Stochastic Optimal Motion Planning and Estimation for the Attitude Kinematics on SO(3)

Taeyoung Lee

Abstract—This paper investigates stochastic motion planning and estimation for the attitude kinematics of a rigid body. A Fokker-Planck equation on the special orthogonal group is numerically solved via noncommutative harmonic analysis to obtain computational tools to propagate a probability density function over the attitude kinematics. Based on this, a stochastic optimal control problem is formulated for motion planning, and a Bayesian framework is applied for estimation. The proposed intrinsic, geometric formulation does not require the common assumption that uncertainties are Gaussian or localized. It can be also applied to complex rotational maneuvers of a rigid body without singularities in a unified way. The desirable properties are illustrated by numerical examples.

I. INTRODUCTION

Motion planning is formulated as generating curves in a configuration manifold satisfying given boundary conditions. Lie group is a differential manifold that has group structures, and the configuration space of complex interconnected multi-body systems can be represented by a Lie group [1]. Motion planning for deterministic systems on a nonlinear manifold has been studied with interpolation-based techniques [2], [3] and variational methods [4]. Stochastic motion planning for simple kinematic systems on a manifold, such as a planar cart is considered in [5].

Estimation involves determining the configuration of a dynamic system based on measurements and prior knowledge of its configuration. For attitude estimation problems, the extended Kalman filter has been the workhorse. In particular, extended Kalman filter developed in terms of quaternions has been extremely popular (see, for example [6], [7]).

Quaternions do not exhibit singularities unlike other minimal attitude representations, but they suffer from ambiguity as a single attitude is represented by two antipodal quaternions. It has been shown that attitude control systems based on quaternions may be sensitive to measurement errors or they may exhibit unwinding where a rigid body unnecessarily rotates along the opposite direction through a large angle even if the initial attitude error is small [8]. Recently, attitude estimation on the special orthogonal group SO(3) is addressed in [9], and a nonlinear attitude observer for deterministic attitude kinematics is developed in [10].

For successful implementation of stochastic motion planning and estimation, characterizing the evolution of uncertainties accurately is critical. Most of the existing approaches are based on the assumptions that uncertainties follow Gaussian distribution, and variances are sufficiently small such that they can be propagated along the linearized equations of motion. However, the Gaussian assumption greatly reduces the amount of information that is contained in the true density, and the linearization is valid only in small neighborhoods of the reference trajectory. These put fundamental limitation on the performance of stochastic motion planning and estimation, especially for complex maneuvers with large, non-Gaussian uncertainties. There are many applications of the extended Kalman filter where unsatisfactory results have been reported [11].

In this paper, a Fokker-Planck equation on a Lie group is presented, and it is solved according to noncommutative harmonic analysis [12]. This provides a computational framework to propagate non-Gaussian probability density functions in a coordinate-free fashion over attitude flows. Based on this, stochastic optimization problems are formulated for motion planning, and a Bayesian estimator is constructed for the attitude kinematics equation. The unique properties of the proposed approaches are as follows: (i) non-Gaussian probability distributions with possibly large variances are incorporated for motion planning and estimation of attitude; (ii) they are developed in a coordinate-free fashion to avoid singularities, complexities, and ambiguities that are associated with local parameterizations or quaternions; (iii) the corresponding attitude uncertainties are visualized. Stochastic motion planning and estimation with large uncertainties have not been addressed for attitude flows on SO(3), and while this paper is focused on the attitude kinematics, the proposed approaches are readily applied for the attitude dynamics of a rigid body, or multi-body systems.

II. UNCERTAINTY PROPAGATION ON A LIE GROUP

A. Fokker-Planck Equation on a Lie Group

Consider an n-dimensional Lie group G and its Lie algebra g. Let $L_g : G \rightarrow G$ be the left translation map defined as $L_g h = gh$ for $g, h \in G$. Its tangential map is denoted by $T_h L_g : T_h G \rightarrow T_{gh} G$. A deterministic dynamic system evolving on G can be described as

$$\dot{g} = T_c L_g f(g, t) \triangleq g f(g, t),$$

(1)

where $f : G \times \mathbb{R}_+ \rightarrow g$ describe the left-trivialized vector field [13]. Consider a stochastic system that is perturbed by a Wiener process. It is described by the following stochastic differential equation:

$$g^{-1} dg = f(g, t) dt + H(t) dW,$$

(2)
where $W \in \mathbb{R}^m$ denotes a vector of Wiener process with a matrix $H(t) \in \mathbb{R}^{n \times m}$ that scales and couples noises, and it is assumed that the Lie algebra $\mathfrak{g}$ is identified with $\mathbb{R}^n$. It is defined according to the Ito stochastic calculus [14].

Let $p(g, t) : G \times \mathbb{R}^n \rightarrow \mathbb{R}_+$ be a probability density function that describe the uncertainty distribution of the stochastic system at $t$. The evolution of a probability density function along the solution of (2) is described by the Fokker-Planck equation [14]. Here, we introduce the left-trivialized derivative of a function to define a Fokker-Planck equation on $G$. An infinitesimal variation of $g \in G$ can be written as

$$\delta g = \left. \frac{d}{de} \right|_{e=0} g \exp \epsilon \eta = g \eta \in T_g G,$$

for $\eta \in \mathfrak{g}$. Using this, the derivative of $f(g, t)$ along the direction of $\delta g$ is given by

$$\mathbf{D}_g f(g, t) \cdot (\delta g) = (g^{-1}\mathbf{D} f(g, t)) \cdot (g^{-1}\delta g) \triangleq \delta f(g, t) \cdot \eta,$$

where $\mathbf{D}_g$ denotes the left-trivialized derivative with respect to $g$, and $\delta$ denotes the left-trivialized derivative that can be formally written as $\delta = T^\ast L_g$. Note that $\mathbf{D} f(g, t) \in \mathfrak{g}^*$. Let $\{e_i\}_{1 \leq i \leq n}$ be a basis of set for the Lie algebra $\mathfrak{g}$. The left-trivialized derivative along of $e_i$ as $\mathbf{D}_g$:

$$\mathbf{D}_g f(g, t) \cdot (e_i) = \partial_i f(g, t) \cdot e_i. \quad (3)$$

Using this, the Fokker-Planck equation that describes the evolution of a probability density function along (2) is

$$\frac{\partial p(g, t)}{\partial t} + \sum_{i=1}^{n} \partial_i (f_i(g, t)p(g, t)) - \frac{1}{2} \sum_{i,j=1}^{n} \partial_i \partial_j (p(g, t)H_{ik}(t)H_{kj}(t)) = 0, \quad (4)$$

where the second term represents the effects of advection due to the vector field, and the last term corresponds to diffusion due to noise [15].

### B. Noncommutative Harmonic Analysis

Non-commutative harmonic analysis is a generalization of Fourier analysis on $\mathbb{R}^n$ to nonlinear manifolds [12], [16]. This is particularly useful since any square-integrable ($L_2$) function on a Lie group can be approximated by its Fourier spectrum, up to an arbitrarily prescribed level of accuracy.

Here, probability density functions are transformed via noncommutative harmonic analysis. The Fourier transform $\mathcal{F}[]$ of $p(g, t)$ at each $t$ is given by

$$\mathcal{F}_t[p(g, t)] \triangleq \bar{p}_t(t) = \int_G p(g, t)U_l(g^{-1})dg, \quad (5)$$

where $\bar{p}_t(t)$ is the Fourier transform indexed by $l$, and $U_l$ is irreducible representation of the Lie group, and $l$ is Fourier parameter. Both of $\bar{p}_t$ and $U_l$ are complex-valued matrices, and their size is determined by the index $l$ for a given $G$. The inverse transform is given by

$$p(g, t) = \sum_l \lambda_l \text{tr} [\bar{p}_t(t)U_l(g)], \quad (6)$$

where $\lambda_l$ is a scalar determined by $l$ for a given Lie group. Using these, any arbitrary probability density function can be represented in terms of its Fourier transform $\bar{p}_t(t)$.

Another useful property of noncommutative harmonic analysis is that the Fourier transform of the derivatives of $p(g, t)$ is also represented by $\bar{p}_t(t)$. From (3), we have

$$\mathcal{F}_t[\delta_i p(g, t)] = \int_G \left. \frac{d}{de} \right|_{e=0} p(g \exp \epsilon e_i, t)U_l(g^{-1})dg, \quad (7)$$

Let $h = g \exp \epsilon e_i$. Then, we have $dh|_{e=0} = dg d(\exp \epsilon e_i)|_{e=0} = dg$. Also, using the group property of the irreducible representation given by $U(gh) = U(g)U(h)$ for any $g, h \in G$,

$$U_l(g^{-1}) = U_l(\exp \epsilon e_i h^{-1}) = U_l(\exp \epsilon e_i)U_l(h^{-1}).$$

Substituting these into (7),

$$\mathcal{F}_t[\delta_i p(g, t)] = \left. \frac{d}{de} \right|_{e=0} U_l(\exp \epsilon e_i)\bar{p}_t(t) \triangleq u^l_i \bar{p}_t(t), \quad (8)$$

where $u^l_i$ denotes the derivative of $U_l(g)$ along the direction of $e_i$ when $g$ is the identity element of $G$. Note that $u^l_i$ is also a complex-valued matrix whose size is identical to $U_l$, and it is independent of $g$. In short, the operation of finding derivative along $e_i$ at the group space is equivalent to multiplying the Fourier transform by $u^l_i$ at the transformed space. Using this property repeatedly, the Fokker Planck Equation (4) is transformed into

$$\frac{d}{dt} \bar{p}_t(t) + \sum_{i=1}^{n} u^l_i \mathcal{F}_t[f_i(g, t)p(g, t)]$$

$$- \frac{1}{2} \sum_{i,j=1}^{n} \sum_{k=1}^{m} H_{ik}(t)H^T_{kj}(t)u^l_i u^l_j \bar{p}_t(t) = 0, \quad (9)$$

which is an ordinary differential equation for the Fourier transform [5].

This yields a computation framework to solve the Fokker-Planck equation (4). For a given probability density function $p(g, 0)$ at $t = 0$, its Fourier transform $\bar{p}_t(0)$ is computed by (5), which is numerically integrated according to (9) to obtain $\bar{p}_t(t)$ at $t > 0$. The inverse transform (6) yields the propagated probability density $p(g, t)$. This method is applied to the attitude kinematics on $SO(3)$ in the next subsection.

### C. Uncertainty Propagation for the Attitude Kinematics

Consider the attitude kinematics of a rigid body. The attitude is described by a rotation matrix $R \in SO(3)$, where the special orthogonal group is defined as

$$SO(3) = \{ R \in \mathbb{R}^{3 \times 3} | R^T R = I, \det[R] = 1 \}.$$

Its Lie algebra is denoted by $\mathfrak{so}(3)$, which is the set of $3 \times 3$ skew-symmetric matrices. Rotation matrices are defined such that they correspond to the linear transformation of the representation of a vector from the body-fixed frame to the inertial frame.

A stochastic differential equation for the attitude kinematics is given by

$$R^T dR = \{ \Omega(t) dt + H(t)dW \} \wedge, \quad (10)$$

where $\bar{W}$ is a Wiener process with a matrix $H(t) \in \mathbb{R}^{n \times m}$ that scales and couples noises, and it is assumed that the Lie algebra $\mathfrak{g}$ is identified with $\mathbb{R}^n$. It is defined according to the Ito stochastic calculus [14].
where $\Omega(t) \in \mathbb{R}^3$ is the angular velocity of the rigid body represented with respect to the body-fixed frame, and $W \in \mathbb{R}^m$ denotes a Wiener process. The scaling matrix for noise is given by $H(t) \in \mathbb{R}^{3 \times m}$. Here, the hat map is a linear isomorphism from $\mathbb{R}^3$ to $\mathfrak{so}(3)$, defined such that $\hat{x}y = (x)^{\times}y = x \times y$ for any $x, y \in \mathbb{R}^3$. The inverse of the hat map is defined as the vee map, $(\cdot)^\vee : \mathfrak{so}(3) \to \mathbb{R}^3$ [13].

The stochastic differential equation (10) represents the attitude kinematics where the angular velocity is measured by an unbiased sensor: the sensor measurements correspond to $\Omega(t)$, and the noisy measurement errors are described by the scaling matrix $H(t)$ and the Wiener process $W$.

According to noncommutative harmonic analysis on $\mathbb{SO}(3)$, a probability density function $p(R, t) : \mathbb{SO}(3) \times \mathbb{R}_{+} \to \mathbb{R}_{+}$ can be written as

$$p(R, t) = \sum_{l=0}^{l_{\text{max}}}(2l + 1)\text{tr}[\hat{p}_l(t)U_l(R)].$$

(11)

For the Lie group $\mathbb{SO}(3)$, the Fourier parameters $l$ are integers, and $0 \leq l \leq l_{\text{max}}$ for a selected band limit $l_{\text{max}}$. The $l$-th Fourier transform and the irreducible representation belong to $\mathbb{SO}(3)$, $T^{l}(R) \in \mathbb{C}^{2l+1} \times (2l+1)$ [16].

The corresponding transformed Fokker-Planck equation that is equivalent to (9) can be written as

$$\frac{d}{dt}\hat{p}_l(t) = A_l(t)\hat{p}_l(t),$$

(12)

where the complex-valued matrix $A_l(t) \in \mathbb{C}^{2l+1} \times (2l+1)$ is defined as

$$A_l(t) = -\sum_{i=1}^{3} \Omega_i(t)u_i^l + \frac{1}{2} \sum_{i,j=1}^{m} \sum_{k=1}^{3} H_{ik}(t)H_{kj}^T(t)u_i^l u_j^l,$$

where $\Omega_i(t)$ denotes the $i$-th element of $\Omega(t)$. As the angular velocity measurement is assumed to be independent of $R$ at (10), the transformed Fokker-Planck equation is simplified as a linear time-varying matrix ordinary differential equation of the Fourier transform $\hat{p}_l(t)$ at (12).

D. Probability on $\mathbb{SO}(3)$ and Visualization

Probability theory on $\mathbb{SO}(3)$ is substantially different from $\mathbb{R}^n$ as it is a compact nonlinear manifold. As such, several properties of a probability density function on $\mathbb{SO}(3)$ are discussed. The von Mises distribution is a probability distribution on the unit-circle, $S^1 = \{q \in \mathbb{R}^2 | ||q|| = 1\}$, that is similar to the normal distribution on $\mathbb{R}^n$ [17]:

$$p(\theta) = \frac{1}{2\pi l_0(\kappa)} \exp(\kappa \cos(\theta - \theta_0)),$$

(13)

where $\theta$ is a parameterization of $S^1$, and $\theta_0$ is the mean value. Here, $l_0(\kappa) = \sum_{i=0}^{\infty} \frac{(1/4\kappa)^i}{(2i)!} 2i$, and the parameter $\kappa \in \mathbb{R}_{+}$ determines the shape of the distribution. This is analogous to a normal distribution in $\theta$ with mean $\theta_0$ and variance $\frac{1}{\kappa}$.

For given two rotation matrices $R, R_o \in \mathbb{SO}(3)$, the quantity $\frac{1}{2}(\text{tr}[R_o^T R] - 1)$ represents the cosine of the rotation angle between them. Using this, we introduce a probability density function on $\mathbb{SO}(3)$ from the von Mises distribution:

$$p(R) = \frac{1}{c} \exp \left\{ \frac{1}{2} \kappa (\text{tr}[R_o^T R] - 1) \right\},$$

(14)

where the constant $c$ is chosen such that $\int p(R) \, dR = 1$.

For a given probability density, the mean attitude $R_{\text{mean}} \in \mathbb{SO}(3)$ and the variance $V \in \mathbb{R}^{3 \times 3}$ are chosen as follows:

$$R_{\text{mean}} = \arg \max_{R \in \mathbb{SO}(3)} p(R),$$

(15)

$$V = \int_{\mathbb{SO}(3)} \eta(R)\eta^T(R)p(R) \, dR,$$

(16)

where $\eta(R) = \log(R_o^T R) \in \mathbb{R}^3$. Unlike probability densities on $\mathbb{R}^n$, the mean attitude may not be defined on $\mathbb{SO}(3)$ [15]. At the given definition of variance, the logarithm is not well defined if the rotation angle between $R_{\text{mean}}$ and $R$ is $180^\circ$. However, the measure of such rotation matrices is zero in $\mathbb{SO}(3)$, and this eliminates the effects of singularity.

Visualizing a probability density is useful to understand uncertainty distribution intuitively. However, it is challenging to illustrate $p(R)$ graphically, as probability depends on three-dimensional attitudes. A new approach to visualize attitude uncertainties is proposed in [18]. It is based on the fact that each column of a rotation matrix corresponds to the direction of each body-fixed axis. Therefore, we can construct a marginal probability density of each body-fixed axis from $p(R)$. For example, the probability density function for the first body-fixed axis $b_1 \in \mathbb{R}^3$ (with $||b_1|| = 1$), evolving on the two-dimensional unit-sphere $S^2 = \{q \in \mathbb{R}^3 | ||q|| = 1\}$, can be obtained as

$$p_1(b_1) = \int_{B_1} p(R) dR,$$

where $B_1 = \{R \in \mathbb{SO}(3) | Rc_1 = b_1\}$ and $e_1 = [1, 0, 0]^T \in \mathbb{R}^3$. As probability $p_1(b_1)$ is given as a function of two-dimensional direction of the first body-fixed axis, it can be visualized on a unit-sphere by color shading. And, this is repeated for other axes to visualize attitude uncertainty distribution on a unit-sphere.

E. Numerical Examples

Numerical examples are presented for the attitude kinematics. The initial probability density is chosen as (14) with $R_o = I$ and $\kappa = 8$. Fourier transform, inverse transform, and solution of the Fokker-Planck equation on $\mathbb{SO}(3)$ are computed by gcc using the Message Passing Interface (MPI) library. The following two cases are considered: (i) diffusion only, and (ii) diffusion and advection.

**Diffusion:** At the first case, the angular velocity is $\Omega(t) = 0$, and therefore there is no effect of advection. The matrix $H$ at (10) is chosen as $H = \text{diag}[0.2, 0.1, 0.1]$, i.e., the rotations about the first body-fixed axis are perturbed more severely than rotations about the other axes.

The propagated probability densities are illustrated at Fig. 1. As $\Omega(t) = 0$, the mean attitude illustrated by red arrows does not rotate. Overall, the maximum value of probability
density decreases. But, the given \( H \) yields different uncertainty distributions for each body-fixed axis. Let \( H_i \) be the \( i \)-th element of \( H \). Since \( H_1 > H_2, H_3 \), the direction of the first body-fixed axis is more certain than the other axes (higher, reddish density), and the distributions of the second, and the third body-fixed axes become elongated along the great circle passing through \( b_2 \) and \( b_3 \). Since \( H_2 = H_3 \), the shape of uncertainty distribution for the first body-fixed axis remains circular. The last subfigure illustrates the square root of the diagonal elements of \( V \). They increase over time due to noise, but the first element increases faster as \( H_1 > H_2, H_3 \).

**Diffusion and Advection:** At the second case, the angular velocity is chosen as \( \Omega(t) = [5, 0, 0]^T \text{deg/sec} \), and the matrix \( H \) is same as the first case. As shown at Fig. 2, the mean attitude, illustrated by red arrows, rotates about the first body-fixed axis by 90° over 18 seconds, but the dispersion of uncertainties about the mean attitude is similar to the first case as \( H \) is identical.

This approach does not require any restriction on the type of uncertainty distribution or the magnitude of variance, as it is based on the solutions of the Fokker-Planck equation. It is also constricted intrinsically in a coordinate-independent fashion via noncommutative harmonic analysis. We apply this global uncertainty propagation method for stochastic motion planning and estimation with large uncertainties at the next sections.

### III. Stochastic Motion Planning on SO(3)

There are various motion planning schemes developed for robotic systems in uncertain environments [19], [20]. But, most approaches are based on small, localized uncertainties with Gaussian distribution that is propagated along a linearized dynamic model. The effects of a general class of uncertainties with large variances are not well incorporated in motion planning.

However, a planned motion that is feasible to prescribed stochastic constraints with large uncertainties could be drastically different from trajectories constructed under the assumption that uncertainties are small. To address this issue, a stochastic motion planning scheme is proposed based on the numerical solution of the Fokker-Planck equation that is developed at the previous section.

#### A. Stochastic Optimization

Stochastic motion planning is addressed by formulating a stochastic optimization problem. The stochastic differential equation is extended as follows to incorporate the effects of control inputs \( u(t) \):

\[
g^{-1} dg = f(g, u, t) dt + H(t) dW, \tag{17}
\]

where the vector field \( f \) is dependent of the input. An initial condition is prescribed by a probability density \( p(g, 0) \) at \( t = 0 \). A terminal condition is defined such that the probability that the terminal configuration belongs to a given target set, namely \( C_{target} \subset G \) is greater than a certain threshold, i.e., \( \text{Prob}[g(t_f) \in C_{target}] \geq 1 - \epsilon \) for a small constant \( \epsilon > 0 \). Constraints are also imposed in a stochastic framework such that the probability to enter a forbidden set, namely \( C_{forbid} \) is less than a small threshold always, i.e., \( \text{Prob}[g(t) \in C_{forbid}] \leq \epsilon, \forall t \in [t_0,t_f] \). The objective is to choose a control input trajectory such that the expected value of certain performance criterion, namely \( J \) is maximized, while satisfying the above constraints:

\[
u(t) = \arg \max [E[J(t, g(t), u(t))]].
\]

A useful property is that once an initial probability density function \( p(g, 0) \) and a control input trajectory \( u(t) \) are prescribed, all of the constraints and the objective function can be computed from the propagated density \( p(g, t) \) without imposing any restriction on uncertainty distribution. Then, numerical parameter optimization tools can be applied to find the optimal input trajectory.

#### B. Numerical Examples

Consider a rigid body that is constrained to rotate about a fixed spherical joint (see Fig. 3). The stochastic differential equation is given by (10), where \( \Omega(t) \) is considered as a
control input. This corresponds to the attitude dynamics of
a rigid body where the angular velocity can be arbitrarily
treated as a 3D pendulum, that is a rigid body which is
arbitrarily presented as a 3D pendulum, that is a rigidody where the angular velocity can be arbitrarily
controlled, but it is subjected to noise. This can be imple-
mented if there exists an angular velocity control system
whose timescale is sufficiently faster than the desired attitude
maneuver.

The initial time and the terminal time are chose as \( t_0 = 0, \)
\( t_f = 1 \) second. The initial probability density is given as (14)
with \( \kappa = 8 \) and \( R_0 = I \), i.e., the initial attitude distribution
is centered at the identity. The desired terminal attitude is
\[
R_{\text{target}} = \exp(2\pi/3\hat{e}_1) \exp(\pi/6\hat{e}_3),
\]
where \( e_1 = [1, 0, 0]^T, e_3 = [0, 0, 1]^T \in \mathbb{R}^3 \). Instead of
imposing the target attitude as an inequality constraint, the
probability density to reach the target attitude is maximized:
\[
\Omega(t) = \arg \max \{ p(R_{\text{target}}, t_f) \}. \tag{18}
\]

There are two obstacles located at \( r_1 = [-0.11, -0.98, 0.15]^T, \)
\( r_2 = [-0.31, -0.79, 0.53]^T \in \mathbb{R}^3 \). The rigid body is assumed to be
elongated along its third body-fixed axis such that collision with obstacles is avoided
if the third-body fixed axis does not rotate close to \( r_1 \)
and \( r_2 \). The corresponding inequality constraint to avoid
collision is given by
\[
\text{Prob}[R(t) \in C_{\text{forbid}}] \leq \epsilon, \quad \forall t \in [t_0, t_f] \tag{19}
\]
where \( C_{\text{forbid}} = \{ R \in SO(3) | \min \{ (R \hat{e}_3 \cdot r_1, (R \hat{e}_3 \cdot r_2) > \cos \frac{\pi}{2} \} \} \).

The angular velocity is linearly parameterized by 33 points
that are uniformly distributed over \([t_0, t_f]\) and the matrix \( H \)
is chosen as \( H = 0.01 \text{diag}[2, 3, 1] \). The objective function
and the constraint are computed by the solution of the
Fokker-Planck equation based on noncommutative harmonic
analysis, and they are optimized by a sequential quadratic
programming routine in gcc. Two cases are considered for
varying thresholds of the inequality constraint (19): \( \epsilon = 0.3 \)
(less conservative) and \( \epsilon = 0.1 \) (more conservative).

The optimized probability densities evaluated at the target
attitude for both cases are similar: \( p(R_{\text{target}}, t_f) = 106.77 \)
(\( \epsilon = 0.3 \)), 106.47 (\( \epsilon = 0.1 \)). The corresponding optimized
angular velocity and snapshots of a sample trajectory with
\( R(0) = I \) are illustrated at Fig. 3. When \( \epsilon = 0.3 \), the
rigid body rotates between two obstacles along a shorter
path with relatively smaller angular velocities. But, when
the inequality constraint to avoid collision is more conservative,
i.e., \( \epsilon = 0.1 \), the rigid body rotates around both obstacles,
while the third body-fixed axis follows a longer path with
larger angular velocities. These illustrate that stochastically
optimized trajectories can be substantially changed according
to the amount of uncertainties considered.

IV. GLOBAL ESTIMATION ON SO(3)

The extended Kalman filter and its variations have been
widely used for estimation. But, they explicitly rely on the
assumptions that a nonlinear dynamic system can be well-
approximated by its linearized dynamic model relative to a
prescribed reference trajectory and priori distributions are
Gaussian such that the corresponding posteriori densities are
also Gaussian. However, the Gaussian assumption greatly
reduces the amount of information that is contained in the
true density, and the linearization is valid only in small
neighborhoods of the reference trajectory. Therefore, the
performance of an extended Kalman filter is limited when the
initial estimate is poor or there exist large uncertainties [11].
In this section, the preceding uncertainty propagation
approach is applied to construct global nonlinear Bayesian
estimation techniques on a Lie group.

A. Nonlinear Bayesian Estimation

Bayesian estimation approach is summarized as follows.
Consider the stochastic differential equation given by (2).
Let \( p(g, 0) \) be an initial probability density. It is propagated
by (9) until \( t > 0 \) to obtain a priori probability density \( p^-(g, t) \).
Suppose that a function of the configuration, namely \( Z(g) \in \mathbb{R}^p \)
is measured by a sensor and the corresponding sensor measurement is given by
\( z \in \mathbb{R}^p \). Assuming that the measurement noise is independent to the process
noise, a posterior density namely \( p^+(t, g) \) that integrates the predicted
stochastic information with the measurement is constructed according to Bayes’ rule as
\[
p^+(g, t|z) = \frac{1}{c} p^-(g, t)p_{z|g}(z|g),
\]
where \( c \) is a normalizing constant that is chosen such that
\( \int p^+(g, t|z)dg = 1 \). This process is repeated whenever new
measurement data are available.

B. Numerical Example

This is applied to an attitude estimation problem. The true
attitude dynamics is modeled as a 3D pendulum, that is a

Fig. 3. Stochastic motion planning of a rigid body
rigid body rotating about a fixed pivot under the uniform gravitational potential [21]. The true initial attitude and the angular velocity are chosen as \( R_{\text{true}}(0) = \text{diag}[-1, 1, -1] \in \text{SO}(3) \) and \( \Omega_{\text{true}}(0) = [0.5, -0.5, 0.4]^T \in \mathbb{R}^3 \), which yield complex attitude and angular velocity trajectories as the initial condition is close to the unstable, inverted equilibrium.

The initial probability density for the attitude is given by (14) with \( R_o = I \) and \( \kappa = 8 \). Therefore, the initially estimated mean attitude has 180° error compared with the true initial attitude. It is assumed that the angular velocity is measured frequently at every 0.01 seconds with \( \dot{H} = \text{diag}[0.3, 0.1, 0.1] \). The attitude is measured intermittently at every 0.5 seconds, and the probability density for the attitude measurement errors is also modeled as (14) where \( R_o \) corresponds to the true attitude and \( \kappa = 14 \). This corresponds to the case where extended Kalman filters may not perform well: the initial attitude estimate is falsely too confident about a completely incorrect attitude; the attitude is measured infrequently; and the attitude trajectory is complicated and the corresponding linearized dynamic model is inaccurate.

The resulting estimation error is illustrated at Fig. 4. The initial estimation error is 180°, but it reduces significantly when the first attitude measurement is available at \( t = 0.5 \). The estimation error remains under 2° afterwards. The estimated attitude probability densities are visualized at Fig. 5, where the true attitudes are illustrated by black arrows, and the estimated mean attitudes are illustrated by red arrows (after \( t = 0.8 \), they are indistinguishable at the given figures as they become very close). Initially, the estimated error is opposite to the true attitude, but the estimation error is reduced at \( t = 0.8 \). The variance is relatively large at \( t = 1.6 \) (less certain), and it becomes smaller at \( t = 7.2 \) (more certain).

These illustrate desirable convergence properties of the proposed attitude estimator for large initial estimation errors and large uncertainties. Compared with extended Kalman filters, the proposed approach does not have stochastic information loss caused by Gaussian approximations, and it is effective for any type of probability distributions.

REFERENCES