A Note on the Extended Kalman Filter on a Manifold

Yixiao Ge and Pieter van Goor and Robert Mahony

Abstract— The kinematics of many control systems, especially in the robotics field, naturally live on smooth manifolds. Most classical state-estimation algorithms, including the extended Kalman filter, are posed on Euclidean space. Although any filter algorithm can be adapted to a manifold setting by implementing it in local coordinates and ignoring the geometric structure, it has always been clear that there would be advantages in taking the geometric structure into consideration in developing the algorithm. In this paper, we argue that the minimum geometric structure required to adapt the extended Kalman filter to a manifold is that of an affine connection. With this structure, we show that a naive coordinate implementation of the EKF fails to account for geometry of the manifold in the update step and in the reset step. We provide geometric modifications to the classical EKF based on parallel transport of the measurement covariance (for the update) and *a-posteriori* state covariance (for the reset) that address these limitations. Preliminary results for attitude estimation with two directional measurements demonstrate that the proposed modifications significantly improve the transient behavior of the filter.

I. INTRODUCTION

Over the past sixty years, the extended Kalman filter (EKF) has been the industry standard for state estimation problems when the system dynamics are governed by non-linear equations [1][2]. Although the classical formulation of EKF is given in global Euclidean space, there have been many papers that adapt the Kalman filter methodology to systems that live on smooth manifolds. Although any choice of local coordinates provides a representation in which the classical extended Kalman filter can be implemented, from as early as the 1970s, authors were demonstrating the advantage of choosing local coordinate charts that encode geometric structure, such as Riemannian metrics or homogeneous symmetries, in a natural manner [3][4]. Such charts have specific structure that can lead to lower linearisation error for the filter improving performance. The approach comes at the cost of using a new chart for every iteration of the algorithm, since the nice geometric properties of the chart usually only hold at a single point, the origin or reference point of the chart. This structure is also the foundation of the ⊞ ('boxplus') and ⊟ ('boxminus') operators used for modelling state displacement on manifold, introduced in [5]. Similar techniques were used in [6] and [7]. One example of local coordinates centered at a point are the normal coordinates comprising geodesics in a star shaped neighbourhood associated with an affine connection [4]. This is particularly of interest for Riemannian manifolds where Levi-Civita connection is the unique torsion free connection that preserves the metric [3][8][9]. However, normal coordinates can be defined for an arbitrary affine connection and do not require a Riemannian metric, for example, one-parameter subgroups on a Lie-group state-space.

A parallel, and closely related, research thread considers geometric structure induced by global symmetry properties of the system state-space. A key early problem that motivated this perspective was the question of attitude estimation in aerospace applications. The use of quaternions or rotation matrices for orientation representation led to the multiplica-tive extended Kalman filter (MEKF) [10]. This methodology uses an extended Kalman filter based on a linearisation of a global error defined using the group structure of the state-space. Although the MEKF became an industry standard for attitude estimation in aerospace [11] the approach was not extended beyond the quaternion and rotation groups until picked up again after the turn of the century. The advent of remotely piloted aerial vehicle led to a revaluation of the attitude filtering problem [12][13]. Bonnabel *et al.* proposed a general theory for the Invariant Extended Kalman Filter (IEKF) for systems on Lie-groups in a series of works [14][15], motivated originally by attitude estimation. They identified a class of 'group affine' systems for which they showed that the IEKF provides an exact linearisation of the prediction step of the EKF [16] leading to global convergence and high performance. This observation is also seen in the work by Long *et al.* [17]. Bourmaud *et al.* [18][19] pro-posed the continuous-discrete EKF (CD-EKF) for systems on connected unimodular Lie groups using concentrated Gaussian distribution. Mahony *et al.* [20][21][22] proposed the equivariant filter (EqF), a general Kalman filter design methodology for systems on homogeneous spaces. This design perspective has been widely adapted by modern filters and applied to a range of real-world problems including visual inertial odometry [23][24], inertial navigation systems [25][26], homography tracking [27], etc. The design ap-proach can be interpreted as using local coordinates induced by one of the Cartan-Schouten affine connections on the Lie-group. These are the affine connections for which the geodesics are the one-parameter subgroups, or exponentials, on the Lie-group (or their projections on the homogeneous space). Although this insight draws a strong analogy with the more general research theme discussed earlier, the Lie-group and homogeneous space structure fundamental in this approach (for common real-world problems) carries global matrix algebraic structure that is core to the mathematical formulations and efficiency of the filters.

In this paper, we consider the nonlinear systems where the state space and output space are both smooth manifolds that admit an affine connection and present a general error-state extended Kalman filter design methodology for such

systems. An affine connection is the minimum geometric structure that captures the key properties of geometry. We argue that this is the minimum requirement on manifold structure that is necessary to consider to develop EKF algorithms that are not simply an implementation of the classical Euclidean EKF in local coordinates. With an affine connection, one can define the concepts of geodesic and parallel transport that we show are fundamental in deriving high performance Kalman filter algorithms. We frame the filter development in the context of using concentrated Gaussian distributions [19][22] as local approximations of the information state of the filter. Using this formalisation we can define concepts of mean and covariance for the approximate information state in normal coordinates on the manifold without requiring these concepts to be well defined for the full information state on the manifold. We go on to show that the geometric structure of the manifold should be taken into account explicitly in two places in a Kalman filter update. The first in the Bayesian update step to parallel transport the covariance of the generative noise measurement model into coordinates adapted to the state. Then second in the reset step to parallel transport the state covariance estimate generated by the Bayesian update to the new filter estimate coordinates. There are several prior works that have considered geometric modification to the reset step for EKF algorithms on manifolds [23][28][29][20][22], however, the authors are not aware of any prior work in the EKF algorithms for geometric modifications of the update step. In addition, we believe the geometric insight in the present paper to be novel. We provide simulations to demonstrate the advantage of the proposed geometric modifications during the transient response of the filter. This is the period of the filter evolution when not taking account the geometry of the manifold impacts most on the filter performance. Once the filter has converged to steady-state tracking, the local linearisation error is small and all EKF algorithms have similar performance.

This paper includes seven sections alongside the introduction and conclusion. In Section II notation is defined and preliminary mathematics is discussed. In Section III the system and its stochastic model are defined. In Section IV, we present a conventional error-state EKF construction based on normal coordinates on smooth manifolds. In Section V, we propose novel geometric modifications in the filter dynamics to compensate the coordinate transform during the measurement update step and the error reset step. The example of attitude estimation with directional measurements is presented in Section VI.

II. PRELIMINARIES

A. Manifold and Affine Connection

Let $\mathcal{M}$ be a smooth manifold with dimension m. The tangent space at a point $\xi \in \mathcal{M}$ is denoted $T_\xi \mathcal{M}$. The tangent bundle is denoted $T\mathcal{M}$. Given a differentiable function between smooth manifolds $h : \mathcal{M} \to \mathcal{N}$, its derivative at ξ°

is written as

$$D_{\xi|\xi^\circ} h(\xi) : T_{\xi^\circ} \mathcal{M} \to T_{h(\xi^\circ)} \mathcal{N}.$$

The notation $Dh(\xi) : T\mathcal{M} \to T\mathcal{N}$ denotes the differential of h with an implicit base point.

Let $\mathfrak{X}(\mathcal{M})$ denote the space of differentiable vector fields over $\mathcal{M}$. Let $C^\infty(\mathcal{M})$ denote the class of infinitely differentiable functions on $\mathcal{M}$. An affine connection [30] is an operator $\nabla : \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M})$ written $(X, Y) \mapsto \nabla_X Y$, that satisfies

- (*Linear in X*) $\nabla_{f_1 X_1 + f_2 X_2} Y = f_1 \nabla_{X_1} Y + f_2 \nabla_{X_2} Y$,
- (*Linear in Y*) $\nabla_X (a_1 Y_1 + a_2 Y_2) = a_1 \nabla_X Y_1 + a_2 \nabla_X Y_2$,
- (*Product rule*) $\nabla_X (fY) = f \nabla_X Y + (Xf) Y$,

for all $f_1, f_2 \in C^\infty(\mathcal{M})$, $a_1, a_2 \in \mathbb{R}$ and $X, Y \in \mathfrak{X}(\mathcal{M})$. This gives a notion of directional derivative of a vector field defined on the manifold.

A curve $\gamma : I \to \mathcal{M}$ is called geodesic [30] if

$$\nabla_{\dot{\gamma}(t)} \dot{\gamma}(t) = 0 \tag{1}$$

for any $t \in I$ where I is a maximal open interval in $\mathbb{R}$ containing 0. For any $\hat{\xi} \in \mathcal{M}$ and $v \in T_{\hat{\xi}} \mathcal{M}$, there exists a unique maximal geodesic $\gamma : [0, t(v, \hat{\xi})) \to \mathcal{M}$ that satisfies $\gamma_v(0) = \hat{\xi}$ and $\dot{\gamma}_v(0) = v$ [30, Collorary 4.28]. The exponential mapping $\exp_{\hat{\xi}} : W_{\hat{\xi}} \subset T_{\hat{\xi}} \mathcal{M} \to \mathcal{M}$ is defined as mapping each tangent vector $v \in T_{\hat{\xi}} \mathcal{M}$ to the value of its geodesic at time 1; that is,

$$\exp_{\hat{\xi}}(v) = \gamma_v(1) \tag{2}$$

where $W_{\hat{\xi}}$ is the largest open subset of $T_{\hat{\xi}}$ for which $\exp$ is a diffeomorphism. Let $U_{\hat{\xi}} = \exp_{\hat{\xi}}(W_{\hat{\xi}})$ and note that $U_{\hat{\xi}}$ is open by construction. Let $\iota_{\hat{\xi}} : T_{\hat{\xi}} \mathcal{M} \to \mathbb{R}^m$ provide a linear isomorphism between $T_{\hat{\xi}} \mathcal{M}$ and $\mathbb{R}^m$ for each $\hat{\xi}$. Then the *normal* coordinates on $\mathcal{M}$ are defined by

$$\vartheta_{\hat{\xi}} := \iota_{\hat{\xi}} \circ \exp_{\hat{\xi}}^{-1} : U_{\hat{\xi}} \to \mathbb{R}^m. \tag{3}$$

A vector field is an assignment $\xi \mapsto X_\xi$ for every $\xi \in \mathcal{M}$. A vector field X is *parallel* along γ with respect to the connection ∇ if

$$\nabla_{\dot{\gamma}(t)} X_{\gamma(t)} = 0 \tag{4}$$

for all t. Given any vector $X_{\gamma(0)} \in T_{\gamma(0)} \mathcal{M}$ and smooth curve $\gamma(t)$ then there is a unique family of vectors $X_{\gamma(t)}$ that satisfy (4) and this correspondence induces an invertible linear map $\mathbf{P}_{\gamma(t)} : T_{\gamma(0)} \mathcal{M} \to T_{\gamma(t)} \mathcal{M}$ between tangent spaces by

$$\mathbf{P}_{\gamma(t)} X_{\gamma(0)} := X_{\gamma(t)}$$

for all $X_{\gamma(0)} \in T_{\gamma(0)} \mathcal{M}$. For a fixed time T, the inverse of the parallel transport $\mathbf{P}_{\gamma(T)}$ along a curve $\gamma(t)$ is equal to the parallel transport $\mathbf{P}_{\gamma'(T)}$ along the reversed curve $\gamma'(t) := \gamma(T - t)$. Parallel transport of vector fields induces parallel transport of tensor operators $\Sigma_{\gamma(0)} : T_{\gamma(0)} \mathcal{M} \times T_{\gamma(0)} \mathcal{M} \to \mathbb{R}$ by

$$\mathbf{P}_{\gamma(t)} \Sigma_{\gamma(0)} (X_{\gamma(t)}, Y_{\gamma(t)}) := \Sigma_{\gamma(0)} (\mathbf{P}_{\gamma(t)}^{-1} X_{\gamma(t)}, \mathbf{P}_{\gamma(t)}^{-1} Y_{\gamma(t)}).$$

B. The ⊞/⊟ Operators

Building on the established literature we will use the ⊞ and ⊟ operator notation introduced in [31], [5] to model small $\mathbb{R}^m$ 'perturbations acting on $\mathcal{M}$. Recalling the normal coordinates (3), define the *box plus* $\boxplus : \mathcal{M} \times \mathbb{R}^m \to \mathcal{M}$ and *box minus* $\boxminus : \mathcal{M} \times \mathcal{M} \to \mathbb{R}^m$ operators by

$$\xi \boxplus u = \vartheta_\xi^{-1}(u), \tag{5}$$

$$\zeta \boxminus \xi = \vartheta_\xi(\zeta), \tag{6}$$

for all $\xi \in \mathcal{M}$, $\zeta \in W_\xi$ and $u \in \iota(U_\xi)$. Both the $\boxplus$ and $\boxminus$ operators are associated with geodesic curves on the manifold and in this sense are the natural generalisation of straight lines on Euclidean space. We believe that this is the most natural geometric definition of these operators. Any other definition introduces local coordinates that are not adapted in the natural sense to the geometry of the manifold.

It is straightforward to verify that these proposed operators satisfy

$$\xi \boxplus 0 = \xi, \tag{7a}$$

$$\xi \boxplus (\zeta \boxminus \xi) = \zeta, \tag{7b}$$

$$(\xi \boxplus u) \boxminus \xi = u, \tag{7c}$$

the first three of the four requirements of the original definition proposed in [5, Def. 1]. The fourth axiom in [5, Def. 1] requires that

$$|(\hat{\xi} \boxplus \delta_1) \boxminus (\hat{\xi} \boxplus \delta_2)|^2 \leq |\delta_1 - \delta_2|^2. \tag{8}$$

However, in the normal coordinates (on a Riemannian manifold) one has

$$|(\hat{\xi} \boxplus \delta_1) \boxminus (\hat{\xi} \boxplus \delta_2)|^2 = |\delta_1 - \delta_2|^2 - \frac{1}{3}\text{Ric}(\delta_1, \delta_2) + \mathbb{O}(|\delta|^3),$$

where Ric is the Ricci curvature tensor. For manifolds with non-negative curvature; that is, $\text{Ric} \geq 0$ is positive semi-definite, then (8) holds. For any manifold with negative curvature, however, axiom 4 from [5, Def. 1] will fail locally. This axiom was used in [5] to prove properties of the mean of the true information state on the manifold $\mathcal{M}$ associated with properties of the mean of distributions defined in the $\mathbb{R}^m$ coordinates. In general manifolds the concept of mean and covariance are unclear and instead we will work entirely with concentrated Gaussian distribution approximations (§ III-A) of the true information state. These approximations do have well-defined mean and covariance parametrisation that can be used in the filter algorithm. The results developed in this paper are general and not restricted to manifolds with non-negative curvature.

III. PROBLEM FORMULATION

A. Stochastic model

It is always possible to define a volume measure on a general manifold using a partition of unity construction. The class of probability distributions for the information state considered will be those that are integrable with respect to such a measure. Even with a well-defined concept of probability distribution, concepts such as mean and covariance

are not well defined on a general manifold. There are many works that use geometric structure of the manifold such as a Riemannian metric [3][4], or Lie-group structures [19][32] to define equivalent concepts. These constructions, however, are not necessary for the formulation of Kalman filter algorithms. Rather such algorithms need only a definition of a class of *approximating* distributions that can be parameterised by mean and covariance *parameters*. In particular, it is not necessary that the mean and covariance parameters used as state in the filter correspond to statistics of the true distribution or even of the approximate distribution, only that the distribution generated by the filter parameterisations is close in some sense to the true distribution.

In the remainder of the paper we assume that both the systems state-space $\mathcal{M}$ and the output space $\mathcal{N}$ admit affine connections and that we work with normal coordinates

$$\vartheta_{\hat{\xi}} : \mathcal{M} \to \mathbb{R}^m \tag{9}$$

$$\varphi_{\hat{y}} : \mathcal{N} \to \mathbb{R}^n \tag{10}$$

We approximate a general distribution $p : \mathcal{M} \to \mathbb{R}_+$ around $\hat{\xi} \in \mathcal{M}$ by a *concentrated Gaussian distribution* [32]

$$\mathbf{N}_{\hat{\xi}}(\xi|\mu, \Sigma) := \alpha \exp(-\frac{1}{2}(\vartheta_{\hat{\xi}}(\xi) - \mu)^\top \Sigma^{-1}(\vartheta_{\hat{\xi}}(\xi) - \mu)), \tag{11}$$

where

$$\alpha := \left| \int_{U_{\hat{\xi}}} \exp(-\frac{1}{2}(\vartheta_{\hat{\xi}}(\xi) - \mu)^\top \Sigma^{-1}(\vartheta_{\hat{\xi}}(\xi) - \mu)) \mathrm{d}\xi \right|^{-1}.$$

is a normalizing factor, $\mu \in \mathbb{R}^m$ is a mean vector parameter and $\Sigma \in \mathbb{S}_+(m)$ is a positive-definite symmetric $m \times m$ covariance matrix parameter. Note that the support for the distribution $\mathbf{N}_{\hat{\xi}}(\xi|\mu, \Sigma)$ is contained in the open set $U_{\hat{\xi}} \subset \mathcal{M}$. Within this set, the distribution in local coordinates $x = \vartheta_{\hat{\xi}}(\xi)$ looks like a trimmed Gaussian and the first and second order statistics μ and Σ are well defined. Although the mean and covariance have natural interpretations as *parameters* in the concentrated Gaussian, they do not correspond to the statistical mean and variance of the distribution on $\mathcal{M}$, or even in $\mathbb{R}^m$ coordinates due to the trimmed nature of the distribution. This does not prevent them being used to parameterise the approximate distribution and derive an extended Kalman filter. The filter formulation is now done within the class of concentrated Gaussian distribution and its validity will depend on the validity of the approximation $p(\xi) \approx \mathbf{N}_{\hat{\xi}}(\xi|\mu, \Sigma)$. Clearly for a large range of engineering applications where the *a-posteriori* state distribution is locally Gaussian in the normal coordinates this approximation will work extremely well.

B. System definition

In this work, we consider a nonlinear discrete-time system living on a smooth manifold $\mathcal{M}$. The system function is given by

$$\xi_{k+1} = F(\xi_k, u_{k+1} + \kappa_{k+1}^I) \boxplus \kappa_{k+1}^P,$$
$$\kappa_{k+1}^I \sim \mathbf{N}(0, R_{k+1}^I), \quad \kappa_{k+1}^P \sim \mathbf{N}(0, R_{k+1}^P), \tag{12}$$

where $\xi \in \mathcal{M}$ and $u \in \mathbb{L}$ are the system state and input, respectively. These systems usually admit two types of noise, the input noise κ_{k+1}^I and the processing noise κ_{k+1}^P, modelled as Gaussian processes on the linear input space $\mathbb{L}$ and the tangent space $T_{\xi_{k+1}}\mathcal{M}$. In this work, we combine these noise terms through linearisation,

$$\xi_{k+1} = F(\xi_k, u_{k+1}) \boxplus (\kappa_{k+1}^P + B_{k+1}\kappa_{k+1}^I), \qquad (13)$$

where $B_{k+1} = D_{u|u_{k+1}}F(\xi_k, u)$ is the differential of the system function with respect to the input signal. The resulting simplified noise model is

$$\xi_{k+1} = F(\xi_k, u_{k+1}) \boxplus \kappa_{k+1}, \quad \kappa_{k+1} \sim \mathbf{N}(0, R_{k+1}), \qquad (14)$$

where the total covariance $R_{k+1} = R_{k+1}^P + B_{k+1}R_{k+1}^I B_{k+1}^\top$ captures the combined effect of processing noise and input noise.

The configuration output

$$y_{k+1} = h(\xi_{k+1}) \boxplus v_{k+1}, \quad v_{k+1} \sim \mathbf{N}(0, Q_{k+1}), \qquad (15)$$

is given by a function $h : \mathcal{M} \to \mathcal{N}$, where $\mathcal{N}$ is a smooth manifold termed the *output space*. The disturbance v_{k+1} is modelled as a Gaussian process in the normal coordinate around $h(\xi_{k+1})$ on $\mathcal{N}$.

IV. EXTENDED KALMAN FILTER ON MANIFOLD

In this section, we provide an overview of the EKF design methodology on smooth manifolds and indicate points at which the geometric structure of the manifold is not taken into account in classical treatments. We derive the filter in the error-state formulation, since this allows for a more geometric analysis. The error-state Kalman filter considers propagating the information state of an error ε_k between true ξ_k and nominal $\hat{\xi}$ [33] and reconstructs the filter state $\hat{\xi}$ from the error update.

1) Error state: The *local* error in normal coordinates $\varepsilon_k \in T_{\hat{\xi}_k}\mathcal{M}$ is given by

$$\varepsilon_k = \xi_k \boxminus \hat{\xi}_k, \qquad (16)$$

where $\xi_k, \hat{\xi}_k \in \mathcal{M}$ are the true and estimated system state, respectively. In general, on systems with symmetry the error state is often defined globally using a group structure and the local error state is the local linearisation of this construction [16][21].

2) Stochastic approximation: The information state for the filter at time k, that approximates the true *a-posteriori* distribution of the information state, is a concentrated Gaussian

$$\xi_{k|k} \sim \mathbf{N}_{\hat{\xi}_{k|k}}(\xi|0, \Sigma_{k|k})$$

where $k|k$ indicates the state at time k conditioned on information (inputs and measurements) up to and including time k. This construction corresponds to

$$\xi_{k|k} \boxminus \hat{\xi}_{k|k} \sim \mathbf{N}_{\mathrm{trim}}(0, \Sigma_{k|k})$$

as a random variable in $\mathbb{R}^m$, where $\mathbf{N}_{\mathrm{trim}}$ denotes the trimmed Gaussian associated with the domain of definition of the normal coordinates.

3) Prediction: The reference point $\hat{\xi}_{k|k}$ is updated using the full nonlinear model of the system

$$\hat{\xi}_{k+1|k} = F(\hat{\xi}_{k|k}, u).$$

The predicted error is defined as

$$\varepsilon_{k+1|k} := \xi_{k+1} \boxminus \hat{\xi}_{k+1|k}. \qquad (17)$$

We use the linearisation of the state dynamics to compute an update equation for $\varepsilon_{k+1|k}$.

Lemma 4.1: The linearised dynamics of $\varepsilon_{k+1|k}$ is given by

$$\varepsilon_{k+1|k} = A_{k+1}\varepsilon_{k|k} + \kappa_{k+1} + \mathbb{O}(|\varepsilon_{k|k}, \kappa_{k+1}|^2), \qquad (18)$$

where $\kappa_{k+1} \sim \mathbf{N}(0, R_{k+1})$ and A_{k+1} is given by

$$A_{k+1} := D\vartheta_{\hat{\xi}_{k+1|k}}(\hat{\xi}_{k+1|k}) \cdot DF_{u_{k+1}}(\hat{\xi}_{k|k}) \cdot D\vartheta_{\hat{\xi}_{k|k}}^{-1}(0). \qquad (19)$$

Proof: The predicted error can be written

$$\begin{aligned}
\varepsilon_{k+1|k} &= \xi_{k+1} \boxminus \hat{\xi}_{k+1|k}, \\
&= (F(\xi_k, u_{k+1}) \boxplus \kappa_{k+1}) \boxminus F(\hat{\xi}_k, u_{k+1}), \\
&= (F(\hat{\xi}_k \boxplus \varepsilon_{k|k}, u_{k+1}) \boxplus \kappa_{k+1}) \boxminus F(\hat{\xi}_k, u_{k+1}), \quad (20)
\end{aligned}$$

Discarding $\mathbb{O}(|\varepsilon_{k|k}||\kappa_{k+1}|)$ quadratic terms one has

$$\varepsilon_{k+1|k} = \vartheta_{\hat{\xi}_{k+1|k}}(F_{u_{k+1}}(\vartheta_{\hat{\xi}_{k|k}}^{-1}(\varepsilon_{k|k}))) + \kappa_{k+1}. \qquad (21)$$

The formula for A_{k+1} follows by applying the chain rule of differentiation and evaluating at $\varepsilon_{k|k} = 0$. ∎

The predicted state error is distributed according to a Gaussian with mean zero and covariance [1]

$$\xi_{k+1|k} \sim \mathbf{N}_{\hat{\xi}_{k+1|k}}(\xi|0, \Sigma_{k+1|k}).$$

where

$$\Sigma_{k+1|k} = A_{k+1}\Sigma_{k|k}A_{k+1}^\top + R_{k+1}. \qquad (22)$$

This covariance update (22) hides two changes of coordinates. Firstly, the covariance parameter $\Sigma_{k|k}$ is defined in normal coordinates at $\hat{\xi}_{k|k}$ while $\Sigma_{k+1|k}$ is defined in normal coordinates at $\hat{\xi}_{k+1|k}$. Fortunately, the state update intrinsically captures this change, and $A_{k+1}\Sigma_{k|k}A_{k+1}^\top$ is the correct covariance transformation up to linearisation error. The second change of coordinates is associated with the noise process κ_{k+1} in (20) defined around the propagation of the true state $F(\xi_k, u)$. However, in (22) the covariance R_{k+1} is applied directly in coordinates centered at $\hat{\xi}_{k+1|k}$. This implicit change of coordinates is further discussed in Section V-A.

4) Update: The update step involves Bayesian fusion of the prior $\xi_{k+1} \sim \mathbf{N}_{\hat{\xi}_{k+1|k}}(\xi|0, \Sigma_{k+1|k})$ with a measurement $y_{k+1} \in \mathcal{N}$ associated with a generative noise model $y_{k+1} \sim \mathbf{N}_{h(\xi_{k+1})}(y|0, Q_{k+1})$ (15).

Let $\hat{y}_{k+1|k} = h(\hat{\xi}_{k+1|k})$. Define the innovation as $\tilde{y}_{k+1} := y_{k+1} \boxminus \hat{y}_{k+1|k}$. The estimated measurement $\hat{y}_{k+1}$ is considered as an independent signal and $h(\xi_{k+1})$ is considered a function of the error state $\varepsilon_{k+1|k}$. The update equation uses the linearisation of the innovation with respect to $\varepsilon_{k+1|k}$ at zero.

Lemma 4.2: The linearisation of the innovation is given by

$$\tilde{y}_{k+1} = C_{k+1}\varepsilon_{k+1|k} + v_{k+1} + \mathbb{O}(|\varepsilon_{k+1|k}, v_{k+1}|^2), \quad (23)$$

where $v_{k+1} \sim \mathbf{N}(0, Q_{k+1})$ and C_{k+1}, termed the *output matrix*, is given by

$$C_{k+1} = \mathrm{D}\varphi_{\hat{y}_{k+1}}(\hat{y}_{k+1}) \cdot \mathrm{D}h(\hat{\xi}_{k+1|k}) \cdot \mathrm{D}\vartheta^{-1}_{\hat{\xi}_{k+1|k}}(0). \quad (24)$$

Proof: One has

$$
\begin{aligned}
\tilde{y}_{k+1} &:= y_{k+1} \boxminus \hat{y}_{k+1|k} \\
&= (h(\xi_{k+1}) \boxplus v_{k+1}) \boxminus \hat{y}_{k+1|k} \\
&= (h(\hat{\xi}_{k+1|k} \boxplus (\xi_{k+1} \boxminus \hat{\xi}_{k+1|k})) \boxplus v_{k+1}) \boxminus \hat{y}_{k+1|k} \\
&= (h(\hat{\xi}_{k+1|k} \boxplus \varepsilon_{k+1|k}) \boxplus v_{k+1}) \boxminus \hat{y}_{k+1|k}, \quad (25)
\end{aligned}
$$

Discarding $\mathbb{O}(|\varepsilon_{k+1|k}||v_{k+1}|)$ quadratic terms yields

$$\tilde{y}_{k+1} = \varphi_{h(\hat{\xi}_{k+1|k})}(h(\vartheta^{-1}_{\hat{\xi}_{k+1|k}}\varepsilon_{k+1|k})) + v_{k+1}. \quad (26)$$

Linearising at $\varepsilon_{k+1|k} = 0$ and applying the chain rule yields C_{k+1} ∎

The Kalman update [1] for the *a-posteriori* distribution is given by

$$\xi_{k+1|k+1} \sim \mathbf{N}_{\hat{\xi}_{k+1|k}}(\xi|\mu_{k+1}, \Sigma_{k+1|k+1}) \quad (27)$$

where

$$K_{k+1} = \Sigma_{k+1|k}C^{\top}_{k+1}(C_{k+1}\Sigma_{k+1|k}C^{\top}_{k+1} + Q_{k+1})^{-1}, \quad (28)$$
$$\mu_{k+1} = K_{k+1}\tilde{y}_{k+1}, \quad (29)$$
$$\Sigma_{k+1|k+1} = (I - K_{k+1}C_{k+1})\Sigma_{k+1|k}. \quad (30)$$

Similarly to the prediction step, this update step hides an additional change of coordinate. The generative noise process is defined around the true output $y_{k+1} = h(\xi_{k+1})$, but it is applied around the estimated output $\hat{y}_{k+1} = h(\hat{\xi}_{k+1})$. We discuss how one can address this using parallel transport in Section V-B.

5) Reset: The reset step in an extended Kalman filter is geometrically a change of coordinates. One sets

$$\hat{\xi}_{k+1|k+1} = \hat{\xi}_{k+1|k} \boxplus \mu_{k+1}.$$

The resulting *posteriori* distribution is taken to be

$$\xi_{k+1|k+1} = \mathbf{N}_{\hat{\xi}_{k+1|k+1}}(\xi|0, \Sigma_{k+1|k+1}).$$

However, since the covariance $\Sigma_{k+1|k+1}$ was defined in coordinates centered at $\hat{\xi}_{k+1|k}$ it is clear that this second statement does not follow directly. We address this in Section V-C

V. GEOMETRIC INSIGHT

In this section, we integrate the geometric perspective into the EKF design methodology presented in Section IV, and propose novel geometric modifications in the filter dynamics.

A. Geometric Prediction

In the prediction step (20), linearisation of the error dynamics implicitly relocates the noise process κ_{k+1} from the true state prediction $F(\xi_k, u_{k+1})$ to the estimated state prediction $F(\hat{\xi}_{k|k}, u_{k+1})$. To correct this, the covariance prediction (22) should be modified to

$$\Sigma_{k+1|k} = A_{k+1}\Sigma_{k|k}A^{\top}_{k+1} + R^{+}_{k+1},$$

where R^{+}_{k+1} is a solution to the distribution transformation

$$\mathbf{N}_{F(\xi_k,u)}(\xi|0, R_{k+1}) \approx \mathbf{N}_{F(\hat{\xi}_{k|k},u)}(\xi|F(\xi_k, u) \boxminus F(\hat{\xi}_{k|k}, u), R^{+}_{k+1}).$$

However, solving for R^{+}_{k+1} requires knowledge of $F(\xi_k, u)$, and by extension ξ_k, which is unavailable in practice. We do not address this issue in the present paper.
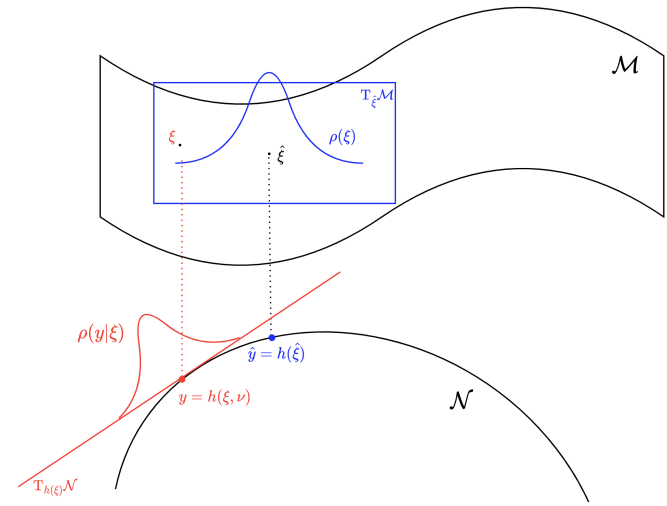
B. Geometric Update

Fig. 1. Demonstration of prior and measurement likelihood in normal coordinates. The prior $p(\xi_{k+1})$ is defined on the normal coordinate around $\hat{\xi}_{k+1|k}$, shown in blue. The measurement likelihood $p(y_{k+1}|\xi_{k+1})$ is defined on the normal coordinate $h(\xi_{k+1})$, shown in red.

As stated in Section IV-.4, the measurement fusion problem is solved by linearising the innovation $y_{k+1} \boxminus \hat{y}_{k+1|k}$ in terms of $\varepsilon_{k+1|k}$ around zero. The covariance $C_{k+1}\Sigma_{k+1|k}C^{\top}_{k+1}$ associated with the filter state estimate correctly captures the change of coordinates from those centered at $\hat{\xi}_{k+1|k}$ to those centered at $\hat{y}_{k+1|k} = h(\hat{\xi}_{k+1|k})$, up to linearisation error. However, the linearisation (25) also implicitly changes the base point of the noise process v_{k+1} from the true output $h(\xi_{k+1})$ to the estimated output $h(\hat{\xi}_{k+1})$, introducing error into the Kalman gain computation (28). To address this, we propose a novel geometric update step in the filter design, which translates the likelihood distribution from $\mathrm{T}_{h(\xi_{k+1})}\mathcal{N}$ to $\mathrm{T}_{\hat{y}_{k+1}}\mathcal{N}$ in a manner compatible with the affine connection of $\mathcal{N}$. Our goal is to approximate

$$y_{k+1} \sim \mathbf{N}_{h(\xi_{k+1})}(y|0, Q) \approx \mathbf{N}_{\hat{y}_{k+1}}(y|h(\xi_{k+1}) \boxminus \hat{y}_{k+1}, Q^{+}), \quad (31)$$

where Q^+ denotes the transformation of the original measurement covariance Q to $T_{\hat{y}_{k+1}}\mathcal{N}$. By modelling Q as a (2,0)-tensor on $T_{h(\xi_{k+1})}\mathcal{N}$, it follows that

$$Q^+ = \mathbf{P}_{\gamma(1)}Q, \tag{32}$$

where $\gamma(t) = h(\xi_{k+1}) \boxplus t(\hat{y}_{k+1|k} \boxminus h(\xi_{k+1}))$ is the geodesic from the true output $h(\xi_{k+1})$ to the estimated output $\hat{y}_{k+1}$. With this new covariance Q^+, the standard filter update step can be applied.

The parallel transport (32) requires knowledge of the true output $h(\xi_{k+1})$ which is not available in practice. Here we propose two choices of approximation.

1) Measurement y_{k+1}: The measurement y_{k+1} can be used as an approximation of $h(\xi_{k+1})$. As shown in Fig 1, the likelihood distribution is centred at $h(\xi_{k+1})$. When the state estimate is poor relative to the measurement, one may take the likelihood to be centred at y_{k+1}, and thus approximate the parallel transport from $h(\xi_{k+1})$ to $\hat{y}_{k+1}$ by the one from y_{k+1} to $\hat{y}_{k+1}$.

2) Naive posteriori: Another alternative is given by the posterior estimated by the original filter update; that is, without geometric update. Following the process in Section IV-.4, one can compute an estimate of the posterior $\hat{\xi}_{k+1|k+1}$. This new estimate can be used to approximate the true output y_{k+1} and repeat the update step to perform the geometric update. This process of estimating a more accurate posterior can be iterated to achieve better performance, at a higher computational cost. Algorithm 1 provides pseudocode for this iterated update.

Remark 5.1: Some existing filter designs, such as the iterated Kalman filter [34] and the iterated EKF on Lie groups [35], use an iterative scheme to generate a more accurate estimation of the output matrix C_{k+1} than can be obtained through a single linearisation step. In contrast, Algorithm 1 uses iteration to better approximate the measurement noise covariance Q_{k+1} in the correct coordinates, while the output matrix is computed in a single step.

C. Geometric Reset

There have been several works on the covariance reset in error-state Kalman filters. It was first mentioned by Markley [36] in the context of multiplicative EKF, recently generalised by Muller *et al.* [28][29]. The authors proposed the covariance reset step using parallel transport in [20][22] for filtering on homogeneous spaces. The same concept can be extended onto a smooth manifold.

At the end of the update step, the posterior distribution (27) is a concentrated Gaussian about the predicted state $\hat{\xi}_{k+1|k}$, but with a non-zero mean μ_{k+1} and updated covariance $\Sigma_{k+1|k+1}$. However, the next filter iteration requires that the state estimate is expressed in coordinate centred at $\hat{\xi}_{k+1|k+1} = \hat{\xi}_{k+1|k} \boxplus \mu_{k+1}$ so that the mean of the distribution is zero. The goal of the reset step is to identify $\Sigma^+_{k+1|k+1}$ such that

$$\xi_{k+1} \sim \mathbf{N}_{\hat{\xi}_{k+1|k}}(\xi|\mu_{k+1}, \Sigma_{k+1|k+1}) \approx \mathbf{N}_{\hat{\xi}_{k+1|k+1}}(\xi|0, \Sigma^+_{k+1|k+1}). \tag{33}$$

Algorithm 1: Iterated geometric update in the proposed EKF

Input: prior $(\hat{\xi}_{k+1|k}, \Sigma_{k+1|k})$, likelihood Q_{k+1}
Input: measurement y_{k+1}
$Q^0_{k+1} \leftarrow Q_{k+1}$,
for `i in range (num of iter)` **do**

$\quad K'_{k+1} = \Sigma_{k+1|k}C^\top_{k+1}(C_{k+1}\Sigma_{k+1|k}C^\top_{k+1} + Q^i_{k+1})^{-1}$,

$\quad \mu'_{k+1} = K'_{k+1}\tilde{y}_{k+1}$,

$\quad \hat{\xi}'_{k+1|k+1} = \hat{\xi}_{k+1|k} \boxplus \mu'_{k+1}$,

$\quad \vec{y}'_{k+1} = h(\hat{\xi}'_{k+1|k+1})$,

$\quad Q^{i+1}_{k+1} = \mathbf{P}_{\gamma(t)}Q_{k+1}$ where $\gamma(t) = \vec{y}'_{k+1} \boxplus t(\hat{y}_{k+1} \boxminus \vec{y}'_{k+1})$,

end
Update:

$\quad K_{k+1} = \Sigma_{k+1|k}C^\top_{k+1}(C_{k+1}\Sigma_{k+1|k}C^\top_{k+1} + Q^{i+1}_{k+1})^{-1}$,

$\quad \mu_{k+1} = K_{k+1}\tilde{y}_{k+1}$,

$\quad \Sigma_{k+1|k+1} = (I - K_{k+1}C_{k+1})\Sigma_{k+1|k}$.

Similarly to the geometric update, this may be solved using parallel transport on $\mathcal{M}$. The covariance Σ is modelled as a (2,0)-tensor on $T\mathcal{M}$. Then the reset covariance $\Sigma^+_{k+1|k+1}$ is found to be

$$\Sigma^+_{k+1|k+1} = \mathbf{P}_{\gamma(1)}\Sigma_{k+1|k+1}, \tag{34}$$

where $\gamma(t)$ is the geodesic curve $\gamma(t) = \hat{\xi}_{k+1|k} \boxplus t\mu_{k+1}$. Unlike the geometric update step, all the information required to solve (34) exactly is available in practice.

VI. SIMULATION

In this section, we use an example of attitude estimation from two directional measurements to demonstrate the algorithm performance. The results of different choices of reference are compared.

A. System Formulation

We consider an attitude estimation problem with gyroscope input and measurements of two known directions. Let $\{G\}$ and $\{I\}$ denote the global reference frame and the body-fixed reference frame, respectively. Let $^G R \in \mathbf{SO}(3)$ denote the rigid body orientation of a moving rigid platform. The onboard gyroscope, which has the same orientation as the platform, returns the bias-free angular velocity $^I\omega \in \mathbb{R}^3$. With non-rotating, flat earth assumption, the deterministic system kinematics is given by

$$^G R_{k+1} = {}^G R_k \exp\left({}^I\omega^\wedge \delta t\right). \tag{35}$$

For the estimation problem, we consider the measurements of two known directions d_1, d_2. The output space is then

defined to be $\mathcal{N} := S^2 \times S^2$. The configuration output is written

$$h({}^G R) = ({}^G R^\top d_1, {}^G R^\top d_2), \qquad (36)$$

where d_1 and d_2 satisfy $d_1 \times d_2 \neq 0$; i.e. the state is always observable.

B. Implementation

We simulate an oscillatory trajectory for attitude estimation (35). The state ${}^G R$ is initialized with identity rotation matrix. The angular velocity input is defined to be ${}^I \omega = (0.1 \times \cos(\tau), 0.1 \times \sin(\tau), 0.1 \times \sin(\tau))$ rad/s. The trajectory is realized using Euler integration at time step $\delta t = 0.02 s$. The estimator has an onboard gyroscope which reads the angular velocity but is corrupted by piecewise constant zero-mean white Gaussian noise with variance 0.02 $(\text{rad/s})^2$ per axis. Additionally, there are sensors that provide measurements of two known directions $d_1 = (0, 1, 0)$ and $d_2 = (1/\sqrt{2}, 0, 1/\sqrt{2})$. The directional measurement (36) are corrupted by Gaussian noise with zero-mean and non-homogeneous covariance $\text{diag}(0.01, 0.03, 0.05)$ rad^2.

We implement the geometric extended Kalman filter described in Section IV, where the manifold $\mathbf{SO}(3)$ is equipped with the Cartan-Schouten 0-connection [37]. Consequently, the normal coordinates are exactly the Lie group exponential coordinates, and the extended Kalman filter is equivalent to the Equivariant filter described in [38]. The parallel transport on the state space $\mathbf{SO}(3)$ is solved using the Cartan-Shouten 0-connection on Lie groups. On the output space $S^2 \times S^2$, we use the canonical affine connection induced by the Cartan-Shouten 0-connection on the Symmetry Lie group $\mathbf{SO}(3)$ [37]. For comparison, we study the proposed filter with and without our proposed geometric update and covariance reset. All filters are initialized by sampling the concentrated Gaussian distribution $\hat{R}_0 \sim \mathbf{N}_{I_3}(0, 1.5^2 I_3)$.

C. Simulation Results

To compare their performance, we plot the error in the attitude estimate as well as the filter energy. The filter energy is $\frac{1}{m} \varepsilon^\top \Sigma^{-1} \varepsilon$ with m being the dimension of the system state and follows a χ^2 distribution. The expected value of filter energy is 1, while a smaller or larger value indicates the filter is under-confident or over-confident in its estimation, respectively.

Figure 2 shows the performance of the filter with and without the proposed geometric modifications. Due to the large initial error in the state estimate, relative to the measurements, the proposed geometric update is seen to make a significant improvement to the convergence of the filter in the transient period. The black trajectory is obtained by using the true output for parallel transport in the geometric update, and is presented only as a reference since this is not possible in practice. It has the best performance over all four implementations. The green and purple trajectories, which are using the naive posterior and measurement for parallel transport, respectively, perform similarly in the transient period, and both outperform the conventional EKF implementation (in
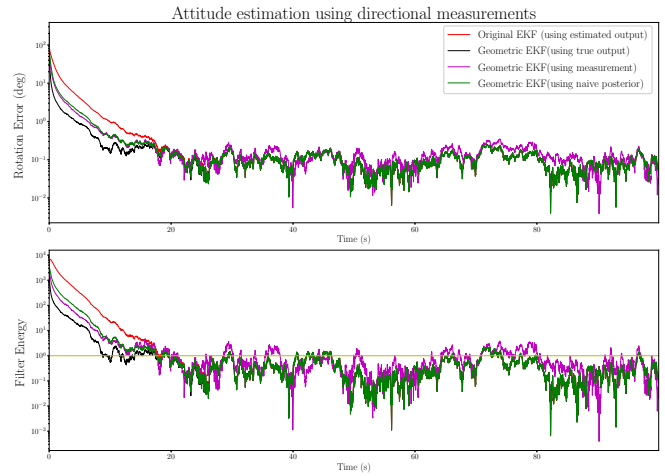

Fig. 2. The rotation estimation error and filter energy are shown for various EKF implementations. The original EKF implementation (—) discussed in Section IV is compared with the EKF with covariance reset and geometric update obtained from the true output (—), the measured output (—), and the naive posterior (—). The yellow horizontal line in the second subplot is for filter energy 1.

red). In terms of the asymptotic behavior, however, the EKF using measurements for parallel transport generally performs worse than all alternatives. The reason is that once the filter is converged, the state estimate tends to be more accurate than the measurement in approximating the true output. In this case, using the measurement as an approximation of the true output will only decrease filter performance.
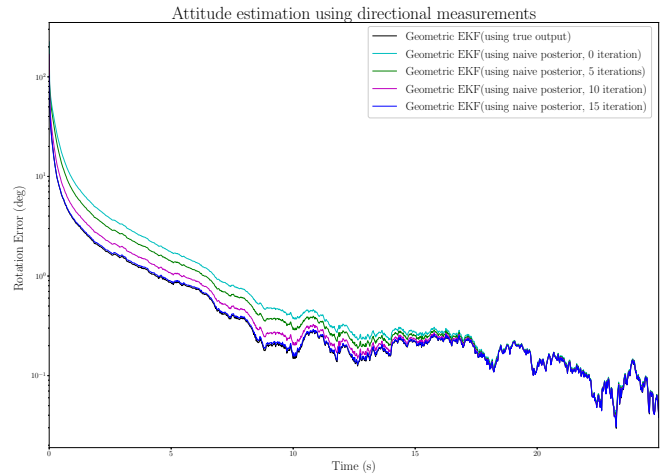

Fig. 3. Comparison of the performance of EKF implementations with geometric update using 0 (—), 5 (—), 10 (—), and 15 (—) iterations. An EKF using the true output for the geometric update (—) is included as a reference for comparison.

Figure 3 shows the results of EKF implementations with iterated geometric updates discussed in Algorithm 1. We implement the EKF with 0, 5, 10, and 15 iterations in the geometric update. As in Figure 2, an EKF where the true output is used for the geometric update is provided as a reference for comparison. One observes that the performance of the EKF with iterated update improves as more iterations

are used, and approaches that of the EKF with true output update. Each iteration requires additional computation and this must be balanced with improved accuracy; however, these results clearly demonstrate the advantage of including the proposed geometric update in an EKF design.

VII. CONCLUSIONS

This paper presents an error-state extended Kalman filter design methodology for smooth manifolds with affine connections. The proposed algorithm includes additional geometric modifications in the filter update and error reset steps by applying parallel transport to the state and measurement covariance matrices. The theory is applied to an example problem of attitude estimation with two directional measurements. The simulation results demonstrate the convergence of the proposed EKF, and show the improvements in performance gained from applying the proposed geometric modifications.

REFERENCES

[1] R. E. Kalman, "A new approach to linear filtering and prediction problems," 1960.

[2] G. L. Smith, S. F. Schmidt, and L. A. McGee, *Application of statistical filter theory to the optimal estimation of position and velocity on board a circumlunar vehicle*. National Aeronautics and Space Administration, 1962.

[3] T. E. Duncan, "Some filtering results in riemann manifolds," *Information and Control*, vol. 35, no. 3, pp. 182–195, 1977.

[4] S. Ng and P. Caines, "Nonlinear filtering in riemannian manifolds," *IMA journal of mathematical control and information*, vol. 2, no. 1, pp. 25–36, 1985.

[5] C. Hertzberg, R. Wagner, U. Frese, and L. Schröder, "Integrating generic sensor fusion algorithms with sound state representations through encapsulation of manifolds," *Information Fusion*, vol. 14, no. 1, pp. 57–77, 2013.

[6] J. Clemens and K. Schill, "Extended kalman filter with manifold state representation for navigating a maneuverable melting probe," in *2016 19Th International Conference On Information Fusion (FUSION)*. IEEE, 2016, pp. 1789–1796.

[7] M. Brossard, A. Barrau, and S. Bonnabel, "A code for unscented kalman filtering on manifolds (ukf-m)," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 5701–5708.

[8] S. Hauberg, F. Lauze, and K. S. Pedersen, "Unscented kalman filtering on riemannian manifolds," *Journal of mathematical imaging and vision*, vol. 46, pp. 103–120, 2013.

[9] S. Said and J. H. Manton, "On filtering with observation in a manifold: Reduction to a classical filtering problem," *SIAM Journal on Control and Optimization*, vol. 51, no. 1, pp. 767–783, 2013.

[10] E. J. Lefferts, F. L. Markley, and M. D. Shuster, "Kalman filtering for spacecraft attitude estimation," *Journal of Guidance, Control, and Dynamics*, vol. 5, no. 5, pp. 417–429, 1982.

[11] F. L. Markley and J. L. Crassidis, *Fundamentals of spacecraft attitude determination and control*. Springer, 2014, vol. 1286.

[12] J. Thienel and R. Sanner, "A coupled nonlinear spacecraft attitude controller and observer with an unknown constant gyro bias and gyro noise," *IEEE Transactions on Automatic Control*, vol. 48, no. 11, pp. 2011–2015, 2003.

[13] R. Mahony, T. Hamel, and J.-M. Pflimlin, "Nonlinear complementary filters on the special orthogonal group," *IEEE Transactions on Automatic Control*, vol. 53, no. 5, pp. 1203–1218, 2008.

[14] S. Bonnabel, "Left-invariant extended kalman filter and attitude estimation," in *2007 46th IEEE Conference on Decision and Control*, 2007, pp. 1027–1032.

[15] S. Bonnable, P. Martin, and E. Salaün, "Invariant extended kalman filter: theory and application to a velocity-aided attitude estimation problem," in *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, 2009, pp. 1297–1304.

[16] A. Barrau and S. Bonnabel, "The invariant extended kalman filter as a stable observer," *IEEE Transactions on Automatic Control*, vol. 62, no. 4, pp. 1797–1812, 2017.

[17] A. W. Long, K. C. Wolfe, M. J. Mashner, G. S. Chirikjian *et al.*, "The banana distribution is gaussian: A localization study with exponential coordinates," *Robotics: Science and Systems VIII*, vol. 265, 2013.

[18] G. Bourmaud, R. Mégret, A. Giremus, and Y. Berthoumieu, "Discrete extended kalman filter on lie groups," in *21st European Signal Processing Conference (EUSIPCO 2013)*. IEEE, 2013, pp. 1–5.

[19] G. Bourmaud, R. Mégret, M. Arnaudon, and A. Giremus, "Continuous-discrete extended kalman filter on matrix lie groups using concentrated gaussian distributions," *Journal of Mathematical Imaging and Vision*, vol. 51, no. 1, pp. 209–228, Jan 2015.

[20] R. Mahony, P. van Goor, and T. Hamel, "Observer design for nonlinear systems with equivariance," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 5, no. 1, p. null, 2022. [Online]. Available: https://doi.org/10.1146/annurev-control-061520-010324

[21] P. van Goor, T. Hamel, and R. Mahony, "Equivariant filter (eqf)," *IEEE Transactions on Automatic Control*, 2022.

[22] Y. Ge, P. van Goor, and R. Mahony, "Equivariant filter design for discrete-time systems," in *2022 IEEE 61st Conference on Decision and Control (CDC)*. IEEE, 2022, pp. 1243–1250.

[23] G. Loianno, M. Watterson, and V. Kumar, "Visual inertial odometry for quadrotors on se (3)," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2016, pp. 1544–1551.

[24] P. van Goor and R. Mahony, "Eqvio: An equivariant filter for visual inertial odometry," *arXiv preprint arXiv:2205.01980*, 2022.

[25] H. A. Hashim, M. Abouheaf, and M. A. Abido, "Geometric stochastic filter with guaranteed performance for autonomous navigation based on imu and feature sensor fusion," *Control Engineering Practice*, vol. 116, p. 104926, 2021.

[26] A. Fornasier, Y. Ng, R. Mahony, and S. Weiss, "Equivariant filter design for inertial navigation systems with input measurement biases," in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 4333–4339.

[27] T. Hamel, R. Mahony, J. Trumpf, P. Morin, and M.-D. Hua, "Homography estimation on the special linear group based on direct point correspondence," in *2011 50th IEEE Conference on Decision and Control and European Control Conference*. IEEE, 2011, pp. 7902–7908.

[28] M. W. Mueller, M. Hehn, and R. D'Andrea, "Covariance correction step for kalman filtering with an attitude," *Journal of Guidance, Control, and Dynamics*, vol. 40, no. 9, pp. 2301–2306, 2017.

[29] R. Gill, M. W. Mueller, and R. D'Andrea, "Full-order solution to the attitude reset problem for kalman filtering of attitudes," *Journal of Guidance, Control, and Dynamics*, vol. 43, no. 7, pp. 1232–1246, 2020.

[30] J. M. Lee, *Introduction to Riemannian manifolds*. Springer, 2018, vol. 2.

[31] A. auf Mannigfaltigkeiten and C. Hertzberg, "A framework for sparse, non-linear least squares problems on manifolds."

[32] Y. Wang and G. S. Chirikjian, "Error propagation on the euclidean group with applications to manipulator kinematics," *IEEE Transactions on Robotics*, vol. 22, no. 4, pp. 591–602, 2006.

[33] J. Sola, "Quaternion kinematics for the error-state kalman filter," *arXiv preprint arXiv:1711.02508*, 2017.

[34] H. W. Sorenson, "Kalman filtering techniques," in *Advances in control systems*. Elsevier, 1966, vol. 3, pp. 219–292.

[35] G. Bourmaud, R. Mégret, A. Giremus, and Y. Berthoumieu, "From intrinsic optimization to iterated extended kalman filtering on lie groups," *Journal of Mathematical Imaging and Vision*, vol. 55, pp. 284–303, 2016.

[36] F. L. Markley, "Attitude error representations for kalman filtering," *Journal of guidance, control, and dynamics*, vol. 26, no. 2, pp. 311–317, 2003.

[37] K. Nomizu, "Invariant affine connections on homogeneous spaces," *American Journal of Mathematics*, vol. 76, no. 1, pp. 33–65, 1954.

[38] A. Fornasier, Y. Ng, C. Brommer, C. Böhm, R. Mahony, and S. Weiss, "Overcoming bias: Equivariant filter design for biased attitude estimation with online calibration," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 12 118–12 125, 2022.