Robotica (2009) volume 27, pp. 861–872. © Cambridge University Press 2009
doi:10.1017/S0263574708005262

Deconvolution on the Euclidean motion group and planar robotic manipulator design

Peter T. Kim†*, Yan Liu‡, Zhi-Ming Luo† and Yunfeng Wang§

†*Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario N1G 2W1 Canada.*
‡*Google New York, 76 9th Ave. 4th Floor New York, NY 10011.*
§*Department of Mechanical Engineering, The College of New Jersey, Ewing, NJ 08628-0718, USA.*

(Received in Final Form: December 2, 2008. First published online: January 15, 2009)

SUMMARY
Several problems of practical interest in robotics can be modelled as the convolution of functions on the Euclidean motion group. These include the evaluation of reachable positions and orientations at the distal end of a robot manipulator arm. A natural inverse problem arises when one wishes to design rather than to model manipulators. Namely, by considering a serial-chain robot arm as a concatenation of segments, we examine how statistics of known segments can be used to select, or design, the remainder of the structure so as to attain the desired statistical properties of the whole structure. This is then a deconvolution density estimation problem for the Euclidean motion group. We prove several results about the convergence of these deconvolution estimators to the true underlying density under certain smoothness assumptions. A practical implementation to the design of planar robot arms is demonstrated.

KEYWORDS: Degenerate diffusion; Fourier analysis; Gaussian distribution; Inverse problem; Irreducible representations; Kinematics; Manipulator arm.

1. Introduction

A robotic manipulator arm is a device that is used to position and orient objects. A manipulator is constructed of rigid links and actuators, such as motors or hydraulic cylinders, which cause all motion of the arm. If the actuators have only a finite number of states, as is the case with stepper motors or pneumatic cylinders, the arm has a finite number of configurations and only a finite number of reference frames (i.e. elements of the rigid-body-motion group) are reachable by the hand. Discrete-state manipulator arms have been studied almost from the beginning of the field of robotics[1,2], and were independently studied in the 1980s in the former Soviet Union.[3] Since the mid 1990s there has been a revival of interest in these devices.[4,5] The set of all reachable positions and orientations is called the workspace. Studying properties of the workspace, as well as designing manipulators that have desired workspace properties are topics which have received considerable attention.[6,7] In the Introduction of ref. [8], a survey of various classes of manipulators is detailed (see also ref. [9]).

* Corresponding author. E-mail: pkim@uoguelph.ca

One can think of a two-dimensional discrete-state manipulator as an object which is constructed out of a series of platforms that are stacked, or cascaded, on top of each other (see for example Fig. 2 in ref. [10]). Since each actuator (which in this case is a pneumatic cylinder in parallel with a viscous dashpot) has two stable states, a very large but finite number of states are reachable by the end. This kind of arm is attractive because it requires no feedback control, and is very inexpensive to construct.

For discretely actuated manipulators, Chirikjian and Ebert-Uphoff observed that the density of reachable frames determines how accurately a random position and orientation can be reached.[11,12] This density information is also important in planning the motions of discretely actuated manipulator arms.[13] Density is calculated in principle by dividing a compact subset of the Euclidean motion group containing the workspace into finite but small volume elements.[11,12] The number of positions and orientations reachable by the end of the manipulator which lie in each volume element is stored. Dividing this number by the volume element size gives the average density in each element.

It is an important aspect of the manipulator design problem to specify the density of reachable frames throughout the workspace. That is, areas which must be reached with great accuracy should have high density, and those areas of the workspace which are less important need less density. However, to compute this workspace density function using brute force enumeration of states is computationally intractable for a large number of segments, e.g. it requires a huge number of evaluations of the kinematic equations relating actuator state to the resulting end frame for a manipulator.

In the current presentation, we will restrict the discussion to planar robot arms. In this context, the concept of convolution of real-valued functions on $\mathbb{SE}(2)$, the three-dimensional Euclidean motion group, provides a powerful computational tool for computing this density efficiently.[11,12] Let us denote an arbitrary element of $\mathbb{SE}(2)$ as x. If we imagine that the manipulator is divided into two connected parts, a density function $k(x)$ can be associated with those frames reachable by the end of the lower half of the manipulator, and a density function $f(x)$ can be associated with the end of the upper half of the manipulator. The density k is defined relative to the base frame, and f treats the frame at the end of the lower

segment as the base frame. That is $f(x) = k(x)$ when the manipulator is cut into two equal parts and there are an even number of identical modules. However, k and f will not be the same function in more general scenarios. By adjusting kinematic parameters such that actuator strokes are limited or extended, the set of reachable frames (and thus the density) is altered. This is achieved mechanically by simply inserting or removing rigid stoppers that specify the physical actuator length corresponding to the discrete states.

While it may not be possible to calculate all of the frames to compute the density function of the workspace, it is often feasible to compute a smaller number of frames for each of the two segments. The density of the whole workspace is then generated by the convolution of these two functions:

$$\int_{\mathbb{SE}(2)} k(y) f(y^{-1}x)\,\mathrm{d}y = h(x).$$

Our objective then is to reconstruct the density f based on observations from h when k is known.

We would also like to add that the implementation of the proposed deconvolution method essentially involves a non-linear optimization with respect to a low-dimensional parameter space using gradient descent. The main technical contribution of this paper involves the fact that because we are approaching this problem in the natural geometry of the Euclidean motion group $\mathbb{SE}(2)$, we are essentially embedding the problem in the constraint set and therefore are not burdened by invoking boundary conditions if approached as a constraint optimization problem. This allows us to obtain error bounds with respect to the sample size and truncation level without having to specify boundary conditions.

We will now provide a summary of this paper. In Section 2, we review the representation theory of the Euclidean motion group of the plane and the corresponding Fourier analysis. In Section 3, we go over diffusions on $\mathbb{SE}(2)$ and, in particular, the Gaussian and degenerate distributions are introduced. In Section 4, we state precisely the deconvolution problem along with the deconvolution density estimator, while in Section 5 we present numerical results along with a demonstration of a practical implementation of the deconvolution procedure. In Section 6, we present some technical results which demonstrate some desirable theoretical properties of our procedure. Finally, all proofs and numerical discussions are collected in the Appendix section.

2. Notation

We will use this section to set the notations; we will closely follow those of the Section III in ref. [8].

The Euclidean motion group $\mathbb{SE}(2)$ is the semidirect product of the special 2×2 orthogonal group $\mathbb{SO}(2)$ with $\mathbb{R}^2$. That is $\mathbb{SE}(2) = \mathbb{SO}(2) \rtimes \mathbb{R}^2$. We denote elements of $\mathbb{SE}(2)$ as $g = (A, \mathbf{a}) \in \mathbb{SE}(2)$ where $A \in \mathbb{SO}(2)$ and $\mathbf{a} \in \mathbb{R}^2$. For any $g_1 = (A_1, \mathbf{a}_1)$ and $g_2 = (A_2, \mathbf{a}_2)$ in $\mathbb{SE}(2)$, the group law is written as $g_1 g_2 = (A_1 A_2, \mathbf{a}_1 + A_1 \mathbf{a}_2)$, and $g^{-1} = (A', -A'\mathbf{a})$, where superscript $'$ means transpose. The unit element is $(\mathbf{I}_2, \mathbf{0}) \in \mathbb{SE}(2)$, where $\mathbf{I}_\nu$ is the $\nu \times \nu$ identity matrix and $\mathbf{0}' = (0, 0)$.

It is convenient to think of elements in $\mathbb{SE}(2)$ as 3×3 matrices

$$g = \begin{pmatrix} A & \mathbf{a} \\ \mathbf{0}' & 1 \end{pmatrix}, \tag{2.1}$$

where $A \in \mathbb{SO}(2)$, $\mathbf{a} \in \mathbb{R}^2$. In polar coordinates, (2.1) would be written as

$$g(r, \theta, \phi) = \begin{pmatrix} \cos\phi & -\sin\phi & r\cos\theta \\ \sin\phi & \cos\phi & r\sin\theta \\ 0 & 0 & 1 \end{pmatrix}, \tag{2.2}$$

where $\phi, \theta \in [0, 2\pi)$ and $r \geq 0$. Then the group action and inverse is just matrix multiplication and inversion of (2.2), respectively.

For the unit element $(\mathbf{I}_2, \mathbf{0}) \in \mathbb{SE}(2)$, define the Lie algebra $\mathfrak{se}(2)$ as the tangent space of $\mathbb{SE}(2)$ at the unit element. By differentiating and evaluating Eq. (2.2) at the unit element, one can see that $\mathfrak{se}(2) = \mathfrak{so}(2) + \mathbb{R}^2$, a vector space sum, where $\mathfrak{so}(2)$ is the Lie algebra of $\mathbb{SO}(2)$, i.e. the tangent space of the latter at $\mathbf{I}_2$. In particular, we have the following matrices as a basis for $\mathfrak{se}(2)$:

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

$$\mathfrak{X}_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \tag{2.3}$$

where $\mathfrak{X}_1$ is a basis for $\mathfrak{so}(2)$ and $\mathfrak{X}_2$ and $\mathfrak{X}_3$ are the basis for $\mathbb{R}^2$.

We note that (2.3), collectively, represent the coordinate axes of $\mathfrak{se}(2)$, the tangent space of $\mathbb{SE}(2)$ at the unit element. An arbitrary $\mathfrak{X} \in \mathfrak{se}(2)$ will therefore be a linear combination of (2.3) and thus for small displacements from the unit element, one can consider the one-parameter subgroup $\exp(t\mathfrak{X})$ of $\mathfrak{se}(2)$, where exp: $\mathfrak{se}(2) \to \mathbb{SE}(2)$ is the exponential map. The left invariant vector field on $\mathfrak{se}(2)$ can now be defined by

$$\tilde{\mathfrak{X}} f = \frac{\mathrm{d}}{\mathrm{d}t} f(g \exp(t\mathfrak{X}))\Big|_{t=0},$$

where $f : \mathbb{SE}(2) \to \mathbb{R}$. This is the directional derivative of f at $g \in \mathbb{SE}(2)$ in the direction $\exp(t\mathfrak{X})$. Thus, with respect to the bases (2.3), we have

$$\exp(t\mathfrak{X}_1) = \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

https://doi.org/10.1017/S0263574708005262 Published online by Cambridge University Press

$$\exp(t\mathfrak{X}_2) = \begin{pmatrix} 1 & 0 & t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\exp(t\mathfrak{X}_3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}, \tag{2.4}$$

for $t \in \mathbb{R}$. In polar coordinates, the left invariant vector fields are

$$\begin{aligned} \tilde{\mathfrak{X}}_1 &= \frac{\partial}{\partial \phi}, \\ \tilde{\mathfrak{X}}_2 &= \cos(\phi - \theta)\frac{\partial}{\partial r} + \frac{\sin(\phi - \theta)}{r}\frac{\partial}{\partial \theta}, \\ \tilde{\mathfrak{X}}_3 &= -\sin(\phi - \theta)\frac{\partial}{\partial r} + \frac{\cos(\phi - \theta)}{r}\frac{\partial}{\partial \theta}. \end{aligned} \tag{2.5}$$

In rectangular coordinates, the left invariant vector fields are

$$\begin{aligned} \tilde{\mathfrak{X}}_1 &= \frac{\partial}{\partial \phi}, \\ \tilde{\mathfrak{X}}_2 &= \cos(\phi)\frac{\partial}{\partial x} - \sin(\phi)\frac{\partial}{\partial y}, \\ \tilde{\mathfrak{X}}_3 &= \sin(\phi)\frac{\partial}{\partial x} + \cos(\phi)\frac{\partial}{\partial y}. \end{aligned} \tag{2.6}$$

Thus by using either (2.5) or (2.6), we can define differential operators on $\mathbb{SE}(2)$ as we will do in Section 3.

2.1. Irreducible unitary representations

The collection of inequivalent irreducible unitary representations of $\mathbb{SE}(2)$ denoted by $\widehat{\mathbb{SE}(2)}$ is characterized by a positive real number $p \in \mathbb{R}_+ = [0, \infty)$. The unitary representation $U(g, p)$ has the property

$$U(g, p) = U(A, \mathbf{a}, p) = U(\mathbf{I}_2, \mathbf{a}, p)\, U(A, 0, p).$$

After an appropriate orthonormal basis of the space of square integrable functions on the unit circle, $L^2(\mathbb{S}^1)$, is chosen, the matrix elements of $U(g, p)$ are expressed as

$$u_{\ell m}(g, p) = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{e}^{-\mathrm{i}\ell\psi}\mathrm{e}^{-\mathrm{i}(a_1 p\cos\psi + a_2 p\sin\psi)}\mathrm{e}^{\mathrm{i}m(\psi-\phi)}\mathrm{d}\psi,$$

where $\ell, m \in \mathbb{Z}$, $\mathrm{i}^2 = -1$ and $g = (A, \mathbf{a}) = (\phi, a_1, a_2)$ parameterized in rectangular coordinates, or

$$u_{\ell m}(g, p) = \mathrm{i}^{m-\ell}\mathrm{e}^{-\mathrm{i}[m\phi + (\ell - m)\theta]} J_{m-\ell}(pa),$$

where $\ell, m \in \mathbb{Z}$, $g = (A, \mathbf{a}) = (\phi, \theta, a)$, $a = (a_1^2 + a_2^2)^{1/2}$, parameterized in polar coordinates and $J_\nu(x)$ is the νth order Bessel function. The matrix elements satisfy the orthogonality relation

$$\int_{\mathbb{SE}(2)} \overline{u_{\ell_1 m_1}(g, p_1)} u_{\ell m}(g, p)\mathrm{d}g = \frac{4\pi^2}{p}\delta_{\ell_1 \ell}\delta_{m_1 m}\delta(p_1 - p),$$

where $\ell, \ell_1, m, m_1 \in \mathbb{Z}$, $p, p_1 \in \mathbb{R}_+$ and δ is the Dirac or Kronecker delta function.

2.2. Fourier analysis on $\mathbb{SE}(2)$

Let $f \in L^1(\mathbb{SE}(2)) \cap L^2(\mathbb{SE}(2))$, the integrable and square integrable functions on $\mathbb{SE}(2)$. We define the Fourier transform on $\mathbb{SE}(2)$ with respect to an irreducible unitary representation as

$$\hat{f}(p) = \int_{\mathbb{SE}(2)} f(g)U(g^{-1}, p)\,\mathrm{d}g. \tag{2.7}$$

The matrix elements of the transform are given as

$$\hat{f}_{\ell m}(p) = \int_{\mathbb{SE}(2)} f(g)u_{\ell m}(g^{-1}, p)\,\mathrm{d}g,$$

where $\ell, m \in \mathbb{Z}$ and $p \in \mathbb{R}_+$.

Fourier inversion is defined by

$$f(g) = \int_0^\infty \mathrm{tr}\,(\hat{f}(p)U(g, p))p\,\mathrm{d}p, \tag{2.8}$$

where 'tr' stands for the trace of the object in question. Explicitly,

$$f(g) = \sum_{\ell=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\int_0^\infty \hat{f}_{\ell m}(p)u_{m\ell}(g, p)p\,\mathrm{d}p.$$

A useful property of Fourier transforms is that convolution of two functions in the Fourier domain turns out to be ordinary matrix multiplication. Indeed, let $f, k \in L^1(\mathbb{SE}(2)) \cap L^2(\mathbb{SE}(2))$. Define the convolution,

$$(k * f)(g) = \int_{\mathbb{SE}(2)} k(x)f(x^{-1}g)\,\mathrm{d}x, \tag{2.9}$$

for $g \in \mathbb{SE}(2)$. We have the following convolution property for $f, k \in L^1(\mathbb{SE}(2)) \cap L^2(\mathbb{SE}(2))$

$$\widehat{k * f} = \hat{f}\hat{k}. \tag{2.10}$$

In particular, for each $p \in \mathbb{R}_+$,

$$\left(\widehat{k * f}\right)_{\ell m}(p) = \sum_{q=-\infty}^{\infty}\hat{f}_{\ell q}(p)\hat{k}_{qm}(p),$$

where $\ell, m \in \mathbb{Z}$ and $p \in \mathbb{R}_+$.

3. Diffusions and Distributions on $\mathbb{SE}(2)$

We note that given the left invariant vector fields associated with the basis of the Lie algebra $\mathfrak{se}(2)$, any left invariant differential operator can be written as sums of products of these vector fields. A particularly important class is the second-order partial differential operators where the interest is due to the Markov (memoryless) property associated with them and is sometimes referred to as diffusions.

https://doi.org/10.1017/S0263574708005262 Published online by Cambridge University Press

In particular, based on the left invariant vector fields, Eq. (2.5) or (2.6), a second-order differential operator can be expressed in general as

$$\mathcal{D} = \sum_{i,j=1}^{3} \sigma_{ij} \tilde{\mathfrak{X}}_i \tilde{\mathfrak{X}}_j + \sum_{i=1}^{3} \tilde{\mathfrak{X}}_i, \tag{3.1}$$

where $\mu = (\mu_1, \mu_2, \mu_3)' \in \mathbb{R}^3$ and $\Sigma = (\sigma_{ij})$ is a positive semi-definite symmetric 3×3 matrix. For the purposes of this paper, we will assume that μ and Σ are fixed, although they do not have to be in general.

For some smooth $\omega : \mathbb{SE}(2) \times \mathbb{R}_+ \to \mathbb{R}$ the differential operator (3.1) induces the following partial differential equation or diffusion:

$$\frac{\partial}{\partial t}\omega = \mathcal{D}\omega, \tag{3.2}$$

where $\omega(x, 0) = f(x)$ is some initial condition $f(x)$, $x \in \mathbb{SE}(2)$. In the case where f is a Dirac delta function, the solution to Eq. (3.2) is called the fundamental solution. The above is also referred to as the Fokker–Planck or Kolmogorov equations.

Let $C^\infty(\mathbb{SE}(2))$ be the set of infinitely continuous differentiable functions on $\mathbb{SE}(2)$. For $f \in C^\infty(\mathbb{SE}(2))$, let

$$\widehat{\tilde{\mathfrak{X}}_j f}(p) = \widehat{\tilde{\mathfrak{X}}}_j(p) \hat{f}(p), \tag{3.3}$$

where

$$\begin{aligned} \widehat{\tilde{\mathfrak{X}}}_1(p) &= -\mathrm{i}\left(\ell \delta_{\ell,m}\right), \\ \widehat{\tilde{\mathfrak{X}}}_2(p) &= \mathrm{i}\frac{p}{2}\left(\delta_{\ell,m+1} + \delta_{\ell,m-1}\right), \\ \widehat{\tilde{\mathfrak{X}}}_3(p) &= \frac{p}{2}\left(\delta_{\ell,m+1} - \delta_{\ell,m-1}\right), \end{aligned} \tag{3.4}$$

where $\mathrm{i}^2 = -1$ and $p \in \mathbb{R}_+$.

Applying (3.3) and (3.4) to (3.1), we obtain

$$K(p) = \sum_{i,j=1}^{3} \sigma_{ij} \widehat{\tilde{\mathfrak{X}}}_i(p) \widehat{\tilde{\mathfrak{X}}}_j(p) + \sum_{i=1}^{3} \mu_i \widehat{\tilde{\mathfrak{X}}}_i(p), \tag{3.5}$$

for $p \in \mathbb{R}_+$. Thus applying the Fourier transform to (3.2) results in the ordinary differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\omega} = K(p)\hat{\omega},$$

which has as solution

$$\hat{\omega}(p, t) = \mathrm{e}^{tK(p)} \hat{f}(p),$$

for $t \in \mathbb{R}_+$. By Fourier inversion (2.8) the fundamental solution is

$$k_t(x) = \int_0^\infty \mathrm{tr}\left(\mathrm{e}^{tK(p)} U(x, p)\right) p \mathrm{d}p, \tag{3.6}$$

for $t \in \mathbb{R}_+$.

The above is a general description of the flow diagram illustrated in Fig. 3 of ref. [8].

3.1. *The Gaussian distribution on* $\mathbb{SE}(2)$

In the situation where $\mu = (0, 0, 0)'$ and $\Sigma = \mathbf{I}_3$, then Eq. (3.1) is of course the Laplacian on $\mathbb{SE}(2)$ which we denote by

$$\Delta = \tilde{\mathfrak{X}}_1^2 + \tilde{\mathfrak{X}}_2^2 + \tilde{\mathfrak{X}}_3^2.$$

Consequently, the corresponding diffusion is the heat equation on $\mathbb{SE}(2)$,

$$\frac{\partial}{\partial t}\omega = \Delta\omega,$$

where $\omega(x, 0) = f(x)$.

In this situation the corresponding Eq. (3.5) is represented as

$$G(p) = \sum_{j=1}^{3} \widehat{\tilde{\mathfrak{X}}}_j^2(p),$$

for $p \in \mathbb{R}_+$. Consequently,

$$\mathrm{e}^{tG(p)} = \left(\mathrm{e}^{-t(m^2+p^2)}\delta_{\ell m}\right),$$

so that the fundamental solution to Eq. (3.2) is

$$g_t(x) = \int_0^\infty \sum_{m=-\infty}^{\infty} \left(\mathrm{e}^{-t(m^2+p^2)} u_{mm}(x, p)\right) p \mathrm{d}p,$$

for $t \in \mathbb{R}_+$ and $x \in \mathbb{SE}(2)$ which we will call the Gaussian distribution on $\mathbb{SE}(2)$. It is worth pointing out that through a change of variables, we can reduce (3.1) to this case when Σ is of full rank.

3.2. *Degenerate distributions on* $\mathbb{SE}(2)$

In the situation where Σ is not of full rank, it is still possible to use (3.6). Again, by a change of variables, we can assume Σ is diagonal with diagonal entries $\sigma_{11}, \sigma_{22}, \sigma_{33} \geq 0$, and where $\sigma_{jj} = 0$ for some $j = 1, 2, 3$. This would result in a class of degenerate diffusions associated with a class of degenerate second-order differential operators otherwise known as hypoelliptic operators, a term first used by Hörmander[14] in a much more general setting than the current one.

Of particular interest is when σ_{22} or σ_{33} vanish, which has been extensively studied in the robotics literature, by Chirikjian *et. al.* (see refs. [8, 15, 16]). In addition to the robotics literature, the above author recognized that probability density functions in position and orientation associated with certain kinds of polymer chains can be described as a degenerate diffusion process on motion groups. This kind of process (and the resulting probability densities) has subsequently been observed in phenomena ranging from phase noise in optical communication

https://doi.org/10.1017/S0263574708005262 Published online by Cambridge University Press

systems[17], dead-reckoning error distributions in non-holonomic mobile robots[18] and the application closest to the current paper: the workspace densities of manipulators .[8]

For physical reasons, to be explained in Section 5.1, it is assumed that $\sigma_{11}, \sigma_{33} > 0$ while $\sigma_{22} = 0$, with $\mu \in \mathbb{R}^3$. With this in mind, define

$$C(p) = \sigma_{11}\widehat{\mathfrak{X}}_1^2(p) + \sigma_{33}\widehat{\mathfrak{X}}_3^2(p) + \sum_{i=1}^{3} \mu_i \widehat{\mathfrak{X}}_i(p),$$

for $p \in \mathbb{R}_+$. Define the resulting fundamental solution

$$c_t(x) = \int_0^\infty \mathrm{tr}\left(\mathrm{e}^{tC(p)}U(x, p)\right) p\mathrm{d}p, \tag{3.7}$$

for $t \in \mathbb{R}_+$.

For some vector space $\mathbb{V}$ with norm $\|\cdot\|$, and an operator $A : \mathbb{V} \to \mathbb{V}$, define the operator norm by

$$\|A\|_{\mathrm{op}}^2 = \sup_{\|v\|=1} \|Av\|^2. \tag{3.8}$$

In the case where $A = (a_{\ell m})_{\ell,m\in\mathbb{Z}}$, for some $\mathrm{T} > 0$ fixed, we will denote its $(2\mathrm{T}+1) \times (2\mathrm{T}+1)$ finite dimensional compression by $A_{\mathrm{T}} = (a_{\ell m})_{|\ell|,|m|\le T}$.

We have the following useful property whose proof is in Appendix 3. The main motivation of this result is in the bounding of the mean integrated squared error of the degenerate diffusion illustrated in Section 5 and is explicitly stated as Corollary 6.3.

Theorem 3.1. *For the degenerate distribution* (3.7)*,*

$$\mathrm{e}^{-\mathrm{t}\max\{\sigma_{11},\sigma_{33}\}(\mathrm{T}^2+3\mathrm{p}^2/2)} \le \|\mathrm{e}^{\mathrm{tC_T(p)}}\|_{\mathrm{op}}^2 \le \mathrm{e}^{-\mathrm{t}\min\{\sigma_{11},\sigma_{33}\}(\mathrm{T}^2+\mathrm{p}^2/2)},$$

where $p, t, T \in \mathbb{R}_+$.

4. Deconvolution

Let X, Y, Z be $\mathbb{SE}(2)$ random elements and suppose we observe Y according to

$$Y = ZX, \tag{4.1}$$

where Y and X are assumed independent. Let X, Y, Z have densities f, h, k, respectively. Through (4.1), the relation among the densities can be described by convolution, $h = k * f$, (2.9).

Now consider the Fourier transforms of each of the densities, where by (2.7), we have $\hat{f}(p)$, $\hat{h}(p)$ and $\hat{k}(p)$ for each $p \in \mathbb{R}_+$, respectively. By (2.10) we have,

$$\hat{h}(p) = \hat{f}(p)\hat{k}(p),$$

$p \in \mathbb{R}_+$. In coordinates, we have

$$\hat{h}_{\ell m}(p) = \sum_q \hat{f}_{\ell q}(p)\hat{k}_{qm}(p),$$

$p \in \mathbb{R}^+$, $\ell, m \in \mathbb{Z}$. Define $(\hat{k}(p))^{-1} = (\hat{k}_{qm}^{-1}(p))_{q,m\in\mathbb{Z}}$ as the (formal) inverse to $\hat{k}(p)$, $p \in \mathbb{R}^+$. Then

$$\hat{f}_{\ell m}(p) = \sum_q \hat{h}_{\ell q}(p)\hat{k}_{qm}^{-1}(p), \tag{4.2}$$

for $p \in \mathbb{R}^+$. The reason we have said formal inverse is because $\hat{k}(p)$ is often a compact (infinite-dimensional) operator hence the inverse would not exist. One would therefore have to regularize the problem through truncation as done later in the paper.

4.1. Statistical estimation

Statistically, (4.1) is describing the non-Euclidean analogue of observations Y made up of the measurement X we desire but cannot observe, multiplied by another random quantity Z. In the context of manipulator design, if we imagine that the manipulator is divided into two connected parts, then a density function $k(z)$ can be associated with those frames reachable by the end of the lower half of the manipulator, which is assumed known, and a density function $f(x)$ can be associated with the end of the upper half of the manipulator for which we do not know. Our interest is in the unknown f while it is assumed that k is known, or estimated by some prior experiment. Since f is unknown, h is also unknown, hence $\hat{h}$ is unknown. Nevertheless, we assume that a random sample $Y_1, \ldots, Y_n$ is available. This will allow us to construct an empirical version:

$$\hat{h}_{\ell m}^n(p) = \frac{1}{n}\sum_{j=1}^{n} u_{\ell m}\left(Y_j^{-1}, p\right), \tag{4.3}$$

for $|\ell|, |m| \le T$ and $p \le T$. By (4.2) an estimator for $\hat{f}$ is therefore

$$\hat{f}_{\ell m}^n(p) = \sum_{|q|\le T} \hat{h}_{\ell q}^n(p)\hat{k}_{T,qm}^{-1}(p), \tag{4.4}$$

where $k_{T,qm}^{-1}(p)$ are the elements of the inverse to the compression $k_T(p)$ for $p \in \mathbb{R}^+$, $|\ell|, |m| \le T$. We can then produce a non-parametric deconvolution density estimator of f by (2.8) and the Fourier inversion is

$$f^n(x) = \int_0^T \sum_{|\ell|,|m|\le T} \hat{f}_{\ell m}^n(p)u_{m\ell}(g, p)p\mathrm{d}p. \tag{4.5}$$

We note that the truncation $T > 0$ is there for the purpose of reducing the general infinite-dimensional problem down to a finite-dimensional one. The theoretical properties of (4.5) will be examined in Section 6. What we first want to establish are some practical issues with respect to the implementation.

5. Practical Implementation

In this section, we will put to use the statistical estimation of Section 4.1 into a practical robot design problem originally proposed in ref. [10] and informally discussed in Section 1.

https://doi.org/10.1017/S0263574708005262 Published online by Cambridge University Press

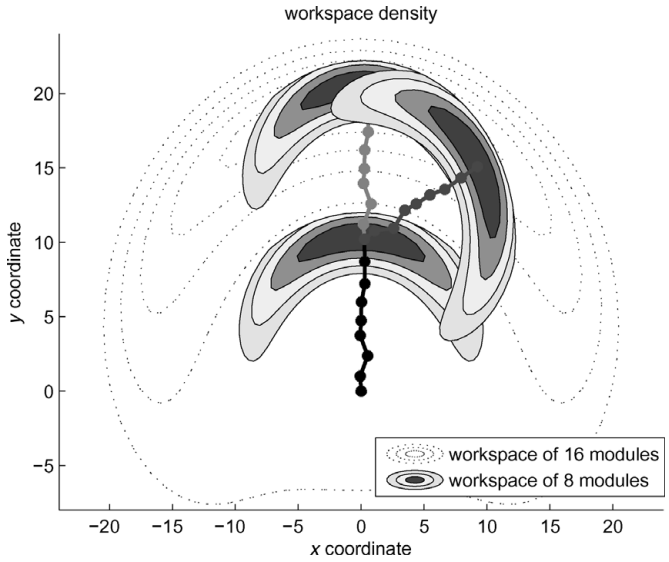

Fig. 1. Scatter plot of workspace density.

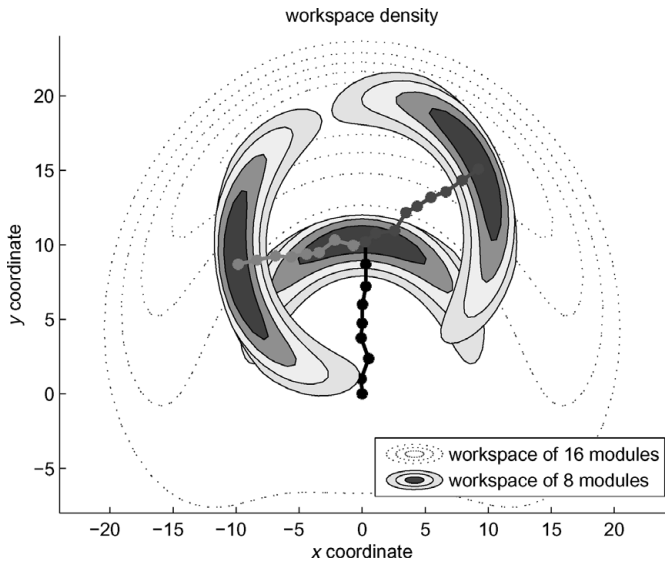

Fig. 2. Scatter plot of workspace density.

In general, suppose that one of the halves of the manipulator has been designed ($k(z)$ is specified and generates the Z data), and the problem is to design the other half of the manipulator (find $f(x)$ that generates the X data) so that the density function for the workspace of the whole manipulator comes as close as possible to a desired density function $h(y)$ that generates the Y data. One must then solve the inverse problem

$$f * k(y) = h(y),$$

for $f(x)$. Once $f(x)$ is known, the methods developed in ref. [4] can be used to find the appropriate kinematic parameters in the manipulator arm.

To be concrete, let us explicitly detail the design problem where we will make use of the following workspace density figures for illustration as has been established in ref. [8].

Both figures are illustrations of a 16-module robot design problem where 8 modules (black) have already been designed with the ensuing workspace probability density shown as coloured contours at the terminal module, and the blue-dotted lines are the contours of the workspace density established at the beginning of the robot design. In Fig. 1, if the remaining eight modules are to be constructed with no change in robot design, then a typical sample path would be the eight-module (red) path with the coloured contours at the end representing the probability density of the ensuing workspace.

It is however possible that after eight modules (black) have been constructed, a change in the original robot design needs to be implemented whereby a different region of the workspace needs emphasis. In Fig. 1, the eight-module (blue) path represents such a change where the ensuing workspace density is shown by the solid coloured contours at it's terminal module. This is further emphasized in Fig. 2 where after the eight module (black) manipulator design has been built, the remaining eight modules (both red and blue) are typical sample paths of the ensuing construction emphasizing different probability density weights in different regions of the workspace.

We note that the characteristics of the above represents the manner in which a robot design problem could actually be conceived of physically. Here, we detail the manner in which the kinematic parameters can be adjusted to achieve such outcomes.

5.1. Properties of degenerate distribution and sampling

For illustrative purposes and in line with ref. [8], let us write the degenerate diffusion as

$$\frac{\partial \omega}{\partial t} = \left(\beta \tilde{\mathfrak{X}}_1^2 + \epsilon \tilde{\mathfrak{X}}_3^2 + \alpha \tilde{\mathfrak{X}}_1 + \tilde{\mathfrak{X}}_3\right)\omega, \tag{5.1}$$

to describe the evolution of the workspace density function in $\mathbb{SE}(2)$. In particular, with respect to (3.7), $\sigma_{11} = \beta, \sigma_{22} = 0, \sigma_{33} = \epsilon$, $\mu_1 = \alpha$, $\mu_2 = 0$ and $\mu_3 = 1$. The practical reason for this specification is due to the fact that motors between segments either rotate or elongate in one direction. Consequently, variability occurs only in the angular component $\mathfrak{X}_1$ and one of the translational axis $\mathfrak{X}_3$.

The Fourier transform of the solution to (5.1) will be written as

$$\hat{c}_t(p|\alpha, \beta, \epsilon) = \exp\left\{t\left(\beta \widehat{\tilde{\mathfrak{X}}}_1^2(p) + \epsilon \widehat{\tilde{\mathfrak{X}}}_3^2(p) + \alpha \widehat{\tilde{\mathfrak{X}}}_1(p) + \widehat{\tilde{\mathfrak{X}}}_3(p)\right)\right\}, \tag{5.2}$$

for $p \in \mathbb{R}_+$ and the resulting density (3.7). Hence, fundamental solution to (5.1), will be written as

$$c_t(\phi, \theta, r|\alpha, \beta, \epsilon) = \int_0^\infty \mathrm{tr}\left(\hat{c}_t(p|\alpha, \beta, \epsilon) U(\phi, \theta, r, p)\right) p \mathrm{d}p, \tag{5.3}$$

$g = g(\phi, \theta, r) \in \mathbb{SE}(2)$ as in (2.2) with $\phi, \theta \in [0, 2\pi), r \geq 0$ and $t > 0$.

The degenerate distribution (5.3) is also written with respect to the parameters $\beta, \epsilon > 0$ and $\alpha \in \mathbb{R}$. Collectively, the parameters $(\alpha, \beta, \epsilon)$ determine the shape of the three-dimensional workspace density over the domain (ϕ, θ, r)

https://doi.org/10.1017/S0263574708005262 Published online by Cambridge University Press

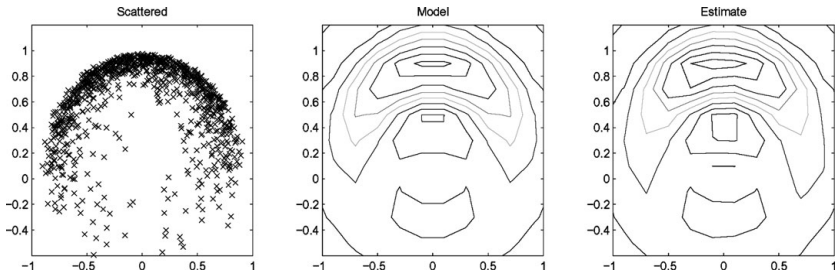

Fig. 3. Scatter plot $\alpha = 0$, $\beta = 1$, $\epsilon = 0$.

in general; however, only the two-dimensional positional workspace density over (θ, r) is needed to illustrate the effects of the parameter. Indeed, let

$$c_t(\theta, r|\alpha, \beta, \epsilon) = \int_0^{2\pi} c_t(\phi, \theta, r|\alpha, \beta, \epsilon)\, \mathrm{d}\phi \qquad (5.4)$$

be the marginal density with respect to the $\mathbb{SO}(2)$ element ϕ.

The effect of increasing $t > 0$ spreads the domain and the function out since it is a diffusion. The extensibility parameter ϵ extends or fattens the area of the workspace as it increases. The flexibility parameter β enlarges the support of the workplace density for larger values, and the symmetry parameter α introduces a shift towards one side depending on whether it is positive or negative. This is detailed in Fig. 5 along with explanations in ref. [8]. We would also like to recall Theorem 1 which provides the truncation bound of Eq. (5.4).

Applying the Euler–Maruyama method (see ref. [19] and Appendix A), one can generate samples from the degenerate distribution (5.3) in $\mathbb{SE}(2)$ given any parameters $(\alpha, \beta, \epsilon)$.

5.2. *Fitting kinematic parameters*

We will generate the sample $Z_1, \ldots, Z_n$ of k and the sample $X_1, \ldots, X_n$ of f, where both are from the degenerate distribution with varying parameters according to Eq. (5.3).

The sample $Y_1, Y_2, \ldots, Y_n$ in $\mathbb{SE}(2)$, defined by $Y = ZX$ will then be generated where the empirical Fourier transform of the convoluted workspace density can be obtained by (4.3). We will then define $\hat{f}^n_{\ell m}(p)$ according to (4.4), whereby we seek to find the optimal values of the parameters α, β and ϵ, by minimizing the cost function

$$\mathcal{C}(\alpha, \beta, \epsilon) = \int_0^P \|\hat{f}^n(p) - \hat{c}_t(p|\alpha, \beta, \epsilon)\|^2 p\, \mathrm{d}p \qquad (5.5)$$

over certain bandwidth (frequency parameters or smoothing parameter of the Fourier density estimation) $P > 0$, where $\hat{c}_t(p|\alpha, \beta, \epsilon)$ is defined in Eq. (5.2).

Minimization of Eq. (5.5) can be solved by taking a gradient descent approach. The basic gradient descent procedure is to compute the numerical derivatives

$$\frac{\partial \mathcal{C}}{\partial x_i} \approx \frac{\mathcal{C}(\mathbf{x} + \varepsilon \mathbf{e}_i) - \mathcal{C}(\mathbf{x})}{\varepsilon},$$

where $\mathbf{x} = (\alpha, \beta, \epsilon)'$. Since the gradient is the direction of steepest ascent, we step in the opposite direction by updating

$$\mathbf{x} \mapsto \mathbf{x} - \varepsilon \sum_{i=1}^{3} \frac{\partial \mathcal{C}}{\partial x_i} \mathbf{e}_i,$$

where $\mathbf{e}_i$, $i = 1, 2, 3$ is the standard basis for $\mathbb{R}^3$. Then we re-evaluate $\mathcal{C}$ at this new value of $\mathbf{x}$, recompute the gradient and iterate until $\mathcal{C}$ is below a specified threshold error.

5.3. *Implementation*

Figure 3 shows the scatter plot of the samples generated by this method with $\alpha = 0$, $\beta = 1$, $\epsilon = 0$ and $t = 1$. The left panel is the scatter plot of the samples ($n = 700$), the middle panel is the contour plot of the marginal density of the model and the right panel is the contour plot of the estimated marginal density based on the sample. This will serve as our density f.

Similarly, the k density is specified by the diffusion model with $\alpha = 0$, $\beta = 0.7$, $\epsilon = 0$ and $t = 0.5$. Figure 4 shows the marginal contour plots of the sample density (left), estimated diffusion density (middle) and the original diffusion density (right).

Here, we report on a numerical experiment consisting of a sample size of $n = 700$. The functions k and f are that of the degenerate distribution (5.2) with kinematic parameters $\alpha = 0$, $\beta = 0.7$, $\epsilon = 0$, $t = 0.5$ and $\alpha = 0$, $\beta = 1$, $\epsilon = 0$, $t = 1$, respectively.

Figure 5 shows the results of the numerical experiment. Starting with Fig. 5(a), the red dots are the scatter plot of samples drawn from $h = k * f$, while the blue dots are the scatter plot of samples drawn from f. Figure 5(b) is the estimated marginal density of f, while Fig. 5(c) is the marginal contour plot of the true f. Finally, Fig. 5(d) is the estimated kinematic density. We note that because of near singularity, a pseudo-inverse is used (see ref. [10] for details of regularization). As one can see, recovery of the true workspace density appears very satisfactory.

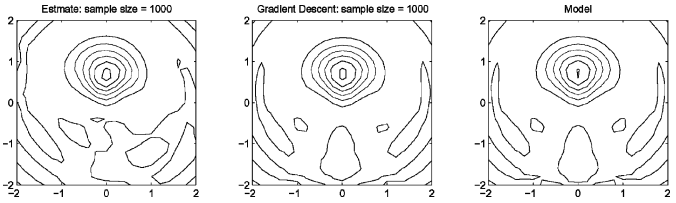

Fig. 4. Contour plot $\alpha = 0$, $\beta = 0.5$, $\epsilon = 0.1$.

https://doi.org/10.1017/S0263574708005262 Published online by Cambridge University Press

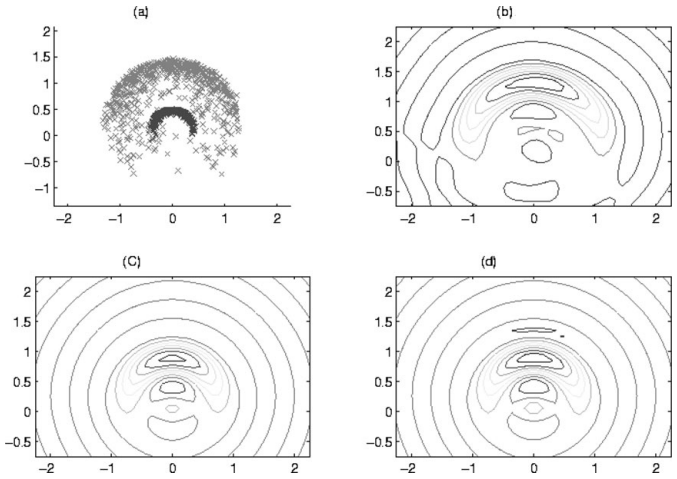

Fig. 5. Sampling deconvolution and kinematic parameter estimation.

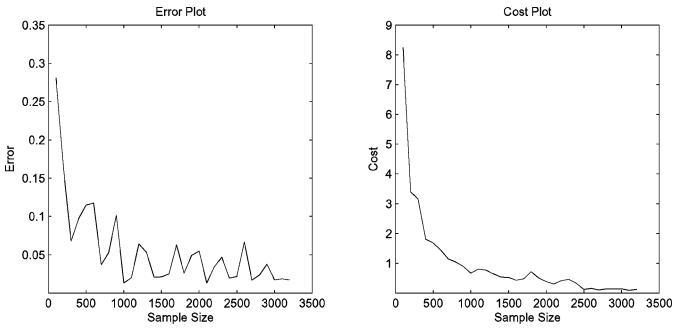

Fig. 6. Error plot $\alpha = 0$, $\beta = 0.5$, $\epsilon = 0.1$.

A number of numerical experiments were performed with equally satisfactory results and therefore, numerical evidence dictates that the deconvolution density estimator (4.5) provide stable and accurate estimates. By using large bandwidth and truncation size of the Fourier matrices, one can also achieve even greater accuracy but of course with higher computational cost. In the above experiment, we used a truncation of $T = 10$. We note, however, that the inversion of the truncated Fourier matrices may introduce larger errors and therefore the truncation effect requires further investigation.

Prior to the conclusion of this section, it is worth pointing out some further numerical features that arise in the implementation. Figure 6 shows the tendencies of the error (left panel) and the value of the cost function (right panel) to decrease with respect to the larger sample sizes which is to be expected. Here, the samples are generated using parameter set: $\{\alpha = 0,\ \beta = 0.5,\ \epsilon = 0.1\}$, the frequency parameter is $p = 0, \ldots, 10$ with step size of 0.2 and the Fourier matrices are truncated at size 10 (21×21 matrix). The error is normalized error and defined as

$$\text{error} = \frac{\|\mathbf{x}_{\text{model}} - \mathbf{x}_{\text{estimate}}\|}{\|\mathbf{x}_{\text{model}}\|}.$$

Another observation is the effect of the bandwidth on the accuracy of the method. As shown in Fig. 7 , the error tends to drop as the bandwidth increases, as expected also. We

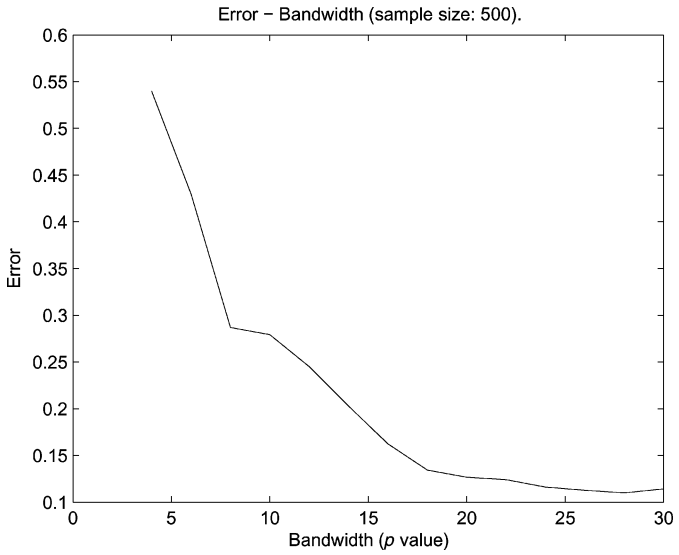

Fig. 7. Bandwidth effect on error.

note that using large bandwidth is particularly important for reducing errors in deconvolution due to the fact that inversion of a truncated Fourier matrix (estimated from the samples) will introduce more error. We will look at some theoretical properties in the next section which can quantify this affect.

6. Theoretical Properties

In this section, we will examine theoretical properties associated with (4.5). The following asymptotic notations will be used. Let $\{a_n\}$ and $\{b_n\}$ denote two real sequences of numbers. We write $a_n \ll b_n$ to mean $a_n \leq Cb_n$ for some $C > 0$, as $n \to \infty$, the Vinogradov notation. We will, however, use the notation $a_n = o(b_n)$ to mean $a_n/b_n \to 0$, as $n \to \infty$. Consequently, the expression, $o(1)$ would mean a sequence converging to 0.

First let us define a class of functions with the property

$$\Theta_s(\mathbb{SE}(2), Q) = \left\{ f : \int_0^\infty \sum_{\ell,m} (1 + \ell^2 + m^2 + p^2)^s \times \left| \hat{f}_{\ell m}(p) \right|^2 p \mathrm{d}p < Q^2 \right\}, \tag{6.1}$$

for $s > 3/2$, a Sobolev-type condition.

Second, let us assume k satisfies

$$\sup_{p \leq T} \left\| \hat{k}_T^{-1}(p) \right\|_{\text{op}}^2 \leq \exp\left\{ \frac{T^\rho}{\gamma} \right\}, \tag{6.2}$$

for some $\rho, \gamma > 0$ and third,

$$\sup_{p \leq T} \sum_{|m| \leq T} \sum_{q'=-\infty}^{\infty} \left| \delta_{q'm} - \sum_{|q| \leq T} \hat{k}_{q'q}(p) \hat{k}_{qm}^{-1}(p) \right|^2 = o(T^{-2s}), \tag{6.3}$$

as $T \to \infty$.

We have the following main result:

https://doi.org/10.1017/S0263574708005262 Published online by Cambridge University Press

Theorem 6.1. *Suppose k satisfies Eqs.* (6.2) *and* (6.3). *If* $f \in \Theta_s(\mathbb{SE}(2), Q)$ *for some* $s > 3/2$ *and*

$$T = \left(\gamma \ln n - \gamma \ln(\ln n)^{\kappa}\right)^{1/\rho},$$

where $\kappa > (2s+3)/\rho$, *then*

$$E\|f^n - f\|^2 \leq 9Q^2 (\gamma \ln n)^{-2s/\rho} (1 + o(1)),$$

as $n \to \infty$.

Condition (6.2) reflects the degree of ill posedness and in the case of the degenerate and Gaussian distributions, $\rho = 2$. Although Eq. (6.3) appears technical, one can see that in the case of diagonal operators, the condition follows. Hence, the Gaussian distribution has that property. In fact, with a little effort one can show that the degenerate distribution also satisfies Eq. (6.3). Therefore, we can apply this result to the cases of the Gaussian and degenerate distributions.

Corollary 6.2. *Suppose k is the Gaussian distribution. If* $f \in \Theta_s(\mathbb{SE}(2), Q)$ *for some* $s > 3/2$ *and*

$$T = \left(\frac{1}{t}\ln n - \frac{1}{t}\ln(\ln n)^{\kappa}\right)^{1/2},$$

where $\kappa > s + 3/2$, *then*

$$E\|f^n - f\|^2 \leq 9Q^2 \left(\frac{1}{t}\ln n\right)^{-s} (1 + o(1)),$$

as $n \to \infty$.

Recalling Theorem 3.1, we have the following result which is useful in describing the relationship of the workspace density to the truncation or threshold errors and kinematic parameters of Section 5.

Corollary 6.3. *Suppose k is the degenerate distribution. If* $f \in \Theta_s(\mathbb{SE}(2), Q)$ *for some* $s > 3/2$ *and*

$$T = \left(\frac{1}{2\max\{\sigma_{11}, \sigma_{33}\}t}\ln n - \frac{1}{2\max\{\sigma_{11}, \sigma_{33}\}t}\ln(\ln n)^{\kappa}\right)^{1/2},$$

where $\kappa > s + 3/2$, *then*

$$E\|f^n - f\|^2 \leq 9Q^2 \left(\frac{1}{2\max\{\sigma_{11}, \sigma_{33}\}t}\ln n\right)^{-s} (1 + o(1)),$$

as $n \to \infty$.

By choosing $Q = \ln n$, we can make everything data dependent. Indeed we have the following.

Corollary 6.4. *Suppose k is the Gaussian distribution. If* $f \in \Theta_s(\mathbb{SE}(2), \ln n)$ *for some* $s > 2$ *and*

$$T = \left(\frac{1}{t}\ln n - \frac{1}{t}\ln(\ln n)^{\kappa}\right)^{1/2},$$

where $\kappa > s + 3/2$, *then*

$$E\|f^n - f\|^2 \leq 9t^s (\ln n)^{-s+2} (1 + o(1)),$$

as $n \to \infty$.

Corollary 6.5. *Suppose k is the degenerate distribution. If* $f \in \Theta_s(\mathbb{SE}(2), \ln n)$ *for some* $s > 2$ *and*

$$T = \left(\frac{1}{2\max\{\sigma_{11}, \sigma_{33}\}t}\ln n - \frac{1}{2\max\{\sigma_{11}, \sigma_{33}\}t}\ln(\ln n)^{\kappa}\right)^{1/2},$$

where $\kappa > s + 3/2$, *then*

$$E\|f^n - f\|^2 \leq 9(2\max\{\sigma_{11}, \sigma_{33}\}t)^s (\ln n)^{-s+2} (1 + o(1)),$$

as $n \to \infty$.

Appendix A: Sampling the degenerate distribution

Consider the stochastic differential equation in $\mathbb{R}^3$:

$$\mathrm{d}X_t = \mathbf{a}\mathrm{d}t + B\mathrm{d}W_t, \tag{A1}$$

where $\mathbf{a}$ and B are constant drift and diffusion terms, X_t is a stochastic process in $\mathbb{R}^3$ and W_t is Brownian motion in $\mathbb{R}^3$. The corresponding Fokker–Planck equation is

$$\frac{\partial f}{\partial t} = -\sum_{i=1}^{3} a_i \frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{3} \frac{\partial^2 f}{\partial x_i \partial y_j} \left(BB'\right)_{ij}.$$

By a change of variables $\mathbf{x} \mapsto \xi$, where $\xi = \xi(\phi, \theta, r)$ as in (2.2), this leads to the diffusion

$$\frac{\partial f}{\partial t} = -\sum_{i=1}^{3} a_i \tilde{\mathfrak{X}}_i f + \sum_{i,j=1}^{3} \sigma_{ij} \tilde{\mathfrak{X}}_i \tilde{\mathfrak{X}}_j f,$$

where $\frac{1}{2}BB' = \Sigma$.

Thus, for a fixed $t \in \mathbb{R}_+$, define $J(\mathbf{x})\mathrm{d}\mathbf{x} = \mathbf{a}\mathrm{d}t + B\mathrm{d}\mathbf{x} = \mathbf{a}\mathrm{d}t + B\mathrm{d}W_t$, where J is the Jacobian of the transformation $\mathbf{x} \mapsto \xi$, as specified in Section III of ref. [8]. Consequently we can recover $\xi = J^{-1}\mathbf{x}$ by sampling from (A1).

Appendix B: Proof to Section 3

We will prove Theorem 3.1 for the (5.1) diffusion since it pertains directly to our numerical experiment. The more general case follows along the same lines.

Let $\mathfrak{X}_0 = (u, 0) \in \mathfrak{se}(2)$. Then $\exp(t\mathfrak{X}_0) = (tu, \mathbf{I}_2)$ and

$$\tilde{\mathfrak{X}}_0 f(x, A) = Au \cdot \nabla_x f(x, A), \tag{B1}$$

where $x, u \in \mathbb{R}^2$, $A \in \mathbb{SO}(2)$, $f : \mathbb{SE}(2) \to \mathbb{R}$ is differentiable, ∇_x is the gradient and '$\cdot$' is the dot product. For $f \in L^2(\mathbb{SE}(2))$, the Fourier transform $\hat{f}(p)$ induces a representation of $L^2(\mathbb{SO}(2))$ by

$$\hat{f}(p)v(U) = \int_{\mathbb{SE}(2)} f(\mathbf{a}, A)\mathrm{e}^{\mathrm{i}p\mathbf{a}\cdot\omega(U)} v(A^{-1}U)\mathrm{d}\mathbf{a}\mathrm{d}A, \tag{B2}$$

https://doi.org/10.1017/S0263574708005262 Published online by Cambridge University Press

where $\omega(U)' = X\mathbf{e}_2$, $\mathbf{e}_2 = (0,1)'$ for $U \in \mathbb{SO}(2)$ and $v \in L^2(\mathbb{SO}(2))$. Consequently,

$$\widehat{\tilde{\mathfrak{X}}_0 f}(p)v(U) = -\mathrm{i}p\hat{f}(p)u \cdot \omega(U)v(U), \tag{B3}$$

and so $C(p)$ of Eq. (5.1) can be written in the equivalent form

$$C(p) = \beta\widehat{\tilde{\mathfrak{X}}}_1^2(p) - \epsilon p^2 V^2 + \alpha\widehat{\tilde{\mathfrak{X}}}_1 - \mathrm{i}pV,$$

where $V(U) = \mathbf{e}_2' U \mathbf{e}_2$. Also define the operator

$$L(p) = -\beta\widehat{\tilde{\mathfrak{X}}}_1^2(p) + \epsilon p^2 V^2.$$

We note that both C and L are operators on the Hilbert space $L^2(\mathbb{SO}(2))$, where $L^2(\mathbb{SO}(2)) = \mathrm{span}\{\mathrm{e}^{in\phi}, 0 \le \phi < 2\pi, n \in \mathbb{Z}\}$. Let $V_T = \mathrm{span}\{\mathrm{e}^{in\phi}, 0 \le \phi < 2\pi, -T \le n \le T\} \subset L^2(\mathbb{SO}(2))$ and denote C_T and L_T as the restriction of C and L on V_T, respectively.

By the Trotter product formula, for any $v \in V_T$,

$$\mathrm{e}^{tC_T}v = \lim_{n\to\infty}\left(\mathrm{e}^{(t/n)\beta\widehat{\tilde{\mathfrak{X}}}_1^2 - (t/n)\epsilon p^2V^2}\mathrm{e}^{(t/n)\alpha\widehat{\tilde{\mathfrak{X}}}_1}\mathrm{e}^{-(t/n)\mathrm{i}pV}\right)^n v.$$

As the factors $\mathrm{e}^{(t/n)\alpha\widehat{\tilde{\mathfrak{X}}}_1}$ and $\mathrm{e}^{-(t/n)\mathrm{i}pV}$ are unitary, hence $\|e^{tC_T}v\|_{\mathrm{op}}$ is bounded by

$$\left\|\mathrm{e}^{(t/n)\beta\widehat{\tilde{\mathfrak{X}}}_1^2 - (t/n)\epsilon p^2V^2}\right\|_{\mathrm{op}}^n \|v\|.$$

Since $\beta\widehat{\tilde{\mathfrak{X}}}_1^2 - \epsilon p^2V^2$ is a negative self-adjoint operator, we have

$$\left\|\mathrm{e}^{(t/n)\beta\widehat{\tilde{\mathfrak{X}}}_1^2 - (t/n)\epsilon p^2V^2}\right\|_{\mathrm{op}}^n = \left\|\mathrm{e}^{t\beta\widehat{\tilde{\mathfrak{X}}}_1^2 - t\epsilon p^2V^2}\right\|_{\mathrm{op}}. \tag{B4}$$

Suppose $a = \min\{\beta,\epsilon\}, b = \max\{\beta,\epsilon\}$. Let $f(\phi) = \sum_{n=-T}^{n=T} a_n\mathrm{e}^{in\phi} \in V_T$ and $\|f(\phi)\| = 1$, i.e. $\sum_{n=-T}^{n=T} a_n^2 = 1$. Then it can be shown that

$$\|L_T f(\phi)\|^2 \le b^2(T^2 + 3p^2/2)^2.$$

Since $\mathrm{e}^{in\phi}$ is a unit vector

$$L_T\mathrm{e}^{in\phi} = (\beta n^2 + \epsilon p^2\sin^2\phi)\mathrm{e}^{in\phi},$$

hence,

$$\begin{aligned}
\|L_T\mathrm{e}^{in\phi}\|^2 &= \frac{1}{2\pi}\int_0^{2\pi}(\beta n^2 + \epsilon p^2\sin^2\phi)^2\mathrm{e}^{in\phi}\mathrm{e}^{-in\phi}\mathrm{d}\phi \\
&= \frac{1}{2\pi}\int_0^{2\pi}(\beta^2 n^4 + 2\beta n^2\epsilon p^2\sin^2\phi \\
&\quad + \epsilon^2 p^4\sin^4\phi)\mathrm{d}\phi \\
&= \beta^2 n^4 + 2\beta n^2\epsilon p^2/2 + \frac{3}{8}\epsilon^2 p^4 \\
&> (\beta n^2 + \epsilon p^2/2)^2 \\
&\ge a^2(n^2 + p^2/2)^2.
\end{aligned}$$

So,

$$\|L_T\|_{\mathrm{op}} > a(T^2 + p^2/2).$$

Consequently,

$$a(T^2 + p^2/2) \le \|L_T\|_{\mathrm{op}} \le b(T^2 + 3p^2/2).$$

Since

$$\|\mathrm{e}^{tC_T}\|_{\mathrm{op}} = \|\mathrm{e}^{-tL_T}\|_{\mathrm{op}},$$

hence,

$$\mathrm{e}^{-tb(T^2+3p^2/2)} \le \|\mathrm{e}^{tC_T}\|_{\mathrm{op}} \le \mathrm{e}^{-ta(T^2+p^2/2)},$$

as required.

Appendix C: Proofs to Section 6

Let us first calculate the bias. We note that

$$\begin{aligned}
f(g) - \mathbb{E}f^n(g) &= \int_0^\infty \sum_\ell \sum_{|m|>T} \hat{f}_{\ell m}(p)u_{m\ell}(g,p)p\mathrm{d}p \\
&\quad + \int_0^\infty \sum_{|\ell|>T}\sum_{|m|\le T} \hat{f}_{\ell m}(p)u_{m\ell}(g,p)p\mathrm{d}p \\
&\quad + \int_T^\infty \sum_{|\ell|\le T}\sum_{|m|\le T} \hat{f}_{\ell m}(p)u_{m\ell}(g,p)p\mathrm{d}p \\
&\quad + \int_0^T \sum_{|\ell|\le T}\sum_{|m|\le T} \left(\hat{f}_{\ell m}(p) - \mathbb{E}\hat{f}_{\ell m}^n\right) \\
&\quad \times u_{m\ell}(g,p)p\mathrm{d}p.
\end{aligned} \tag{C1}$$

Applying Plancherel formula to (C1), we get

$$\begin{aligned}
\|f - \mathbb{E}f^n\| &\le \left\{\int_0^\infty \sum_\ell \sum_{|m|>T} |\hat{f}_{\ell m}(p)|^2 p\mathrm{d}p\right\}^{1/2} \\
&\quad + \left\{\int_0^\infty \sum_{|\ell|>T}\sum_{|m|\le T} |\hat{f}_{\ell m}(p)|^2 p\mathrm{d}p\right\}^{1/2} \\
&\quad + \left\{\int_T^\infty \sum_{|\ell|\le T}\sum_{|m|\le T} |\hat{f}_{\ell m}(p)|^2 p\mathrm{d}p\right\}^{1/2} \\
&\quad + \left\{\int_0^T \sum_{|\ell|\le T}\sum_{|m|\le T} |\hat{f}_{\ell m}(p) \right. \\
&\quad \left. - \mathbb{E}\hat{f}_{\ell m}^n(p)|^2 p\mathrm{d}p\right\}^{1/2}.
\end{aligned} \tag{C2}$$

Now

$$\int_0^\infty \sum_\ell \sum_{|m|>T} |\hat{f}_{\ell m}(p)|^2 p\mathrm{d}p \le Q^2 T^{-2s}, \tag{C3}$$

https://doi.org/10.1017/S0263574708005262 Published online by Cambridge University Press

$$\int_0^\infty \sum_{|\ell|>T} \sum_{|m|\le T} |\hat{f}_{\ell m}(p)|^2 p\mathrm{d}p \le Q^2 T^{-2s}, \tag{C4}$$

and

$$\int_T^\infty \sum_{|\ell|\le T} \sum_{|m|\le T} |\hat{f}_{\ell m}(p)|^2 p\mathrm{d}p \le Q^2 T^{-2s}, \tag{C5}$$

consequently, by (C3), (C4) and (C5), we can bound the first three terms following the inequality in (C2).

As for the last term we have the following:

Lemma C.1. Suppose

$$\sup_{p\le T} \sum_{|m|\le T} \sum_{q'=-\infty}^{\infty} \left| \delta_{q'm} - \sum_{|q|\le T} \hat{k}_{q'q}(p)\hat{k}^{-1}_{qm}(p) \right|^2 = o(T^{-2s}),$$

then

$$\int_0^T \sum_{|\ell|\le T} \sum_{|m|\le T} \left| \hat{f}_{\ell m}(p) - \mathbb{E}\hat{f}^n_{\ell m}(p) \right|^2 pdp = o(T^{-2s})$$

as $T \to \infty$.

Proof. We have

$$\sum_{|\ell|,|m|\le T} \left| \hat{f}_{\ell m}(p) - \mathbb{E}\hat{f}^n_{\ell m}(p) \right|^2 \tag{C6}$$

$$= \sum_{|\ell|,|m|\le T} \left| \hat{f}_{\ell m}(p) - \sum_{|q|\le T} \mathbb{E}\hat{h}^n_{\ell q}(p)\hat{k}^{-1}_{qm}(p) \right|^2 \tag{C7}$$

$$= \sum_{|\ell|,|m|\le T} \left| \hat{f}_{\ell m}(p) - \sum_{|q|\le T} \hat{h}_{\ell q}(p)\hat{k}^{-1}_{qm}(p) \right|^2 \tag{C8}$$

$$= \sum_{|\ell|,|m|\le T} \left| \hat{f}_{\ell m}(p) - \sum_{|q|\le T} \sum_{q'=-\infty}^{\infty} \hat{f}_{\ell q'}(p) \times \hat{k}_{q'q}(p)\hat{k}^{-1}_{qm}(p) \right|^2 \tag{C9}$$

$$= \sum_{|\ell|,|m|\le T} \left| \hat{f}_{\ell m}(p) - \sum_{q'=-\infty}^{\infty} \hat{f}_{\ell q'}(p) \sum_{|q|\le T} \hat{k}_{q'q} \times (p)\hat{k}^{-1}_{qm}(p) \right|^2 \tag{C10}$$

$$= \sum_{|\ell|,|m|\le T} \left| \sum_{q'=-\infty}^{\infty} \hat{f}_{\ell q'}(p) \left(\delta_{q'm} - \sum_{|q|\le T} \hat{k}_{q'q} \times (p)\hat{k}^{-1}_{qm}(p) \right) \right|^2 \tag{C11}$$

$$\le \sum_{|\ell|,|m|\le T} \sum_{q'=-\infty}^{\infty} \left| \hat{f}_{\ell q'}(p) \right|^2 \sum_{q'=-\infty}^{\infty} |\delta_{q'm} - \sum_{|q|\le T} \hat{k}_{q'q}(p)\hat{k}^{-1}_{qm}(p)|^2 \tag{C12}$$

$$= \sum_{|\ell|\le T} \sum_{q'=-\infty}^{\infty} \left| \hat{f}_{\ell q'}(p) \right|^2 \sum_{|m|\le T} \sum_{q'=-\infty}^{\infty} \left| \delta_{q'm} - \sum_{|q|\le T} \hat{k}_{q'q}(p)\hat{k}^{-1}_{qm}(p) \right|^2. \tag{C13}$$

Therefore, we have

$$\begin{aligned}
&\int_0^T \sum_{|\ell|,|m|\le T} \left| \hat{f}_{\ell m}(p) - \mathbb{E}\hat{f}^n_{\ell m}(p) \right|^2 p\mathrm{d}p \\
&\le \sup_{p\le T} \sum_{|m|\le T} \sum_{q'=-\infty}^{\infty} \left| \delta_{q'm} - \sum_{|q|\le T} \hat{k}_{q'q}(p)\hat{k}^{-1}_{qm}(p) \right|^2 \\
&\quad \times \int_0^T \sum_{|\ell|,|m|\le T} |\hat{f}_{\ell m}(p)|^2 p\mathrm{d}p \\
&= \sup_{p\le T} \sum_{|m|\le T} \sum_{q'=-\infty}^{\infty} \left| \delta_{q'm} - \sum_{|q|\le T} \hat{k}_{q'q}(p)\hat{k}^{-1}_{qm}(p) \right|^2 \\
&\quad \times \int_{\mathbb{SE}(2)} |f(x)|^2 \mathrm{d}x \\
&= o(T^{-2s}),
\end{aligned}$$

as $T \to \infty$. □

Consequently, by (C2) and Lemma C1, we obtain the inequality

$$\left\| \mathbb{E}f^n - f \right\|^2 \ll 9Q^2 T^{-2s}(1 + o(1)), \tag{C14}$$

as $T \to \infty$ for $s > 3/2$.

To compute the variance, we note that

$$\begin{aligned}
\| f^n - \mathbb{E}f^n \|^2 &= \int_0^T \sum_{|\ell|,|m|\le T} \left| \hat{f}^n_{\ell m}(p) - \mathbb{E}\hat{f}^n_{\ell m}(p) \right|^2 p\mathrm{d}p \\
&= \int_0^T \sum_{|\ell|,|m|\le T} \left| \sum_{|q|\le T} \left(\hat{f}^n_{\ell q}(p) - \mathbb{E}\hat{f}^n_{\ell,q}(p) \right) \times \hat{k}^{-1}_{qm}(p) \right|^2 p\mathrm{d}p \\
&\le \sup_{p\le T} \left\| \hat{k}^{-1}_T(p) \right\|^2_{\mathrm{op}} \int_0^T \sum_{|\ell|,|m|\le T} \left| \hat{h}^n_{\ell m}(p) - \mathbb{E}\hat{h}^n_{\ell m}(p) \right|^2 p\mathrm{d}p.
\end{aligned} \tag{C15}$$

Now

$$\begin{aligned}
&\sum_{|\ell|,|m|\le T} \mathbb{E} \left| \hat{h}^n_{\ell m}(p) - \mathbb{E}\hat{h}^n_{\ell m}(p) \right|^2 \\
&= \frac{1}{n} \left\{ \sum_{|\ell|,|m|\le T} \mathbb{E} \left| u_\ell(Y^{-1}, p) \right|^2 - \sum_{|\ell|,|m|\le T} \left| \hat{h}_{\ell m}(p) \right|^2 \right\}
\end{aligned}$$

https://doi.org/10.1017/S0263574708005262 Published online by Cambridge University Press

$$\leq \frac{1}{n}\sum_{|\ell|\leq T}\mathbb{E}\sum_{m=-\infty}^{\infty}\left|u_{\ell m}(Y^{-1},p)\right|^2$$
$$= \frac{1}{n}\sum_{|\ell|\leq T}\mathbb{E}\sum_{m=-\infty}^{\infty}u_{\ell m}(Y^{-1},p)u_{m\ell}(Y,p)$$
$$= \frac{1}{n}\sum_{|\ell|\leq T}\mathbb{E}u_{\ell\ell}(e,p)$$
$$= \frac{T}{n}, \tag{C16}$$

here e denotes the unit element of $\mathbb{SE}(2)$. Here we used the homomorphism property $U(g,p)U(g^{-1},p) = U(gg^{-1},p)$ along with the fact that $U(g,p)^* = U(g^{-1},p)$.

By taking expectation of (C15) and applying (C16), we obtain,

$$\mathbb{E}\|f^n - \mathbb{E}f^n\|^2 \leq \sup_{p\leq T}\left\|\hat{k}_{\mathrm{T}}^{-1}(p)\right\|_{\mathrm{op}}^2\frac{T^3}{n}, \tag{C17}$$

as $n\to\infty$.

Now,

$$\mathbb{E}\|f^n - f\|^2 = \mathbb{E}\|f^n - \mathbb{E}f^n\|^2 + \|\mathbb{E}f^n - f\|^2$$
$$\leq \sup_{p\leq T}\left\|\hat{k}_T^{-1}(p)\right\|_{\mathrm{op}}^2\frac{T^3}{n} + 9Q^2T^{-2s}(1+o(1))$$
$$\leq \exp\left\{\frac{T^\rho}{\gamma}\right\}\frac{T^3}{n} + 9Q^2T^{-2s}(1+o(1)). \tag{C18}$$

Now set

$$T = (\gamma\ln n - \gamma\ln(\ln n)^\kappa)^{1/\rho},$$

where $\kappa > (2s+3)/\rho$ and the result follows.

Acknowledgements

The authors would like to thank the editor, an associate editor and the three anonymous referees, for very insightful comments which greatly improved the quality of this paper. Also, the research is supported in part by the Natural Sciences and Engineering Research Council of Canada, Grant DG 46204.

Reference

1. D. L. Pieper, The Kinematics of Manipulators under Computer Control *Ph.D. Dissertation* (Stanford, CA: Stanford University, Oct. 1968).
2. B. Roth, J. Rastegar and V. Scheinman, "On the Design of Computer Controlled Manipulators," *First CISM-IFTMM Symposium on Theory and Practice of Robots and Manipulators* (1973), pp. 93–113.
3. A. Koliskor, "The l-Coordinate Approach to the Industrial Robots Design," *Information Control Problems in Manufacturing Technology 1986. Proceedings of the 5th IFAC/IFIP/IMACS/IFORS Conference*, Suzdal, USSR, (1987) pp. 225–232 (Preprint).
4. G. S. Chirikjian, "Kinematic synthesis of mechanisms and robotic manipulators with binary actuators," *ASME J. Mech. Des.* **117**, 573–580 (Dec. 1995).
5. D. Sen and T. S. Mruthyunjaya, "A discrete state perspective of manipulator workspaces," *Mech. Mach. Theory* **29**(4), 591–605 (1994).
6. U. Basavaraj and J. Duffy, "End-effector motion capabilities of serial manipulators," *Int. J. Rob. Res.* **12**(2), 132–145 (Apr. 1993).
7. J. U. Korein, *A Geometric Investigation of Reach* (MIT Press, Cambridge, MA, 1985).
8. Y. F. Wang and G. S. Chirikjian, "Workspace generation of hyper-redundant manipulators as a diffusion process on $SE(N)$," *IEEE Trans. Rob. Autom.* **20**(3), 399–408 (Jun. 2004).
9. Y. Liu, Probability Density Estimation on Rotation and Motion Groups *Ph.D. Dissertation* (Baltimore, MD: Department of Mechanical Engineering, Johns Hopkins University, Apr. 2007).
10. G. S. Chirikjian, "Fredholm integral equations on the Euclidean motion group," *Inverse Probl.* **12**, 579–599 (Oct. 1996).
11. G. S. Chirikjian and I. Ebert-Uphoff, "Numerical convolution on the Euclidean group with applications to workspace generation," *IEEE Trans. Rob. Autom.* **14**(1), 123–136 (Feb. 1998).
12. I. Ebert-Uphoff and G. S. Chirikjian, "Efficient workspace generation for binary manipulators with many actuators," *J. Rob. Syst.* **12**(6), 383–400 (Jun. 1995).
13. I. Ebert-Uphoff and G. S. Chirikjian, "Inverse Kinematics of Discretely Actuated Hyper-Redundant Manipulators Using Workspace Densities," *Proceedings of the IEEE International Conference on Robotics and Automation ICRA*, Minneapolis, MN (Apr. 1996) pp. 139–145.
14. L. Hörmander, "Hypoelliptic second order differential equations," *Acta Math.* **119**, 147–171 (1967).
15. G. S. Chirikjian, and A. B. Kyatkin, *Engineering Applications of Noncommutative Harmonic Analysis* (CRC Press, Boca Raton, FL, 2001).
16. G. S. Chirikjian and A. B. Kyatkin, "An operational calculus for the Euclidean motion group with applications in robotics and polymer science," *J. Fourier Anal. Appl.* **6**(6), 583–606 (Dec. 2000).
17. Y. Wang, Y. Zhou, D. K. Maslen and G. S. Chirikjian, "Solving the phase-noise Fokker–Planck equation using the motion-group Fourier transform," *IEEE Trans. Commun.* **54**(5), 868–877 (May 2006).
18. Y. Zhou and G. S. Chirikjian, "Probabilistic Models of Dead-Reckoning Error in Nonholonomic Mobile Robots," *Proceedings of the IEEE International Conference on Robotics and Automation ICRA*, Taipei, Taiwan (Sep. 2003).
19. D. J. Higham, "An algorithmic introduction to numerical simulation of stochastic differential equations," *SIAM Rev.* **43**(3), 525–546 (2001).