Uncertainty Propagation on Unimodular Lie Groups Using a Gaussian Approximation

Jikai Ye and Gregory S. Chirikjian

***Abstract*—We discuss the connection between two definitions of stochastic differential equations (SDEs) on unimodular Lie groups and derive the mean and covariance propagation equations in this work. Starting from an SDE defined on Lie groups via Mckean-Gangolli injection, we first convert it to a parametric SDE in exponential coordinates. The coefficient transform method for the conversion is stated for both Itô's and Stratonovich's interpretation of the SDE. Then we derive a mean and covariance fitting formula for probability distributions on Lie groups defined by a concentrated distribution on the exponential coordinate. It is used to derive the mean and covariance propagation equations for the SDE defined by injection, which coincides with the equations derived from a Fokker-Planck equation. The Gaussian distribution constructed from the mean and covariance can be used for calculating the cost function of diffusion models on Lie groups.**

I. Introduction

Stochastic differential equations (SDEs) play an important role in many fields of engineering, from noise modeling in dynamical systems to diffusion models in deep learning [6]. When the state variable $\boldsymbol{x}$ lives in Euclidean space, the general form of a Itô SDE is

$$d\boldsymbol{x} = \boldsymbol{h}(\boldsymbol{x}, t)dt + H(t)d\boldsymbol{W} \tag{1}$$

and the propagation of the mean $\boldsymbol{\mu}(t)$ and covariance $\Sigma(t)$ of the state variable $\boldsymbol{x}$ is well-known [12]

$$\begin{cases} \dot{\boldsymbol{\mu}} = \langle \boldsymbol{h} \rangle \\ \dot{\Sigma} = \langle \boldsymbol{h}(\boldsymbol{x}-\boldsymbol{\mu})^T + (\boldsymbol{x}-\boldsymbol{\mu})\boldsymbol{h}^T \rangle + HH^T \end{cases} \tag{2}$$

where $\langle \varphi \rangle \doteq \mathbb{E}(\varphi)$. A famous example is the Ornstein–Uhlenbeck process where $\boldsymbol{h}(\boldsymbol{x}, t) = A\boldsymbol{x} + \boldsymbol{b}$. When the initial state is $\boldsymbol{x}_0$, the conditional probability density function of $\boldsymbol{x}_t$, $p_{t|0}(\boldsymbol{x}_t|\boldsymbol{x}_0)$, will remain a Gaussian during the process and the mean and covariance can be calculated by (2). This property is utilized in calculating the cost function for score-based generative models [6]

$$\ell_t(\boldsymbol{s}_\theta) = \mathbb{E}[\|\boldsymbol{s}_\theta(\boldsymbol{x}_t) - \nabla_{x_t} \log p_{t|0}(\boldsymbol{x}_t|\boldsymbol{x}_0)\|^2] \tag{3}$$

where $p_{t|0}(\boldsymbol{x}_t|\boldsymbol{x}_0)$ has an analytic expression and is thus easy to compute.

In robotic tasks, however, the state space usually has a Lie group structure, *e.g.* $SO(3)$ and $SE(3)$. The expression of the probability density function $p_{t|0}(g_t|g_0)$ on Lie groups is not well understood. This paper derives a mean and covariance propagation equation in this setting from a stochastic differential equation perspective, which provides more insight into the equation derived by a Fokker-Planck equation on unimodular Lie groups in [16]. The connection between different ways of defining stochastic differential equations on Lie groups is also stated. The Gaussian distribution constructed from the propagated mean and covariance can be used to calculate the cost function for diffusion models on Lie groups as in (3), which have found applications in robotics tasks, such as 6D grasping [15] and robot motion generation [5].

When the state variable lives in a Lie group, the noisy dynamic model is a stochastic differential equation (SDE) on the Lie group. There exist several ways to define such an SDE. In [14, 10], a matrix Lie group is considered as a submanifold embedded in $GL(n)$, and the SDE is defined on $GL(n)$ directly. Another way is to use the McKean-Gangolli injection [11, 3] in which an infinitesimal stochastic process on the Lie algebra is projected to the Lie group. These two types of SDEs can be interconverted if the coefficients satisfy a simple relationship [14]. A third way is to parametrize the group and define an SDE on the parameter space. In this paper, we start with the second type of SDE and convert it to the third type. When using exponential coordinates for parametrization and Stratonovich's interpretation, the coefficient conversion is very simple, while when using Ito's interpretation, an additional drift term shows up. Note that the Fokker-Planck equation corresponding to these two equivalent SDEs is derived in [16], which describes the evolution of the probability density function of the state variable.

We proceed to review the mean and covariance propagation methods. In extended Kalman filter type of methods, the mean is propagated using the dynamic model without noise [1, 2] and the propagation of covariance is derived by expansion and truncation. In the literature that employs the unscented transform, the mean and covariance propagation can be calculated by propagating sigma points on the group and performing optimization [7, 9], by propagating sigma points on the tangent space and projecting the mean and covariance back to the group [9, 13], or by propagating the mean using the deterministic dynamic model [3, 4] and the covariance by the unscented transform. In this work, we take the second approach which utilizes the tangent space for propagation. Different from previous work that projects mean and covariance alone, we project quantities on the tangent space back to the group while taking into account the influence of the probability distribution. Also, we do not make approximations in the derivation and arrive at propagation equations in the form of (2) assuming the initial probability distribution is concentrated. A propagation equation of the same form has been derived previously in [16] from a Fokker-Planck equation's perspective and the approximate equation based on it has been demonstrated by experiments. We offer another derivation that provides more insight into the meaning of each term in the equation from an SDE perspective.

Contributions: i) We state the equivalence between a non-parametric stochastic differential equation (SDE) on a Lie group defined by Mckean-Gangolli injection and a parametric SDE defined on exponential coordinates of the Lie group. ii) A formula with error analysis is derived for fitting the group-theoretic mean and covariance of a probability defined on exponential coordinates of a Lie group. iii) A continuous-time mean and covariance propagation equation is derived using exponential coordinates and the mean and covariance fitting method, which can then be used to construct a Gaussian distribution to approximate the true distribution for diffusion model training.

II. Background

In this section, we provide a minimal introduction to Lie group theory. The *Einstein summation* convention is used to simplify notations throughout this paper.

An N-dimensional matrix Lie group G is an N-dimensional analytic manifold and also a subgroup of the general linear matrix $GL(n)$ with group product and inverse operation being analytic. The Lie algebra $\mathcal{G}$ of a Lie group is the tangent space at the identity of G equipped with a Lie bracket. In the case of a N-dimensional matrix Lie group, its Lie algebra $\mathcal{G}$ can be understood as a N-dimensional linear space consisting of matrices whose matrix exponential are in G and the Lie bracket is defined by

$$[X, Y] \doteq XY - YX, \ X, Y \in \mathcal{G}. \tag{4}$$

Given a basis of $\mathcal{G}$ as $\{E_i\}_{i=1,2,...,N}$, we can draw equivalence between the Lie algebra $\mathcal{G}$ and $\mathbb{R}^N$ using the '$\wedge$' and '$\vee$' operation: $\boldsymbol{x}^\wedge \doteq \sum_{i=1}^N x_i E_i \in \mathcal{G}, \ \boldsymbol{x} \in \mathbb{R}^N$ and $X^\vee = \boldsymbol{x} \in \mathbb{R}^N$ which is the inverse of '$\wedge$'. This identification of E_i with $\boldsymbol{e}_i$ is equivalent to fixing a metric for G. The little 'ad' operator is defined by $ad_X Y \doteq [X, Y], \ X, Y \in \mathcal{G}$. This operator is a linear operator on Y and can be transformed into a matrix $[ad_X] \in \mathbb{R}^{N\times N}$ that satisfies $[ad_X]\boldsymbol{y} = (ad_X Y)^\vee$.

Since a matrix Lie group is also a manifold, we can locally parametrize it by a subset of $\mathbb{R}^N$ as $g(\boldsymbol{q}) \in G$ where $\boldsymbol{q} \in \mathbb{R}^N$. One parametrization that exists for all matrix Lie groups is the *exponential coordinate*, where parametrization around $\mu \in G$ is obtained by the matrix multiplication and matrix exponential, $g(\boldsymbol{x}) = \mu \exp(\boldsymbol{x}^\wedge)$. The neighborhood of any group element $\mu \in G$ can be parametrized in this way and the parametrization is a local diffeomorphism between $D \subseteq \mathbb{R}^N$ and G. The domain of exponential coordinate, D, is specified case by case, for example $D = \{\boldsymbol{x} \in \mathbb{R}^3 \,|\, \|\boldsymbol{x}\|_2 < \pi\}$ for $SO(3)$.

The left and right Jacobian matrices of G are defined by

$$J_l(\boldsymbol{q}) = \left[\left(\frac{\partial g(\boldsymbol{q})}{\partial q_1} g^{-1} \right)^\vee, \left(\frac{\partial g(\boldsymbol{q})}{\partial q_2} g^{-1} \right)^\vee, ..., \left(\frac{\partial g(\boldsymbol{q})}{\partial q_N} g^{-1} \right)^\vee \right], \tag{5}$$

$$J_r(\boldsymbol{q}) = \left[\left(g^{-1} \frac{\partial g(\boldsymbol{q})}{\partial q_1} \right)^\vee, \left(g^{-1} \frac{\partial g(\boldsymbol{q})}{\partial q_2} \right)^\vee, ..., \left(g^{-1} \frac{\partial g(\boldsymbol{q})}{\partial q_N} \right)^\vee \right], \tag{6}$$

where $g(\boldsymbol{q})$ can be the exponential or any other parametrization. One use of Jacobians is to solve ordinary differential equations on G by parametrization. For example,

$$(g^{-1}\dot{g})^\vee = \boldsymbol{h}(g) \iff \dot{\boldsymbol{q}} = J_r^{-1}(\boldsymbol{q})\boldsymbol{h}(g(\boldsymbol{q})) \tag{7}$$

when the Jacobian matrix is not singular. The Jacobian can also be used to construct two Haar measures on G defined by $d_l g \doteq |\det J_l| d\boldsymbol{q}$ and $d_r g \doteq |\det J_r| d\boldsymbol{q}$. When $|\det J_l| \equiv |\det J_r|$, the group is called *unimodular*, *e. g.* $SE(2)$, $SE(3)$, $SO(3)$, and the two Haar measures are both invariant to left and right shift.

In this paper, we only consider functions whose support is within the domain of exponential coordinates centered at some group element μ, i.e. $\mathrm{supp}(f) \subseteq \mu \exp(D)$, and the integration can be calculated on the exponential coordinate by

$$\int_G f(g) dg \doteq \int_D f(\mu \exp(\boldsymbol{x})) |J_r(\boldsymbol{x})| d\boldsymbol{x}. \tag{8}$$

To simplify equations, we notationally suppress the '$\wedge$' operator in the exponential map, which we will continue to use throughout this paper.

A probability density function $p(g)$ on G is a function that satisfies: i) $p(g) \geq 0$ and ii) $\int_G p(g) dg = 1$. Building on terminology already in use in literature [13, 2], we now formally define the concept of *concentrated distribution* on $\mathbb{R}^N$:

Definition 1. *If a probability distribution on $\mathbb{R}^N$ satisfies the following properties, it is said to be a concentrated distribution for Lie group G: i) the support of the probability density function $\tilde{p}(\boldsymbol{x})$ is within the domain of exponential coordinates of G, i.e. $supp(\tilde{p}) \subseteq D$, and ii) the mean of the distribution is around $\boldsymbol{0}$, i.e. $\|\mathbb{E}(\boldsymbol{x})\|_2 << 1$.*

Given a concentrated distribution on R^N, $\tilde{p}(\boldsymbol{x})$, we can construct a probability distribution on unimodular Lie groups by defining a random variable $g(\boldsymbol{x}) = \mu \exp(\boldsymbol{x}) \in G$. Denote the probability density function of this random variable as $p(g)$. The expectation of a function $f(g)$ can be calculated as

$$\langle f \rangle \doteq \int_G f(g) p(g) dg = \int_{\mathbb{R}^N} f(\mu \exp(\boldsymbol{x})) \tilde{p}(\boldsymbol{x}) d\boldsymbol{x} \tag{9}$$

where

$$p(\mu \exp(\boldsymbol{x})) |J_r(\boldsymbol{x})| = \tilde{p}(\boldsymbol{x}). \tag{10}$$

The *group-theoretic* mean and covariance of $p(g)$ denoted by μ_G and Σ are defined below:

$$\int_G \log^\vee(\mu_G^{-1} g) p(g) dg = \boldsymbol{0} \tag{11}$$

and

$$\Sigma = \int_G [\log^\vee(\mu_G^{-1} g)][\log^\vee(\mu_G^{-1} g)]^T p(g) dg. \tag{12}$$

III. Stochastic Differential Equations on Lie Groups

Suppose we have a N-dimensional matrix Lie group, G, a vector-valued function, $\boldsymbol{h}: G\times\mathbb{R}\rightarrow\mathbb{R}^N$, and a matrix-valued function, $H: G\times\mathbb{R}\rightarrow\mathbb{R}^{N\times N}$. Denote a N-dimensional Wiener process as $\boldsymbol{W}(t)$ which satisfies $\big(\boldsymbol{W}(t+s)-\boldsymbol{W}(t)\big)\sim\mathcal{N}(\boldsymbol{0}, s\cdot\mathbb{I}_{N\times N})$.

Definition 2. *(non-parametric SDE on G) A stochastic differential equation on G can be defined non-parametrically via Mckean-Gangolli injection [11],*

$$g(t+dt)=g(t)\exp\big(\boldsymbol{h}\Big|_{\substack{g=g(t)\\t=t}}dt+H\Big|_{\substack{g=g(t+\kappa dt)\\t=t+\kappa dt}}d\boldsymbol{W}\big),\quad(13)$$

where dt is the infinitesimal increment of time and $d\boldsymbol{W}\doteq\boldsymbol{W}(t+dt)-\boldsymbol{W}(t)$.

The SDE is called Itô's when $\kappa=0$ and Stratonovich's when $\kappa=\frac{1}{2}$, which is consistent with the definitions on Euclidean space [8]. A sample path of the non-parametric SDE (13) starting at $g(0)$ is defined by the following limit

$$g(T)=\lim_{M\rightarrow\infty}g(0)\prod_{i=0}^{M-1}\exp\left[\boldsymbol{h}\Big|_{\substack{g=g(t_i)\\t=t_i}}\Delta t+H\Big|_{\substack{g=g(t_i+\kappa\Delta t)\\t=t_i+\kappa\Delta t}}\Delta\boldsymbol{W}_i\right]\quad(14)$$

where $\Delta t=T/M$, $t_i=i\Delta t$, $\Delta\boldsymbol{W}_i=\boldsymbol{W}(t_{i+1})-\boldsymbol{W}(t_i)$, and the exponential of increment is multiplied on the right sequentially.

Another way to define an SDE on Lie groups is to parameterize group elements and define an SDE on the parameter space:

Definition 3. *(parametric SDE on G) A stochastic differential equation on G can be defined parametrically by parametrizing group elements as $g=g(\boldsymbol{q})$ and writing an SDE on the parameter space as*

$$\boldsymbol{q}(t+dt)=\boldsymbol{q}(t)+(J_r^{-1}\tilde{\boldsymbol{h}})\Big|_{\substack{\boldsymbol{q}=\boldsymbol{q}(t)\\t=t}}dt+(J_r^{-1}\tilde{H})\Big|_{\substack{\boldsymbol{q}=\boldsymbol{q}(t+\kappa dt)\\t=t+\kappa dt}}d\boldsymbol{W}\quad(15)$$

where $\tilde{\boldsymbol{h}}(\boldsymbol{q},t)$ and $\tilde{H}(\boldsymbol{q},t)$ are functions of the parameters and time.

As before, it is called Itô's when $\kappa=0$ and Stratonovich's when $\kappa=1/2$. For a short time, assume the trajectory $\boldsymbol{q}(t)$ is still in the domain of the parametrization, a sample path of the SDE starting at $\boldsymbol{q}(0)$ is defined by the following limit [8]

$$\boldsymbol{q}(T)=\boldsymbol{q}(0)+\lim_{M\rightarrow\infty}\sum_{i=0}^{M-1}\left\{(J_r^{-1}\tilde{\boldsymbol{h}})\Big|_{\substack{\boldsymbol{q}=\boldsymbol{q}(t_i)\\t=t_i}}\Delta t+(J_r^{-1}\tilde{H})\Big|_{\substack{\boldsymbol{q}=\boldsymbol{q}(t_i+\kappa\Delta t)\\t=t_i+\kappa\Delta t}}\Delta\boldsymbol{W}_i\right\}\quad(16)$$

where $\Delta t=T/M$, $t_i=i\Delta t$ and $\Delta\boldsymbol{W}_i=\boldsymbol{W}(t_{i+1})-\boldsymbol{W}(t_i)$. The path is then mapped back to G by $g(t)=g(\boldsymbol{q}(t))$.

When using exponential coordinates to parametrize the group, i.e. $g(\boldsymbol{x})=\mu\exp(\boldsymbol{x})$, the sample paths of equation (13) and equation (15) are related by the following theorem:

Theorem 1. *Using the parametrization $g(\boldsymbol{x})=\mu\exp(\boldsymbol{x})$ in Definition 3, when both equation (13) and equation (15) are interpreted as Ito's SDEs, i.e. $\kappa=0$, their sample paths are equivalent if the following condition holds*

$$\tilde{\boldsymbol{h}}\Big|_{\substack{\boldsymbol{x}=\boldsymbol{x}\\t=t}}=\boldsymbol{h}\Big|_{\substack{g=\mu\exp(\boldsymbol{x})\\t=t}}+\left(\frac{1}{2}J_r\frac{\partial J_r^{-1}}{\partial x_k}HH^TJ_r^{-1}\right)\Bigg|_{\substack{g=\mu\exp(\boldsymbol{x})\\\boldsymbol{x}=\boldsymbol{x}\\t=t}}\boldsymbol{e}_k,$$
$$\tilde{H}\Big|_{\substack{\boldsymbol{x}=\boldsymbol{x}\\t=t}}=H\Big|_{\substack{g=\mu\exp(\boldsymbol{x})\\t=t}}.\quad(17)$$

Remark: At first glimpse, it is surprising that an additional appears when using a parametric SDE to describe a non-parametric SDE. That term comes from the non-linearity of the exponential map which is used in defining the non-parametric SDE. Previous work [10] also observe a similar phenomenon in their definitions of SDE.

Theorem 2. *Using the parametrization $g=\mu\exp(\boldsymbol{x})$ in equation (15), when both the non-parametric and the parametric SDEs are interpreted as Stratonovich's SDEs, i.e. $\kappa=1/2$, their solutions are equivalent when*

$$\tilde{\boldsymbol{h}}\Big|_{\substack{\boldsymbol{x}=\boldsymbol{x}\\t=t}}=\boldsymbol{h}\Big|_{\substack{g=\mu\exp(\boldsymbol{x})\\t=t}}\quad\text{and}\quad\tilde{H}\Big|_{\substack{\boldsymbol{x}=\boldsymbol{x}\\t=t}}=H\Big|_{\substack{g=\mu\exp(\boldsymbol{x})\\t=t}}.\quad(18)$$

Remark: The connection between a parametric and a non-parametric SDE on G is natural when both are interpreted as Stratonovich's. Their corresponding Fokker-Planck equation also takes a simple form [16].

IV. Mean and Covariance Propagation on Lie Groups

In this section, we provide a way to derive the mean and covariance propagation equations on unimodular Lie groups from an SDE perspective. We first present a theorem that estimates the group-theoretic mean and covariance of a probability distribution on the exponential coordinate with error analysis. It is then used to derive a continuous-time mean and covariance propagation equation.

The following theorem gives an estimation of the group-theoretic mean and covariance of the random variable $g(\boldsymbol{x})=\mu\exp(\boldsymbol{x})$ when $\boldsymbol{x}$ obeys a concentrated distribution:

Theorem 3. *Given a random variable $\boldsymbol{x}\in\mathbb{R}^N$ whose probability distribution is concentrated for G. Denote its mean, covariance matrix, and the probability density function by $\boldsymbol{m}$, Σ, and $\tilde{p}(\boldsymbol{x})$. The random variable defined by $g=\mu\exp(\boldsymbol{x})$ obeys a distribution whose group-theoretic mean μ_m and covariance Σ_m are estimated by*

$$\begin{aligned}\boldsymbol{m}'&=\langle J_l^{-1}\rangle^{-1}\boldsymbol{m}\\\mu_m&=\mu\exp(\boldsymbol{m}'+\mathcal{O}(|\boldsymbol{m}'|^2))\\\Sigma_m&=\Sigma-sym\big(\langle J_l^{-1}\boldsymbol{m}'\boldsymbol{x}^T\rangle\big)+\mathcal{O}(|\boldsymbol{m}'|^2)\end{aligned}\quad(19)$$

where $\langle v(\boldsymbol{x})\rangle\doteq\int_{\mathbb{R}^N}v(\boldsymbol{x})\tilde{p}(\boldsymbol{x})d\boldsymbol{x}$ and $sym(A)\doteq A+A^T$.

Remark: This theorem states that when the probability density function $\tilde{p}(\boldsymbol{x})$ is close to a Dirac-delta function, the group-theoretic mean is close to $\mu\exp(\boldsymbol{m})$. However, when the probability is relatively dispersed, this approximate can lead to a $\mathcal{O}(|\boldsymbol{m}|)$ level error.

Suppose we have a stochastic process on Lie group G described by a non-parametric Ito's SDE (13) and assume H is a constant matrix. We aim to derive the propagation equations for the group-theoretic mean $\mu(t)$ and covariance $\Sigma(t)$ of $g(t)$. At time t, we parametrize group elements by $g(t) = \mu(t)\exp(\boldsymbol{x}(t))$ and write the probability density function of $\boldsymbol{x}$ as $\tilde{p}(\boldsymbol{x}, t)$. We assume $\tilde{p}(\boldsymbol{x}, t)$ to be a concentrated distribution for G. Using Theorem 1 and Theorem 3, we derive the following propagation equations:

Theorem 4. *The group-theoretic mean $\mu(t)$ and covariance $\Sigma(t)$ of a stochastic process $g(t)$ described by the Itô's SDE (13) obey the following ordinary differential equations:*

$$(\mu^{-1}\dot{\mu})^\vee = \langle J_l^{-1}\rangle^{-1}\left\langle \frac{1}{2}\frac{\partial J_r^{-1}}{\partial x_k}(HH^T J_r^{-T}\boldsymbol{e}_k) + J_r^{-1}\boldsymbol{h}^c \right\rangle \tag{20}$$

and

$$\dot{\Sigma} = \left\langle sym\left[(\frac{1}{2}\frac{\partial J_r^{-1}}{\partial x_k}(HH^T J_r^{-T})\boldsymbol{e}_k - J_l^{-1}(\mu^{-1}\dot{\mu})^\vee + J_r^{-1}\boldsymbol{h}^c)\boldsymbol{x}^T \right] + J_r^{-1}HH^T J_r^{-T} \right\rangle. \tag{21}$$

where $\boldsymbol{h}^c(\boldsymbol{x}, t) \doteq \boldsymbol{h}(\mu\exp(\boldsymbol{x}), t)$ and $sym(A) \doteq A + A^T$.

In practice, we could calculate the expectation approximately using Taylor's expansion or unscented transform as demonstrated in [16]. After having the mean and covariance, we can construct a concentrated Gaussian distribution $g = \mu(t)\exp(\boldsymbol{x})$, $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{0}, \Sigma(t))$ as an approximate probability distribution. In score-based generative models, it can be used to compute $\nabla \log p_t(g_t|g_0)$ approximately [6].

V. Conclusion

This paper provides a derivation to the mean and covariance propagation equations when the dynamics model is a stochastic differential equation (SDE) on Lie groups. We first derive the relationship between a non-parametric SDE defined by Mckean-Gangolli injection and a parametric SDE on exponential coordinates. Then we derive a mean and covariance fitting formula for probability distributions on Lie groups that are defined by projecting a concentrated distribution on the exponential coordinate to the group. Combining these two tools, we derive the mean and covariance propagation equations for a non-parametric SDE. In the future, we will apply the approximate propagation method to train score-based generative models on Lie groups.

References

[1] Axel Barrau and Silvère Bonnabel. The invariant extended Kalman filter as a stable observer. *IEEE Transactions on Automatic Control*, 62(4):1797–1812, 2016.

[2] Guillaume Bourmaud, Rémi Mégret, Marc Arnaudon, and Audrey Giremus. Continuous-discrete extended Kalman filter on matrix Lie groups using concentrated Gaussian distributions. *Journal of Mathematical Imaging and Vision*, 51:209–228, 2015.

[3] Martin Brossard, Silvere Bonnabel, and Jean-Philippe Condomines. Unscented Kalman filtering on Lie groups. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2485–2491. IEEE, 2017.

[4] Martin Brossard, Axel Barrau, and Silvere Bonnabel. A code for unscented Kalman filtering on manifolds (ukf-m). In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5701–5708. IEEE, 2020.

[5] Joao Carvalho, An T Le, Mark Baierl, Dorothea Koert, and Jan Peters. Motion planning diffusion: Learning and planning of robot motions with diffusion models. In *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1916–1923. IEEE, 2023.

[6] Valentin De Bortoli, Emile Mathieu, Michael Hutchinson, James Thornton, Yee Whye Teh, and Arnaud Doucet. Riemannian score-based generative modelling. *Advances in Neural Information Processing Systems*, 35: 2406–2422, 2022.

[7] James Richard Forbes and David Evan Zlotnik. Sigma point Kalman filtering on matrix Lie groups applied to the slam problem. In *Geometric Science of Information: Third International Conference, GSI 2017, Paris, France, November 7-9, 2017, Proceedings 3*, pages 318–328. Springer, 2017.

[8] Crispin W Gardiner et al. *Handbook of stochastic methods*, volume 3. springer Berlin, 1985.

[9] Søren Hauberg, François Lauze, and Kim Steenstrup Pedersen. Unscented Kalman filtering on Riemannian manifolds. *Journal of mathematical imaging and vision*, 46:103–120, 2013.

[10] Xinghan Li, Jianqi Chen, Han Zhang, Jieqiang Wei, and Junfeng Wu. Errors dynamics in affine group systems. *arXiv preprint arXiv:2307.16597*, 2023.

[11] Henry P McKean. *Stochastic integrals*, volume 353. American Mathematical Soc., 1969.

[12] Simo Särkkä and Juha Sarmavuori. Gaussian filtering and smoothing for continuous-discrete dynamic systems. *Signal Processing*, 93(2):500–510, 2013.

[13] Alexander Meyer Sjøberg and Olav Egeland. Lie algebraic unscented Kalman filter for pose estimation. *IEEE Transactions on Automatic Control*, 67(8):4300–4307, 2021.

[14] Victor Solo and Gregory S Chirikjian. Ito, Stratonovich and geometry. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 3026–3032. IEEE, 2019.

[15] Julen Urain, Niklas Funk, Jan Peters, and Georgia Chalvatzaki. Se (3)-diffusionfields: Learning smooth cost functions for joint grasp and motion optimization through diffusion. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5923–5930. IEEE, 2023.

[16] Jikai Ye, Amitesh S Jayaraman, and Gregory S Chirikjian. Uncertainty propagation on unimodular matrix Lie groups. *arXiv preprint arXiv:2312.03348*, 2023.