -0.1484 \\ -0.0030 & -0.0199 & 0.0424 & 0.1227 & 0.1491 & 0.1113 \\ 0.1347 & 0.0647 & -0.1788 & 0.1491 & 0.2671 & 0.1708 \\ 0.0714 & 0.1127 & -0.1484 & 0.1113 & 0.1708 & 0.1418 \end{pmatrix}$$

$$C_{\text{brute}} = 10^3 \begin{pmatrix} 0.8339 & 0.0327 & -0.6474 & -0.0045 & 0.1324 & 0.0696 \\ 0.0327 & 0.9992 & -0.9840 & -0.0214 & 0.0616 & 0.1105 \\ -0.6474 & -0.9840 & 1.7244 & 0.0460 & -0.1726 & -0.1440 \\ -0.0045 & -0.0214 & 0.0460 & 0.1181 & 0.1417 & 0.1060 \\ 0.1324 & 0.0616 & -0.1726 & 0.1417 & 0.2549 & 0.1623 \\ 0.0696 & 0.1105 & -0.1440 & 0.1060 & 0.1623 & 0.1358 \end{pmatrix}$$

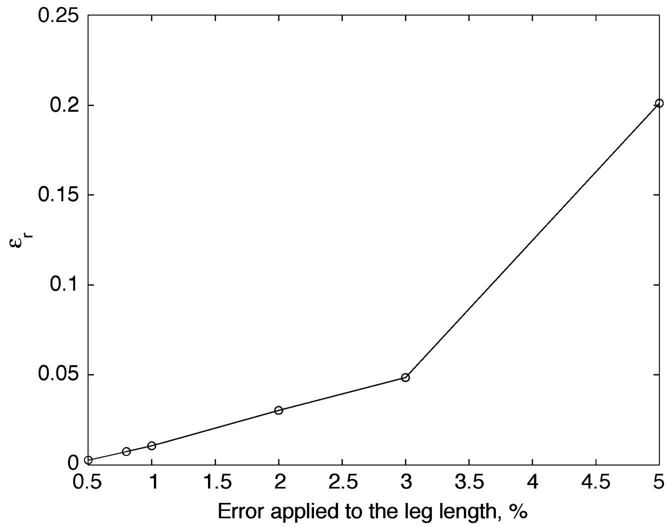

Fig. 3. Deviation of the proposed propagation covariance method relative to brute-force enumeration.

As can be seen, these results are in excellent agreement, which serves as a demonstration and validation of the derived formula for the case of small errors. This agreement is quantified in a single number defined using the Hilbert–Schmidt (Frobenius) norm as

$$\epsilon_r = \frac{||C_{1*2} - C_{\mathrm{brute}}||}{||C_{\mathrm{brute}}||}$$

where ϵ_r is the deviation in the C computed by covariance propagation relative to that generated by brute force, and $||\cdot||$ denotes the Hilbert–Schmidt (Frobenius) norm. For leg lengths with $\pm 1\%$ error, we found $\epsilon_r = 0.0113 = 1.13\%$.

Of course, it is of interest to know what happens in the case of other smaller and larger errors, and so we have repeated this experiment with $\pm 0.5\%$, $\pm 0.8\%$, $\pm 1\%$, $\pm 2\%$, $\pm 3\%$, and $\pm 5\%$ errors on leg lengths. The trend is graphed in Fig. 3. Clearly, the approximations used in the derivation of covariance propagation break down as the errors become large.

The following subsection considers a different example in which the errors are large, and the value of the convolution formulation presented earlier is demonstrated.

B. Large Backlash Propagation in a Revolute Manipulator by Convolution

Consider the three-link planar revolute manipulator shown in Fig. 4. Each rigid link has length L, and each joint has some backlash that is described by a probability distribution $f(\theta - \theta_0)$ centered around the value $\theta_0 = 30$ degrees. The error density for a single link is then of the form

$$\rho\left(g(r,\phi,\theta)\right) = f(\theta-\theta_0)\delta(\phi-\theta)\delta(r-L)/r \qquad (15)$$

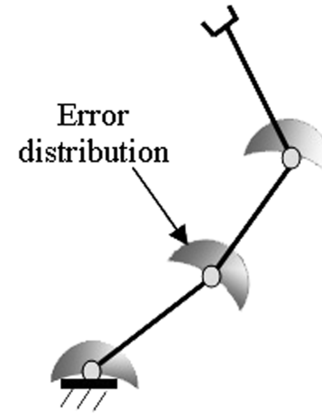

Fig. 4. Three-link planar manipulator with joint backlash.

where $\delta(\cdot)$ is the usual Dirac-delta function in one dimension, and an arbitrary element of $SE(2)$ is parametrized as

$$g(r,\phi,\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & r\cos\phi \\ \sin\theta & \cos\theta & r\sin\phi \\ 0 & 0 & 1 \end{pmatrix}.$$

The associated volume element for this parameterization is $d(g) = r dr d\phi d\theta$, and integration over $G = SE(2)$ is integration over all values of $r \in \mathbb{R}^+$ and $\phi, \theta \in [0, 2\pi]$.

In (15), the delta functions enforce the rigidity of the links, and division by r is due to the r in the volume element. The function $f(\theta)$ has its mode at 0, but the backlashes can be potentially large (i.e., not highly concentrated). For this reason, we cannot take it to be a Gaussian, but rather, a folded Gaussian of the form

$$\begin{aligned} f(\theta,\sigma) &= \frac{1}{2\pi}\sum_{n=-\infty}^{\infty} e^{-n^2\sigma^2} e^{in\theta} \\ &= \frac{1}{\sqrt{2\pi}\sigma}\sum_{n=-\infty}^{\infty} e^{-(\theta-2\pi n)^2/2\sigma^2}. \end{aligned} \qquad (16)$$

The error density that accumulates at the end-effector due to backlashes in each of the joints is computed as the convolution

$$\rho_{ee}(g) = (\rho * \rho * \rho)(g).$$

Computing this numerically by the definition of convolution is not as convenient as using the $SE(2)$-convolution theorem and the corresponding concept of Fourier transform, which is what we shall do here.

The Fourier transform of a function on $G = SE(2)$ is defined as

$$\hat{f}(p) = \int_G f(g) U(g^{-1}, p) d(g) \qquad (17)$$

where $U(g,p)$ is an infinite-dimensional unitary matrix called an irreducible unitary representation (IUR) [4]. It possesses the important homomorphism property $U(g_1 \circ g_2, p) = U(g_1,p)U(g_2,p)$. One can show that the generalization of the classical Fourier transform in (17) admits a convolution theorem due to the homomorphism property of $U(g,p)$, and that the following inverse transform can be used to reconstruct the original function:

$$f(g) = \int_0^\infty \text{trace}\left(\hat{f}(p)U(g,p)\right) pdp. \tag{18}$$

This is because the matrix elements of the full set of IURs form an orthonormal basis with which to expand functions on $SE(2)$.

A number of works, including [4], [27], [28], and [30], have shown that the matrix elements of the IURs for $SE(2)$ can be expressed as

$$u_{mn}\left(g(r,\phi,\theta),p\right) = i^{n-m} e^{-i[n\theta+(m-n)\phi]} J_{n-m}(pr) \tag{19}$$

where $J_\nu(x)$ is the νth-order Bessel function, and m and n take values in the integers.

From this expression, and the fact that $U(g,p)$ is a unitary representation, we have that

$$\begin{aligned} u_{mn}\left(g^{-1}(r,\phi,\theta),p\right) &= u_{mn}^{-1}\left(g(r,\phi,\theta),p\right) \\ &= \overline{u_{nm}\left(g(r,\phi,\theta),p\right)} \\ &= i^{n-m} e^{i[m\theta+(n-m)\phi]} J_{m-n}(pr). \end{aligned} \tag{20}$$

Computing the $SE(2)$-Fourier transform of the one-link backlash-error density in (15), one finds (after the delta functions kill the integrals over r and ϕ) that

$$\begin{aligned} \hat{\rho}_{mn}(p) &= i^{n-m} J_{m-n}(pL) \int_0^{2\pi} f(\theta-\theta_0) e^{in\theta} d\theta \\ &= i^{n-m} J_{m-n}(pL) e^{in\theta_0 - n^2\sigma^2}. \end{aligned} \tag{21}$$

Using the convolution theorem, we compute $\hat{\rho}_{ee}(p) = \hat{\rho}\hat{\rho}\hat{\rho}$, where the matrix elements of $\hat{\rho}$ are given by (21). Then, the original error density $\rho_{ee}(g(r,\phi,\theta))$ can be reconstructed by applying the Fourier-inversion formula (18) to $\hat{\rho}_{ee}(p)$. Since it is difficult to view the error density $\rho(g)$ graphically, the marginal density function $\rho(r,\phi)$ is computed. The marginal density function $\rho(r,\phi)$ is found by just integrating the Fourier reconstruction formula (18) for $\rho(g)$, with respect to θ, as

$$\begin{aligned} \rho(r,\phi) &= \frac{1}{2\pi}\int_0^{2\pi} \rho(g) d\theta \\ &= \sum_{n\in\mathbf{Z}} i^{-n} e^{-in\phi} \int_0^\infty \hat{f}_{0n}(p) J_{-n}(pr) pdp. \end{aligned}$$

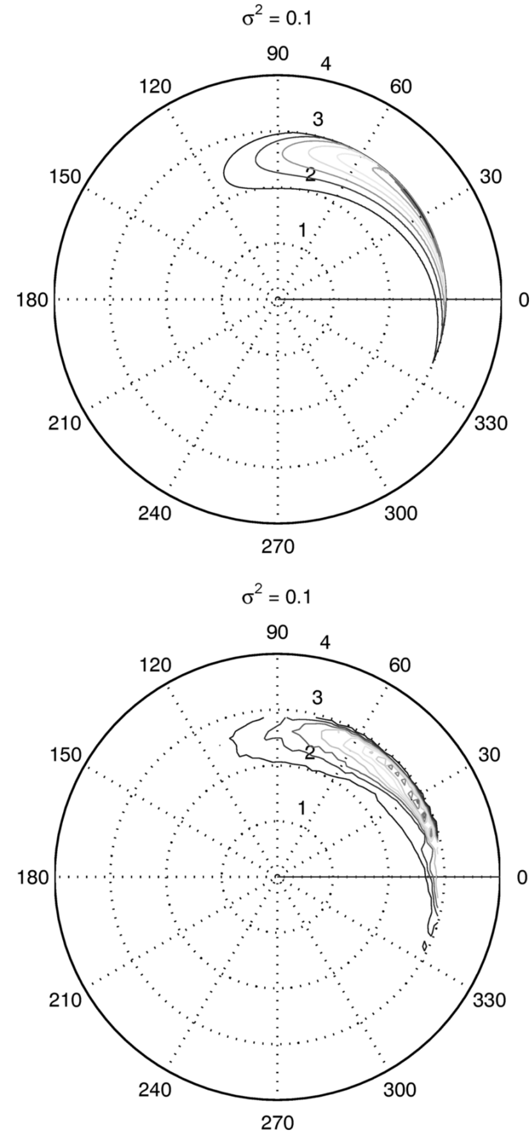

Fig. 5. Marginal error density $\rho(r,\phi)$ for $\sigma^2 = 0.1$.

To validate the results obtained from our convolution-based error-propagation method, the error distribution (16) is sampled and applied to each joint of the manipulator directly. Then, brute-force enumeration is used to obtain the error distribution directly.

The marginal error densities $\rho(r,\phi)$ obtained from both methods are plotted in Figs. 5 and 6, with the top one from the propagation method, and the bottom one from brute force. The variance σ^2 of 0.1 is given in Fig. 5, and σ^2 of 0.3 is given in Fig. 6.

For the above computations, the link length L is taken as 1, and 60 sample points are generated for the distribution (16). The infinite-dimensional matrix function $U(g,p)$ in the $SE(2)$ Fourier transform is truncated at finite values of $|m|$, $|n| \leq 10$ (i.e., the dimension of $U(g,p)$ is $2\times 10+1$). The bandlimited approximation still gives very accurate results, because the magnitude of the Fourier transform of a sufficiently smooth function

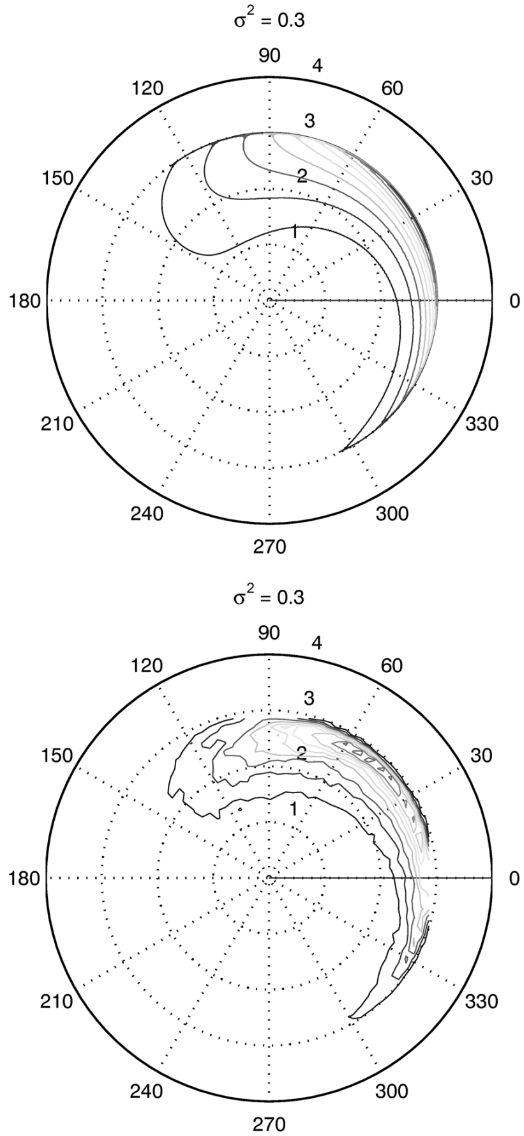

Fig. 6. Marginal error density $\rho(r, \phi)$ for $\sigma^2 = 0.3$.

can be ignored beyond a certain cutoff frequency. The frequency parameter p is sampled in the interval of 300 with an integration step of 0.2. All the calculations in this example took less than 3 min using Matlab with a 1.0-GHz, 516-MB RAM computer.

VI. Conclusion

Quantifying the intuitive notion of how spatial errors "add" has been addressed in this paper. It was shown that even though the concept of a Gaussian distribution does not completely generalize when considering the case of Lie-group-valued arguments, an appropriate concept does exist when considering highly concentrated distributions. This paper worked out the details of how Gaussian distributions are defined in this context, what their properties are, and how they can be applied to compute the propagation of covariances in serial manipulators. Properties of these distributions were proven. The computations performed show that such distributions have the desired closure properties in order for them to be useful in estimation problems.

Appendix

A. Metrics on Rigid-Body Displacements

Several metrics have been proposed in the kinematics literature to measure displacements between rigid bodies [5], [7], [11], [17], [20].

It can be shown that the following is a metric:

$$d(g_1, g_2) = \left(\frac{1}{2}\mathrm{tr}\left[(H(g_1) - H(g_2))\, W (H(g_1) - H(g_2))^T\right]\right)^{\frac{1}{2}} \tag{22}$$

where

$$W = \begin{pmatrix} J & \mathbf{0} \\ \mathbf{0}^T & M \end{pmatrix}$$

contains inertial information about the rigid body that is being moved from g_1 to g_2. In particular, if the body has mass density $\rho(\mathbf{x})$, then $M = \int_{\mathbb{R}^3} \rho(\mathbf{x}) d\mathbf{x}$ and $J = \int_{\mathbb{R}^3} \mathbf{x}\mathbf{x}^T \rho(\mathbf{x}) d\mathbf{x}$. This naturally reconciles the difference in units used to measure translations and rotations. In other words, the body that is undergoing the motion itself defines (through its mass density) how rotations should be weighted relative to translations. The metric $d(g_1, g_2)$ discussed here is left invariant [5]. A compatible inner product satisfying *Definition 1* is

$$(X, Y) = \frac{1}{2}\mathrm{tr}[XWY^T]$$

where $X, Y \in se(3)$.

A second metric that can be used for $SE(3)$ is

$$d'(g_1, g_2) = \left\| \log\left(g_1^{-1} g_2\right)^{\vee} \right\|.$$

In a Lie-thoeretic setting, this may be a more natural metric. However, in some applications, the issue of differences in units between orientational and translational quantities must be addressed. An inner product compatible with this metric is

$$(X, Y) = \frac{1}{2}\mathrm{tr}[XY^T]$$

where $X, Y \in se(3)$.

B. Integration Over Rigid-Body Motions

The body-fixed Jacobian for $SE(3)$ parametrized with $q_1, \ldots, q_6$ is [18], [19]

$$\mathcal{J}(\mathbf{q}) = \left[\left(H^{-1}\frac{\partial H}{\partial q_1}\right)^{\vee}, \cdots, \left(H^{-1}\frac{\partial H}{\partial q_6}\right)^{\vee}\right]. \tag{23}$$

In general, the volume element with which to integrate over motions will be of the form

$$d(H) = |\det(\mathcal{J})|\, dq_1 \cdots dq_6.$$

It can be shown that the same result occurs whether the body-fixed or base-fixed Jacobian is used.

C. Statistics on Groups

The concepts of mean, variance, and covariance are well defined for probability densities in $\mathbb{R}^N$. Let

$$\sigma^2(\mathbf{y}) = \int_{\mathbf{R}^N} d^2(\mathbf{x}, \mathbf{y})\rho(\mathbf{x})d\mathbf{x}$$

where $d(\mathbf{x}, \mathbf{y}) = ||\mathbf{x} - \mathbf{y}||$. Then the mean, or expected value, $E[\mathbf{x}]$ is the value of $\mathbf{y}$ which minimizes $\sigma^2(\mathbf{y})$. The straightforward generalization of this to pdfs on groups is that the expected value $E_d[g_1] \in G$ is the group element which minimizes the function

$$\sigma_d^2(g_1) = \int_G [d(g_1, g_2)]^2 \rho(g_2)d(g_2). \quad (24)$$

Here, $d(g_1, g_2)$ is a metric (not to be confused with the integration measure $d(g)$), and clearly, the center of mass in this case depends on how this metric is defined. Hence, $E_d[g]$ is called the d-mean, and the value $\sigma_d^2(E_d[g])$ is what we will refer to as the d-variance. These definitions are known in the theoretical statistics literature (e.g., see [6]), but are not part of what is generally considered to be standard engineering mathematics. Other issues relating to pdfs on groups in general (and $SO(N)$, in particular) are addressed in [12], [16], and [22].

In traditional statistics in $\mathbb{R}^N$, $N \times N$ covariance matrices with entries of the form

$$\sigma_{ij} = \int_{\mathbf{R}^N} x_i x_j \rho\left(\mathbf{x} + E[\mathbf{x}]\right) d\mathbf{x}$$

play an important role. Here x_i can be viewed as a (signed) distance from the origin to a point on the ith coordinate axis. A natural extension of this concept to the group-theoretic setting is

$$\sigma_{ij}^d = \int_G \operatorname{sgn}[\alpha_i \alpha_j] d\left(e, \exp(\alpha_i \tilde{E}_i)\right) \times d\left(e, \exp(\alpha_j \tilde{E}_j)\right) \rho\left(E_d[g] \circ g\right) dg. \quad (25)$$

We note that when using the metric $d'(g_1, g_2) = \| \log(g_1^{-1} g_2)^\vee \|$, the above covariance formula essentially reduces to the standard definition, since $d'(e, \exp(\alpha_j \tilde{E}_j)) = |\alpha_j|$. Therefore, this is a very natural choice in this context. In the body of this paper, when calculating covariances, we are calculating $\sigma_{ij}^{d'}$.

In contrast, while the other metric is applicable in many contexts, it introduces the weighting matrix W into the calculations. In order to use a covariance calculated with this metric in the definition of the $SE(3)$-Gaussian distribution, the fact that the units in σ_{ij}^d are homogeneous over all values of i and j would need to be modified to account for the fact that x_1, x_2, x_3, x_4, x_5, and x_6 have different units. In other words, if, for example, the weighting matrix in (22) has $W = r^2 I$ and $m = 1$, where r is a scalar and I is the identity, then σ_{ij}^d will be divided by r^2 when both i and j are rotational, and divided by r when either i or j (but not both) are rotational. This reduces the definition of covariances to be exactly the same as in (6) for the case of highly concentrated distributions.

Acknowledgment

The authors would like to thank the reviewers of this paper for the many useful comments that they contributed.

References

[1] D. Blackmore and M. C. Leu, "Analysis of swept volume via Lie groups and differential equations," *Int. J. Robot. Res.*, vol. 11, no. 6, pp. 516–537, Dec. 1992.

[2] R. W. Brockett, "System theory on group manifolds and coset spaces," *SIAM J. Control*, vol. 10, no. 2, pp. 265–284, May 1972.

[3] F. Bullo and R. M. Murray, "Tracking for fully actuated mechanical systems: A geometric framework," *Automatica*, vol. 35, no. 1, pp. 17–34, Jan. 1999.

[4] G. S. Chirikjian and A. B. Kyatkin, *Engineering Applications of Noncommutative Harmonic Analysis*. Boca Raton, FL: CRC, 2001.

[5] G. S. Chirikjian and S. Zhou, "Metrics on motion and deformation of solid models," *J. Mech. Des.*, vol. 120, pp. 252–261, Jun. 1998.

[6] P. Diaconis, *Group Representations in Probability and Statistics*, ser. Lecture Notes-Monograph Series, S. S. Gupta, Ed. Hayward, CA: Inst. Math. Statist., 1988.

[7] P. Fanghella and C. Galletti, "Metric relations and displacement groups in mechanism and robot kinematic," *Trans. ASME J. Mech. Des.*, vol. 117, pp. 470–478, Sep. 1995.

[8] Y. M. Han, "Simultaneous translational and rotational tracking in dynamic, environments: Theoretical and practical viewpoints," *IEEE Trans. Robot. Autom.*, vol. 20, no. 2, pp. 309–318, Apr. 2004.

[9] Y. Han and F. C. Park, "Least squares tracking on the Euclidean group," *IEEE Trans. Autom. Control*, vol. 46, no. 7, pp. 1127–1132, Jul. 2001.

[10] K. Kanatani, *Statistical Optimization for Geometric Computation: Theory and Practice*. Mineola, NY: Dover, 2005.

[11] K. Kazerounian and J. Rastegar, "Object norms: A class of coordinate and metric independent norms for displacement," *ASME Trans. Flexible Mechan., Dyn., Anal.*, vol. 47, pp. 271–275, 1992.

[12] P. T. Kim, "Deconvolution density estimation on $SO(N)$," *Ann. Statist.*, vol. 26, no. 3, pp. 1083–1102, Jun. 1998.

[13] A. B. Kyatkin and G. S. Chirikjian, "Applications of noncommutative harmonic analysis in robotics," in *Proc. 12th CISM-IFToMM Symp. Theory Practice Robots Manip.*, Paris, France, Jul. 1998, pp. 119–126.

[14] N. E. Leonard and P. S. Krishnaprasad, "Motion control on drift-free, left-invariant systems on Lie groups," *IEEE Trans. Autom. Control*, vol. 40, no. 9, pp. 1539–1554, Sep. 1995.

[15] J. T.-H. Lo and L. R. Eshleman, "Exponential Fourier densities on $SO(3)$ and optimal estimation and detection for rotational processes," *SIAM J. Appl. Math.*, vol. 36, no. 1, pp. 73–82, Feb. 1979.

[16] J.T.-Y. Lo and S. K. Ng, "Characterizing Fourier-series representation of probability-distributions on compact Lie groups," *SIAM J. Appl. Math.*, vol. 48, no. 1, pp. 222–228, Feb. 1988.

[17] J. M. R. Martinez and J. Duffy, "On the metrics of rigid body displacement for infinite and finite bodies," *ASME J. Mech. Des.*, vol. 117, pp. 41–47, Mar. 1995.

[18] J. M. McCarthy, *An Introduction to Theoretical Kinematics*. Cambridge, MA: MIT Press, 1990.

[19] R. M. Murray, Z. Li, and S. S. Sastry, *A Mathematical Introduction to Robotic Manipulation*. Ann Arbor, MI: CRC, 1994.

[20] F. C. Park, "Distance metrics on the rigid-body motions with applications to mechanism design," *Trans. ASME J. Mech. Des.*, vol. 117, pp. 48–54, Mar. 1995.

[21] F. C. Park and R. W. Brockett, "Kinematic dexterity of robotic mechanisms," *Int. J. Robot. Res.*, vol. 13, no. 1, pp. 1–15, Feb. 1994.

[22] J. S. Rosenthal, "Random rotations—characters and random-walks on $SO(N)$," *Ann. Probabil.*, vol. 22, no. 1, pp. 398–423, Jan. 1994.

[23] J. M. Selig, *Geometrical Methods in Robotics*. New York: Springer-Verlag, 1996.

[24] P. Smith, T. Drummond, and K. Roussopoulos, "Computing MAP trajectories by representing, propagating and combining PDFs over groups," in *Proc. 9th IEEE Int. Conf. Comput. Vision*, Nice, France, Oct. 2003, vol. II, pp. 1275–1282.

[25] 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.

[26] S. Su 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, Jun. 1992.
[27] M. Sugiura, *Unitary Representations and Harmonic Analysis*, 2nd ed. Amsterdam, The Netherlands: Elsevier, 1990.
[28] J. Talman, *Special Functions*. Amsterdam, The Netherlands: W. A. Benjamin, 1968.
[29] V. S. Varadarajan, *Lie Groups, Lie Algebras, and Their Representations*. New York: Springer-Verlag, 1984.
[30] N. J. Vilenkin and A. U. Klimyk, *Representation of Lie Group and Special Functions*. Amsterdam, The Netherlands: Kluwer, 1991, vol. 1–3.
[31] Y. F. Wang and G. S. Chirikjian, "Workspace generation of hyper-redundant manipulators as a diffusion process on $SE(N)$," *IEEE Trans. Robot. Autom.*, vol. 20, no. 3, pp. 399–408, Jun. 2004.
[32] Y. F. Wang, "An incremental method for forward kinematics of parallel manipulators," in *Proc. IEEE Int. Conf. Robot., Autom. Mechatron.*, Bangkok, Thailand, Jun. 2006, pp. 243–247.
[33] J. T.-Y. Wen and K. Kreutz-Delgado, "The attitude control problem," *IEEE Trans. Autom. Control*, vol. 36, no. 10, pp. 1148–1162, Oct. 1991.

Yunfeng Wang (M'04) received the B.S. and M.S. degrees in mechanical engineering from Tianjin University, Tianjin, China, in 1992 and 1995, respectively, and the M.S.E. degree in electrical and computer engineering and the Ph.D. degree in mechanical engineering from Johns Hopkins University, Baltimore, MD, in 2001 and 2002, respectively.

She is currently an Assistant Professor with the Department of Mechanical Engineering, The College of New Jersey, Ewing, NJ. Her research interests include manipulator kinematics and dynamics, computational methods, autonomous vehicles, and intelligent manufacturing.

Gregory S. Chirikjian (M'93) was born August 16, 1966, in New Brunswick, NJ. He received the B.S.E. degree in engineering mechanics, the M.S.E. degree in mechanical engineering, and the B.A. degree in mathematics, all from The Johns Hopkins University, Baltimore, MD, in 1988, and the Ph.D. degree from the California Institute of Technology, Pasadena, in 1992.

Since 1992, he has been on the faculty of the Department of Mechanical Engineering, Johns Hopkins University, where he is now Professor and Chair. His research interests include the kinematic analysis, motion planning, design, and implementation of hyper-redundant, metamorphic, and binary manipulators. In recent years, he has expanded the scope of his research to include applications of group theory in a variety of engineering disciplines and the mechanics of biological macromolecules.

Dr. Chirikjian is a 1993 National Science Foundation Young Investigator, a 1994 Presidential Faculty Fellow, and a 1996 recipient of the ASME Pi Tau Sigma Gold Medal.