

Article

# Variational Bayesian Iteration-Based Invariant Kalman Filter for Attitude Estimation on Matrix Lie Groups

Jiaolong Wang \*  and Zeyang Chen

Key Laboratory of Advanced Process Control for Light Industry, Institute of Automation, Jiangnan University, Wuxi 214122, China; zeyangchenn@jiangnan.edu.cn

\* Correspondence: jlwang@jiangnan.edu.cn

**Abstract:** Motivated by the rapid progress of aerospace and robotics engineering, the navigation and control systems on matrix Lie groups have been actively studied in recent years. For rigid targets, the attitude estimation problem is a benchmark one with its states defined as rotation matrices on Lie groups. Based on the invariance properties of symmetry groups, the invariant Kalman filter (IKF) has been developed by researchers for matrix Lie group systems; however, the limitation of the IKF is that its estimation performance is prone to be degraded if the given knowledge of the noise statistics is not accurate. For the symmetry Lie group attitude estimation problem, this paper proposes a new variational Bayesian iteration-based adaptive invariant Kalman filter (VBIKF). In the proposed VBIKF, the a priori error covariance is not propagated by the conventional steps but directly calibrated in an iterative manner based on the posterior sequences. The main advantage of the VBIKF is that the statistics parameter of the system process noise is no longer required and so the IKF's hard dependency on accurate process noise statistics can be reduced significantly. The mathematical foundation for the new VBIKF is presented and its superior performance in adaptability and simplicity is further demonstrated by numerical simulations.



**Citation:** Wang, J.; Chen, Z.

Variational Bayesian Iteration-Based Invariant Kalman Filter for Attitude Estimation on Matrix Lie Groups.

*Aerospace* **2021**, *8*, 246. <https://doi.org/10.3390/aerospace8090246>

Academic Editor: Carlo E.D. Riboldi

Received: 22 June 2021

Accepted: 2 September 2021

Published: 3 September 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** attitude estimation; variational Bayesian inference; fixed-point iteration; posterior stochastic sequences; invariant Kalman filter; matrix Lie groups

## 1. Introduction

Attitude determination is a benchmark and important problem in astronautic engineering for the state estimation and control of spacecraft and robotic systems [1–4]. As a widely used attitude estimator, the Kalman filter (KF) is an optimal, numerically efficient, and widely used tool that infers based on all available information, i.e., the dynamic model of the system, sensor data, and the probabilities of both sensor signals and the algorithm's numerical behavior [5,6]. Generally, the state estimation system is usually modeled as vector state space model in Euclidean space and solved with a Kalman filter for vector state variables. In recent years, building navigation and control systems on matrix Lie groups have been actively studied and the properties of the group system provide practical convenience for the definition and simplification of navigation and control system models [7,8]. For engineering applications, the Euclidean system models are usually corrupted by unknown noises or unpredictable disturbances; this is also true for the attitude estimation problems defined on matrix Lie groups and the performance of employed filtering methods is prone to be influenced by inaccurate covariance parameters [6,9,10]. Therefore, for attitude estimation systems on matrix Lie groups, this work aims to study the Kalman filtering problem with unknown noise statistics parameters.

For attitude representation, the widely used quaternions can present a uniform approximation of attitude representation without gimbal-locks [1,9]. The quaternion-based algorithms do not need to compute trigonometric functions and allow for the writing of attitude changes as matrix–vector products. The multiplicative extended Kalman filter is

an ad hoc modification of the extended Kalman Filter that accounts for the constraint of a unit quaternion, but its convergence property is ensured only at equilibrium points and the antipodal quaternions should be identified [11]. Recently, there has been significant research interest in the estimation and control of matrix Lie group systems, including state observers and filters for attitude estimation on  $SO(3)$  and  $SE(3)$  [12,13]. Building attitude kinematics on matrix Lie groups can usually preserve the geometrical property of group systems and the resulting attitude representation does not suffer from the problems of singularities and unwinding [11]. Nevertheless, these studies only generalize the extension to Euclidean Kalman filter theory for attitude estimation on matrix Lie groups and the symmetry properties of attitude estimation models are not sufficiently exploited.

Accounting for the invariance property of Lie groups' attitude kinematics will lead to well-posed problems and boost the performance of attitude estimation algorithms [14]. According to the theory of the invariant Kalman filter (IKF), using the mapping between matrix Lie groups and Lie algebra, attitude estimation systems on matrix Lie groups can be equivalently mapped into a Euclidean vector space and, therefore, the classical Kalman-type estimation methods can be applied to solve the corresponding problems [15,16]. The implementation steps of the IKF are derived based on the invariance and log-linearity of the linear invariant error, which contributes to better filtering convergence on a much bigger set of state trajectories [16]. The necessary probability theory for uncertainty definition on matrix Lie groups has been studied in [17,18] and the concept of a *concentrated Gaussian distribution* has been employed to describe the noisy and random variables on matrix Lie groups [19,20]. Note that the IKF actually obeys the classical Kalman theory and so its estimation performance also heavily depends on the parameters of the noise covariance matrices being accurate; however, in aerospace and satellite engineering, accurate noise statistics parameters are usually not available due to the presence of unpredictable noises and disturbances, which are sure to have a negative influence on the estimation performance of the invariant Kalman filter [21,22].

Motivated by above discussion, it is meaningful to further improve the filtering performance of the invariant Kalman filter for the matrix Lie group attitude estimation problem in the presence of unpredictable noises and disturbances. Note that, for conventional Euclidean space filtering problems, some adaptive techniques have already been developed [23,24] and the most common way is to directly scale or estimate the unknown noise parameters, i.e., the process and observation noise covariance [25]. Recently, variational Bayesian (VB) methods have been applied to the joint estimation of system state and unknown noise parameters [26,27]. A novel adaptive Kalman filter utilizes the VB technique to obtain an approximate inference for inaccurate process and measurement noise covariance [28]. Note that, to the best of our knowledge, for matrix Lie group systems there remains no investigation on adaptive methods for invariant Kalman filtering.

Therefore, for aerospace, satellite, and robotics engineering, this work aims to investigate the benchmark matrix Lie group attitude estimation problem. Note that, in practice, the noise parameter of onboard observation sensors usually depends on the utilized sensing technique and can be determined offline, but it is impractical to accurately evaluate the unpredictable disturbances in the kinematics/dynamics model [23,24]. In this work, the matrix Lie group attitude estimation problem with inaccurate process noise covariance is investigated and a variational Bayesian iteration-based adaptive invariant Kalman filter (VBIKF) is proposed. In the VBIKF, the a priori error covariance is not propagated by the conventional steps but directly calibrated in an iterative manner based on the posterior sequences. The main advantage is that the statistics parameter of the system process noise is no longer required and so the IKF's hard dependency on accurate process noise statistics can be reduced significantly. The mathematical foundation for the new VBIKF is presented and its superior performance in adaptability and simplicity is further demonstrated by numerical simulations.

## 2. Primaries and Problem Definition

This section first presents the essential primaries about matrix Lie groups and the uncertainty definition, which constitute the fundamental theory of the matrix Lie group attitude estimation problem. The attitude estimation problem is described with discussions on the symmetry invariance property of the attitude system. Then, the invariant Kalman filter is introduced and its heavy constraint on the noise parameter is also presented in the problem definition of this work.

### 2.1. Matrix Lie Groups and the Concentrated Gaussian Distribution

In this paper, a matrix Lie group  $G \subset \mathbb{R}^{m \times m}$  is characterized as the set of square, invertible  $m \times m$  matrices satisfying the matrix operations of multiplication and inversion without going outside [29–31].

$$I^d \in G, \forall \chi \in G, \chi^{-1} \in G, \forall \chi_1, \chi_2 \in G, \chi_1 \chi_2 \in G \quad (1)$$

where the  $m \times m$  identity matrix  $I^d$  is the group identity element.

For every matrix Lie group  $G$ , there is an associated Lie algebra  $\mathfrak{g} \subset \mathbb{R}^{m \times m}$  and, from the differential geometric point of view, it is an open neighborhood of  $0_{m \times m}$  in the tangent space  $T_{I^d}G$  of  $G$  at the identity element  $I^d$ . The matrix exponential  $\exp(\cdot)$  and logarithm  $\log(\cdot)$  mappings establish a local diffeomorphism between an open neighborhood of the group identity  $I^d$  in  $G$  and the associated Lie algebra  $\mathfrak{g}$ , i.e.,

$$\exp(\cdot) : \mathfrak{g} \rightarrow G \text{ and } \log(\cdot) : G \rightarrow \mathfrak{g}. \quad (2)$$

As a linear space, the Lie algebra  $\mathfrak{g}$  of  $p$ -dimensional matrix Lie group  $G$  is related to the  $p$ -dimensional Euclidean vector space  $\mathbb{R}^p$  through the natural linear isomorphism

$$(\cdot)^\vee : \mathfrak{g} \rightarrow \mathbb{R}^p \text{ and } (\cdot)^\wedge : \mathbb{R}^p \rightarrow \mathfrak{g}. \quad (3)$$

The uncertainty representation of additive noise in Euclidean vector space cannot be directly applied to matrix Lie group space. The Baker–Campbell–Hausdorff formula [19] provides an approximate to the product of Lie groups in  $\mathbb{R}^p$  [29,30], i.e., for  $a, b \in \mathbb{R}^p$ .

$$\text{BCH}(a, b) = \log(\exp(a^\wedge) \exp(b^\wedge))^\vee = a + b + O(|a, b|^2), \quad (4)$$

where  $O(|a, b|^2)$  is the negligible second- and higher-order terms if  $a$  and  $b$  are small enough. In a *concentrated Gaussian distribution* (CGD) [19,20,32], a random Lie group variable  $\chi \in G$  has a CGD with mean  $\mu \in G$  and covariance  $\Sigma \in \mathbb{R}^{p \times p}$ , i.e.,  $\chi = \chi' \mu \sim \mathcal{G}(\mu, \Sigma)$ , if:

$$\chi' = \exp(\xi^\wedge) \sim \mathcal{G}(I^d, \Sigma), \xi \sim \mathcal{N}^c(0^p, \Sigma) \in \mathbb{R}^p \quad (5)$$

where  $\mathcal{G}(\cdot, \cdot)$  denotes the *concentrated Gaussian distribution* (CGD) and  $\mathcal{N}^c(\cdot, \cdot)$  denotes the classical Gaussian distribution in  $\mathbb{R}^p$ ;  $\chi' \in G$  is a CGD random variable with mean  $I^d$  and covariance  $\Sigma$ ; and  $\xi$  is a zero-mean normally distributed random variable with covariance  $\Sigma$ .

### 2.2. The Attitude Estimation Systems on Special Orthogonal Group $SO(3)$

One motivating example of estimation on matrix Lie groups is the attitude determination on *special orthogonal group*  $SO(3) \subset \mathbb{R}^{3 \times 3}$  for a rigid spacecraft with model [1,14].

$$R_k = Y_k R_{k-1} \Omega_{k-1}, Y_k = \exp((w_k)^\wedge) \quad (6)$$

$$Y_k = \begin{pmatrix} y'_k \\ y''_k \end{pmatrix} = \begin{pmatrix} R_k^T b' + v'_k \\ R_k^T b'' + v''_k \end{pmatrix}, \quad (7)$$

where  $R_k \in G = SO(3)$  denotes the rotation matrix from the body frame to the Earth-fixed frame at time instant  $k$ ;  $\Upsilon_k \sim \mathcal{G}(I^d, \Sigma_k)$  is the left-multiplied random variable on group  $SO(3)$  with  $w_k \sim \mathcal{N}^c(0_{3 \times 1}, \Sigma_w) \in \mathbb{R}^3$  as the corresponding random process noise vector in  $\mathbb{R}^p$ ;  $\Omega_k \in G$  is the rotation control input;  $Y_k \in \mathbb{R}^6$  is a composition of two discrete noisy observations  $y'_k, y''_k \in \mathbb{R}^3$  given the parameter  $b', b'' \in \mathbb{R}^3$ ; and  $v'_k \sim \mathcal{N}^c(0_{3 \times 1}, \Sigma_{v'})$  and  $v''_k \sim \mathcal{N}^c(0^3, \Sigma_{v''})$  are mutually independent isotropic observation noises.

On group  $SO(3)$ , all elements are  $3 \times 3$  real rotation matrices  $R$  satisfying  $RR^T = I_{3 \times 3}$  and determinant  $\det(R) = 1$  [31]. The associated Lie algebra  $\mathfrak{g}$  is actually the space of  $3 \times 3$  real skew-symmetric matrices, i.e.,  $\mathfrak{so}(3) = \{A \in \mathbb{R}^{3 \times 3}, A = -A^T\}$ . Additionally, for  $\zeta = [\zeta_1, \zeta_2, \zeta_3]^T \in \mathbb{R}^3$ , the mapping from the vector to the corresponding skew-symmetric matrix has the form of (8) and its relations to group element  $R$  are (9) and (10) [31].

$$(\zeta)^\wedge = \zeta_\times = \begin{pmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{pmatrix}_\times = \begin{pmatrix} 0 & -\zeta_3 & \zeta_2 \\ \zeta_3 & 0 & -\zeta_1 \\ -\zeta_2 & \zeta_1 & 0 \end{pmatrix} \in \mathfrak{g}, \tag{8}$$

$$R := \exp((\zeta)^\wedge) = \exp(\zeta_\times) = \sum_{m=0}^{\infty} \frac{\zeta_\times^m}{m!} = I_{3 \times 3} + \frac{\sin\|\zeta\|}{\|\zeta\|} \zeta_\times + 2 \frac{\sin(\|\zeta\|/2)^2}{\|\zeta\|^2} \zeta_\times^2, \tag{9}$$

$$\log(R) = \frac{\theta}{2 \sin \theta} (R - R^T) \quad \text{if } \text{tr}(R) \neq -1 \tag{10}$$

where  $\zeta_\times$  denotes the common skew-symmetric operation for vector  $\zeta \in \mathbb{R}^3$ ;  $\|\cdot\|$  represents the standard Euclidean vector norm; and the  $\theta$  in (10) satisfies  $|\theta| < \pi$  and  $1 + 2 \cos \theta = \text{tr}(R)$  with  $\text{tr}(\cdot)$  being the trace of a matrix. For the special case of  $\text{tr}(R) = -1$ , the antipodal solutions to  $\log(R)$  can be determined according to the relation  $R = I_{3 \times 3} + (2/\pi^2)\zeta_\times^2$ .

Let  $\hat{R}_{k-1} \sim \mathcal{G}(R_{k-1}, \Sigma_{k-1})$  be an estimate for the true  $R_{k-1}$  of instant  $k - 1$  and  $\hat{R}_k = \hat{R}_{k-1}\Omega_{k-1}$  be the estimate for the true  $R_k$  satisfying  $\hat{R}_k \sim \mathcal{G}(R_k, \Sigma_k)$ . If we define the multiplied form of *invariant error*  $\zeta_k = (\log(\hat{R}_k R_k^{-1}))^\vee$ , then the motion of  $\zeta_k$  can be obtained as.

$$\begin{aligned} \zeta_k &= (\log(\hat{R}_k R_k^{-1}))^\vee = (\log(\hat{R}_{k-1}\Omega_{k-1}(\Upsilon_k R_{k-1}\Omega_{k-1})^{-1}))^\vee \\ &= (\log(\hat{R}_{k-1} R_{k-1}^{-1} \Upsilon_k^{-1}))^\vee = (\log(\exp(\zeta_{k-1}^\wedge) \exp(-w_k^\wedge)))^\vee \\ &= \text{BCH}(\zeta_{k-1}, -w_k) \doteq \zeta_{k-1} - w_k, \end{aligned} \tag{11}$$

where the last step is a first-order approximation of the BCH formula (4). In the evolution model (11) for  $\zeta_k$ , the true state  $R_{k-1}$  and the control input  $\Omega_{k-1}$  have been cancelled out, which means that the motion of  $\zeta_k$  is independent of the true state trajectory and control input  $\Omega_{k-1}$ , which is called the *trajectory-independent* property of error  $\zeta_k$  [15,16].

### 2.3. The Invariant Kalman Filter for Attitude Estimation

In the invariant Kalman filter (IKF), if given the initial  $\hat{R}_{0|0} \sim \mathcal{G}(R_0, \Sigma_{0|0})$ , the state prediction using the deterministic part of the dynamics (6) is:

$$\hat{R}_{k|k-1} = \hat{R}_{k-1|k-1}\Omega_{k-1}, \tag{12}$$

where  $\hat{R}_{k|k-1}$  and  $\hat{R}_{k-1|k-1}$  denote the prior and posterior error estimate, respectively. Resorting to the equivalence of  $R_k$  and  $\zeta_k$  in the probability distribution, the prior error covariance  $\Sigma_{k|k-1}$  for  $\hat{R}_{k|k-1}$  can be propagated according to the evolution of  $\zeta_k$  in (11).

$$\Sigma_{k|k-1} = \Sigma_{k-1|k-1} + \Sigma_w, \tag{13}$$

where  $\Sigma_{k-1|k-1}$  is the covariance of the posterior estimate  $\hat{\xi}_{k|k-1} = \left(\log(\hat{R}_{k|k-1}R_k^{-1})\right)^\vee$  and  $\hat{R}_{k|k-1}$ .

As to the steps for updating the invariant Kalman filter, the two-vector-based observation  $Y_k \in \mathbb{R}^6$  can be reformulated to define the new innovation vector  $z_k \in \mathbb{R}^6$  [14].

$$z_k = \hat{R}_{k|k-1}Y_k - \begin{pmatrix} b' \\ b'' \end{pmatrix} := H_{\hat{\xi}}\hat{\xi}_{k|k-1} + V_k, \tag{14}$$

where  $H = \begin{pmatrix} b'_\times \\ b''_\times \end{pmatrix}$  and  $V_k = \hat{R}_{k|k-1} \begin{pmatrix} v'_k \\ v''_k \end{pmatrix}$  are respectively the new observation matrix and the noise with  $\Sigma_V = \text{Cov}(V_kV_k^T) = \text{diag}(\Sigma_{v'}, \Sigma_{v''})$ . Then, the steps for the correction of the IKF are:

$$\hat{R}_{k|k} = \exp\left((K_k z_k)^\wedge\right)\hat{R}_{k|k-1}, \tag{15}$$

$$K_k = \Sigma_{k|k-1}H^T / \left(H\Sigma_{k|k-1}H^T + \Sigma_V\right), \tag{16}$$

$$\Sigma_{k|k} = \Sigma_{k|k-1} - K_kH\Sigma_{k|k-1}, \tag{17}$$

where  $K_k$  is the gain matrix correction term, and the vector term  $K_k z_k$  is projected into the matrix Lie group space through the matrix exponential operation in (15) [16,17].

#### 2.4. The Constraint on the Invariant Kalman Filter for Attitude Estimation

In IKF theory, the matrix Lie group system (6) and (7) for rotation  $R_k$  is converted to the Euclidean vector space system (11) and (14) for error  $\xi_k$ ; therefore, according to the equivalence of  $R_k$  and  $\xi_k$  in the probability distribution, the propagation of the covariance and gain parameters in the IKF can mimic that of the classical Kalman filter [14–16].

However, while the IKF resembles the simple and elegant filtering steps of the classical Kalman filter, at the same time it also inherits the hard constraint that the system model and noise statistics parameters should be accurately given; if the noise covariance parameters are not correct, the estimate results are rather prone to being biased. Nevertheless, for some aerospace and astronautics applications, the noise parameter of the observation sensors usually depends on the utilized sensor technique and could be tuned by offline tests; however, it is too difficult to precisely determine the statistics of unpredictable disturbances that are often influenced by the operating environment of different target missions [1,21,23,24].

Note that, for conventional Euclidean space filtering problems, some adaptive techniques have already been developed but, to the best of our knowledge, for matrix Lie group systems there remains no investigation on adaptive methods for invariant Kalman filtering. Therefore, this work focuses on adaptive invariant Kalman filtering methods for matrix Lie group attitude estimation problems without an accurate process noise covariance  $\Sigma_w$  and attempts to remove the IKF’s hard constraint on an a priori and accurate  $\Sigma_w$ .

### 3. Variational Iteration-Based Invariant Kalman Filter for Attitude Estimation

In the invariant Kalman filter, given the prior error covariance matrix  $\Sigma_{k|k-1}$ , the predicted probability density function (PDF)  $p(R_k|Y_{1:k-1})$  and the likelihood probability density function  $p(Y_k|R_k)$  are assumed to be a *concentrated Gaussian distribution*, i.e.,

$$p\left(R_k|Y_{1:k-1}, \Sigma_{k|k-1}\right) = \mathcal{G}\left(\hat{R}_{k|k-1}, \Sigma_{k|k-1}\right), \tag{18}$$

where the prior estimate  $\hat{R}_{k|k-1}$  is obtained with (12) and the prior error covariance  $\Sigma_{k|k-1}$  is propagated through the calculus (13) with  $\Sigma_w$  required. Obviously, if the true value of the process noise covariance  $\Sigma_w$  cannot be accurately determined in advance, the  $\Sigma_{k|k-1}$  is sure to be misled by an incorrect  $\Sigma_w$ , which will in turn degrade the estimation results of the IKF.

### 3.1. Distribution Definition for the Prior Error Covariance

To deal with the trouble caused by an inaccurate  $\Sigma_w$ , in this work we choose to infer the rotation group state  $\hat{R}_{k|k-1}$  together with the prior error covariance  $\Sigma_{k|k-1}$ . The basic idea is that, before calculating the required PDF  $p(R_k | Y_{1:k-1}, \Sigma_{k|k-1})$ , the probability density function of the prior error covariance  $\Sigma_{k|k-1}$  should first be calculated based on all the historical observation sequences, i.e., by calculating the probability density function  $p(\Sigma_{k|k-1} | Y_{1:k-1})$ .

As to the definition of the conjugate prior distribution PDF  $p(\Sigma_{k|k-1} | Y_{1:k-1})$  for  $\Sigma_{k|k-1}$ , in this work we define it as the inverse Winshart distribution, which has been commonly employed as the conjugate for the prior covariance matrix of Gaussian distributions with known mean [28]. The inverse Winshart probability density function of a symmetric positive definite  $d \times d$  random matrix  $B \in \mathbb{R}^{d \times d}$  is formulated as:

$$\text{IW}(B; \lambda, \Psi) = \frac{|\Psi|^{\lambda/2} |B|^{-(\lambda+d-1)/2} \exp\{-0.5 \text{tr}(\Psi B^{-1})\}}{2^{d\lambda/2} \Gamma_d(\lambda/2)}, \quad (19)$$

where IW is the inverse Winshart distribution;  $\lambda$  denotes the degrees of freedom (dof) for the inverse Winshart distribution;  $\Psi$  denotes the inverse scale matrix, which should be a  $d \times d$  symmetric and positive definite matrix;  $|\cdot|$  is the matrix determinant operation; and  $\Gamma_d(\cdot)$  denotes the  $d$ -variate gamma function [28]. For covariance matrix  $B$  with an inverse Winshart distribution, i.e.,  $B \sim \text{IW}(B; \lambda, \Psi)$ , the expectation of the matrix inverse  $E(B^{-1})$  is:

$$E(B^{-1}) = (\lambda - d - 1)\Psi^{-1}, \text{ if } \lambda > d + 1. \quad (20)$$

Therefore, to infer the prior error covariance  $\Sigma_{k|k-1}$ , the following inverse Winshart distribution is employed in this work to describe the distribution of  $\Sigma_{k|k-1}$ .

$$p(\Sigma_{k|k-1} | Y_{1:k-1}) = \text{IW}(\Sigma_{k|k-1}; \hat{\lambda}_{k|k-1}, \hat{\Psi}_{k|k-1}) = \frac{|\hat{\Psi}_{k|k-1}|^{\lambda/2} |\Sigma_{k|k-1}|^{-(\hat{\lambda}_{k|k-1}+d-1)/2} \exp\{-0.5 \text{tr}(\hat{\Psi}_{k|k-1} \Sigma_{k|k-1}^{-1})\}}{2^{d\hat{\lambda}_{k|k-1}/2} \Gamma_d(\hat{\lambda}_{k|k-1}/2)}, \quad (21)$$

where  $\hat{\lambda}_{k|k-1}$ ,  $\hat{\Psi}_{k|k-1}$  are respectively the degree of freedom parameter and the inverse scale matrix parameter for the PDF  $p(\Sigma_{k|k-1} | Y_{1:k-1})$  that needs to be determined.

### 3.2. Variational Bayesian Approximations of Posterior PDF

To infer the rotation group state  $\hat{R}_{k|k-1}$  together with the prior error covariance  $\Sigma_{k|k-1}$ , the joint posterior probability density function  $p(R_k, \Sigma_{k|k-1} | Y_{1:k})$  should be calculated. In this work, we use the variational Bayesian method to obtain an approximation factored from [33], i.e.,

$$p(R_k, \Sigma_{k|k-1} | Y_{1:k}) \approx q(R_k)q(\Sigma_{k|k-1}), \quad (22)$$

where  $q(\cdot)$  denotes the approximate posterior probability density function  $p(\cdot)$ ; and  $q(R_k)$  and  $q(\Sigma_{k|k-1})$  can be determined by minimizing the Kullback–Leibler divergence (KLD) [28,33] between  $q(R_k)q(\Sigma_{k|k-1})$  and  $p(R_k, \Sigma_{k|k-1} | Y_{1:k})$ , i.e.,

$$\{q(R_k), q(\Sigma_{k|k-1})\} = \text{argminKLD}(q(R_k)q(\Sigma_{k|k-1}) || p(R_k, \Sigma_{k|k-1} | Y_{1:k})), \quad (23)$$

where  $\text{KLD}(q(x) || p(x)) \triangleq \int q(x) \log \frac{q(x)}{p(x)} dx$  denotes the KLD between  $q(x)$  and  $p(x)$ .

The optimal solution to (23) can be obtained based on the following equations:

$$\log q(R_k) = E_{\Sigma_{k|k-1}} [\log p(R_k, \Sigma_{k|k-1}, Y_{1:k})] + c_{R_k}, \quad (24)$$

$$\log q\left(\Sigma_{k|k-1}\right) = E_{R_k}\left[\log p\left(R_k, \Sigma_{k|k-1}, Y_{1:k}\right)\right] + c_{\Sigma_{k|k-1}}, \tag{25}$$

where  $E_{\Sigma_{k|k-1}}[\cdot]$  denotes the mathematical expectation only for the variable  $\Sigma_{k|k-1}$  and  $E_{R_k}[\cdot]$  denotes the mathematical expectation only for the variable  $R_k$ ; and  $c_{R_k}, c_{\Sigma_{k|k-1}}$  are the constants with respect to  $R_k$  and  $\Sigma_{k|k-1}$ , respectively. Note that fixed-point iterations are needed in order to solve above two coupled equations; the approximation to the posterior PDF  $q(R_k)$  is updated as  $q^{(i+1)}(R_k)$  at the  $i+1$ th iteration using the approximate PDF  $q^i(\Sigma_{k|k-1})$ , while the approximate  $q(\Sigma_{k|k-1})$  is updated as  $q^{(i+1)}(\Sigma_{k|k-1})$  at the  $i+1$ th iteration using the  $q^i(R_k)$  [33].

The joint PDF  $p\left(R_k, \Sigma_{k|k-1}, Y_{1:k}\right)$  can be factored based on the conditional independence properties as follows:

$$\begin{aligned} p\left(R_k, \Sigma_{k|k-1}, Y_{1:k}\right) &= p\left(Y_k|R_k\right)p\left(R_k|Y_{1:k-1}, \Sigma_{k|k-1}\right)p\left(\Sigma_{k|k-1}|Y_{1:k-1}\right)p\left(Y_{1:k-1}\right) \\ &= p\left(Y_k|R_k\right)\mathcal{G}\left(\hat{R}_{k|k-1}, \Sigma_{k|k-1}\right)\text{IW}\left(\Sigma_{k|k-1}; \hat{\lambda}_{k|k-1}, \hat{\Psi}_{k|k-1}\right)p\left(Y_{1:k-1}\right) \end{aligned} \tag{26}$$

Note that the PDF  $p\left(Y_k|R_k\right)$  is actually the distribution of the noisy observation  $Y_k$  conditioned on the true rotation state  $R_k$  and, according to (7) and (14), we have:

$$p\left(Y_k|R_k\right) = p\left(R_k^T Y_k - \begin{pmatrix} b' \\ b'' \end{pmatrix} \middle| R_k\right) = p\left(z_k|\zeta_k\right). \tag{27}$$

Besides,  $\mathcal{G}\left(\hat{R}_{k|k-1}, \Sigma_{k|k-1}\right)$  is actually the prior distribution of the rotation group state  $R_k$ ; then, according to the equivalence of  $R_k$  and  $\zeta_k$  in the probability distribution, we have:

$$\begin{aligned} \log p\left(R_k, \Sigma_{k|k-1}, Y_{1:k}\right) &= \log\left(p\left(z_k|\zeta_k\right)\mathcal{N}^c\left(\hat{\zeta}_{k|k-1}, \Sigma_{k|k-1}\right)\text{IW}\left(\Sigma_{k|k-1}; \hat{\lambda}_{k|k-1}, \hat{\Psi}_{k|k-1}\right)p\left(Y_{1:k-1}\right)\right) \\ &= -0.5\left(z_k - H\hat{\zeta}_k\right)^T \Sigma_V^{-1}\left(z_k - H\hat{\zeta}_k\right) - 0.5\left(d + \hat{\lambda}_{k|k-1} + 2\right) \log\left|\Sigma_{k|k-1}\right| \\ &\quad - 0.5\left(\hat{\zeta}_k - \hat{\zeta}_{k|k-1}\right)^T \Sigma_{k|k-1}^{-1}\left(\hat{\zeta}_k - \hat{\zeta}_{k|k-1}\right) - 0.5\text{tr}\left(\hat{\Psi}_{k|k-1} \Sigma_{k|k-1}^{-1}\right) + \text{constant}. \end{aligned} \tag{28}$$

Then, using (24) and (25) in (28), we have:

$$\log q^{(i+1)}\left(\Sigma_{k|k-1}\right) = -0.5\left(d + \hat{\lambda}_{k|k-1} + 2\right) \log\left|\Sigma_{k|k-1}\right| - 0.5\text{tr}\left(\left(\Pi_k^{(i)} + \hat{\Psi}_{k|k-1}\right)\Sigma_{k|k-1}^{-1}\right) + c_{\Sigma}, \tag{29}$$

$$\log q^{(i+1)}\left(R_k\right) = -0.5\left(z_k - H\hat{\zeta}_k\right)^T \Sigma_V^{-1}\left(z_k - H\hat{\zeta}_k\right) - 0.5\left(\hat{\zeta}_k - \hat{\zeta}_{k|k-1}\right)^T E^{(i+1)}\left[\Sigma_{k|k-1}^{-1}\right]\left(\hat{\zeta}_k - \hat{\zeta}_{k|k-1}\right) + c_{R}. \tag{30}$$

where  $q^{(i+1)}(\cdot)$  is the iterative approximation of PDF  $q(\cdot)$  at the  $i+1$ th iteration;  $E^{(i+1)}[\cdot]$  represents the mathematical expectation at the  $i$ th iteration; and  $\Pi_k^{(i)}$  is given by:

$$\begin{aligned} \Pi_k^{(i)} &= E^{(i)}\left[\left(\zeta_k - \hat{\zeta}_{k|k-1}\right)\left(\zeta_k - \hat{\zeta}_{k|k-1}\right)^T\right] \\ &= E^{(i)}\left[\left(\zeta_k - \hat{\zeta}_{k|k}^{(i)} + \hat{\zeta}_{k|k}^{(i)} - \hat{\zeta}_{k|k-1}\right) \times \left(\zeta_k - \hat{\zeta}_{k|k}^{(i)} + \hat{\zeta}_{k|k}^{(i)} - \hat{\zeta}_{k|k-1}\right)^T\right] \\ &= E^{(i)}\left[\left(\zeta_k - \hat{\zeta}_{k|k}^{(i)}\right)\left(\zeta_k - \hat{\zeta}_{k|k}^{(i)}\right)^T\right] + \left(\hat{\zeta}_{k|k}^{(i)} - \hat{\zeta}_{k|k-1}\right)\left(\hat{\zeta}_{k|k}^{(i)} - \hat{\zeta}_{k|k-1}\right)^T \\ &= \Sigma_{k|k}^{(i)} + \left(\hat{\zeta}_{k|k}^{(i)} - \hat{\zeta}_{k|k-1}\right)\left(\hat{\zeta}_{k|k}^{(i)} - \hat{\zeta}_{k|k-1}\right)^T \\ &= \Sigma_{k|k}^{(i)} + \Delta_{k|k}^{(i)}\left(\Delta_{k|k}^{(i)}\right)^T, \end{aligned} \tag{31}$$

where  $\Delta_{k|k}^{(i)} = \hat{\xi}_{k|k}^{(i)} - \hat{\xi}_{k|k-1}$  is the correction term at the  $i+1$ th iteration and will be calculated later. Note that, since the PDF of  $\Sigma_{k|k-1}$  is considered to be an inverse Wishart distribution, the updated PDF  $q^{(i+1)}(\Sigma_{k|k-1})$  in (29) with the updated  $\hat{\lambda}_k^{(i+1)}$  and  $\hat{\Psi}_k^{(i+1)}$  can be written as:

$$q^{(i+1)}(\Sigma_{k|k-1}) = IW(\Sigma_{k|k-1}; \hat{\lambda}_k^{(i+1)}, \hat{\Psi}_k^{(i+1)}), \tag{32}$$

$$\hat{\lambda}_k^{(i+1)} = \hat{\lambda}_{k|k-1} + 1, \tag{33}$$

$$\hat{\Psi}_k^{(i+1)} = \Pi_k^{(i)} + \hat{\Psi}_{k|k-1}. \tag{34}$$

Then, according to (20), the  $E^{(i+1)}[\Sigma_{k|k-1}^{-1}]$  in (30) can be calculated as:

$$E^{(i+1)}[\Sigma_{k|k-1}^{-1}] = (\hat{\lambda}_k^{(i+1)} - d - 1) (\hat{\Psi}_k^{(i+1)})^{-1}. \tag{35}$$

### 3.3. The Variational Bayesian Iteration-Based Invariant Kalman Filter

Define the propagated  $p^{(i+1)}(\xi_k | z_{1:k-1})$  at the  $i+1$ th iteration as:

$$p^{(i+1)}(\xi_k | z_{1:k-1}) = \mathcal{N}^c(\hat{\xi}_{k|k-1}, \Sigma_{k|k-1}^{(i+1)}), \tag{36}$$

$$\Sigma_{k|k-1}^{(i+1)} = \{E^{(i+1)}[\Sigma_{k|k-1}^{-1}]\}^{-1}, \tag{37}$$

then the propagated PDF  $p^{(i+1)}(R_k | Y_{1:k-1})$  at the  $i+1$ th iteration can be written as:

$$p^{(i+1)}(R_k | Y_{1:k-1}) = \mathcal{G}(\hat{R}_{k|k-1}, \Sigma_{k|k-1}^{(i+1)}), \tag{38}$$

Then, using (27) and (36)~(38) in (30) yields:

$$q^{(i+1)}(\xi_k) = \frac{p(z_k | \xi_k) p^{(i+1)}(\xi_k | z_{1:k-1})}{\int p(z_k | \xi_k) p^{(i+1)}(\xi_k | z_{1:k-1}) d\xi_k} \tag{39}$$

$$q^{(i+1)}(R_k) = \frac{p(Y_k | R_k) p^{(i+1)}(R_k | Y_{1:k-1})}{\int p(Y_k | R_k) p^{(i+1)}(R_k | Y_{1:k-1}) dR_k} \tag{40}$$

According to the above equations, the  $q^{(i+1)}(\xi_k)$  can be updated as a Gaussian PDF with the mean being  $\hat{\xi}_{k|k}^{(i+1)}$  and the covariance being  $\Sigma_{k|k}^{(i+1)}$ , i.e.,

$$q^{(i+1)}(\xi_k) = \mathcal{N}^c(\hat{\xi}_{k|k}^{(i+1)}, \Sigma_{k|k}^{(i+1)}), \tag{41}$$

and the propagated PDF  $q^{(i+1)}(R_k)$  at the  $i+1$ th iteration can be updated as the *concentrated Gaussian distribution* with the mean being  $\hat{R}_{k|k}^{(i+1)}$  and the covariance being  $\Sigma_{k|k}^{(i+1)}$ , i.e.,

$$q^{(i+1)}(R_k) = \mathcal{G}(\hat{R}_{k|k}^{(i+1)}, \Sigma_{k|k}^{(i+1)}), \tag{42}$$

where the mean  $\hat{R}_{k|k}^{(i+1)}$  and the covariance  $\Sigma_{k|k}^{(i+1)}$  at the  $i+1$ th iteration are calculated as:

$$\hat{R}_{k|k}^{(i+1)} = \exp\left(\left(K_k^{(i+1)} z_k\right)^\wedge\right) \hat{R}_{k|k-1}, \tag{43}$$



$$K_k^{(i+1)} = \Sigma_{k|k-1}^{(i+1)} H^T \left( H \Sigma_{k|k-1}^{(i+1)} H^T + \Sigma_V \right)^{-1}, \tag{44}$$

$$\Sigma_{k|k}^{(i+1)} = \Sigma_{k|k-1}^{(i+1)} - K_k^{(i+1)} H \Sigma_{k|k-1}^{(i+1)}, \tag{45}$$

Additionally, according to the definition of *invariant error*  $\xi_k = \left( \log(\hat{R}_k R_k^{-1}) \right)^\vee$ , the correction term  $\Delta_{k|k}^{(i)}$  that is required in (31) can be calculated as follows:

$$\begin{aligned} \Delta_{k|k}^{(i)} &= \hat{\xi}_{k|k}^{(i)} - \hat{\xi}_{k|k-1} = \left( \log(\hat{R}_{k|k}^{(i)} R_k^{-1}) \right)^\vee - \left( \log(\hat{R}_{k|k-1} R_k^{-1}) \right)^\vee \\ &= \left( \log(\hat{R}_{k|k}^{(i)} R_k^{-1}) - \log(\hat{R}_{k|k-1} R_k^{-1}) \right)^\vee \\ &= \left( \log(\hat{R}_{k|k}^{(i)} \hat{R}_{k|k-1}^{-1}) \right)^\vee \\ &= K_k^{(i+1)} z_k \end{aligned} \tag{46}$$

Then, for  $N$  iterations, the variational Bayesian approximation of posterior PDFs is:

$$q(R_k) \approx q^{(N)}(R_k) = \mathcal{G}(\hat{R}_{k|k}^{(N)}, \Sigma_{k|k}^{(N)}) = \mathcal{G}(\hat{R}_{k|k}, \Sigma_{k|k}), \tag{47}$$

$$q(\Sigma_{k|k-1}) \approx q^{(N)}(\Sigma_{k|k-1}) = \text{IW}(\Sigma_{k|k-1}; \hat{\lambda}_k^{(N)}, \hat{\Psi}_k^{(N)}) = \text{IW}(\Sigma_{k|k-1}; \hat{\lambda}_{k|k}, \hat{\Psi}_{k|k}). \tag{48}$$

Therefore, the filtering steps of proposed approach include the prediction step (12), the initialization of the  $\hat{\lambda}_{k|k-1}$  and  $\hat{\Psi}_{k|k-1}$  for inverse Wishart distribution, and the variational Bayesian iteration steps of (31)~(35), (37), and (43)~(48). The details of the implementation of the proposed method for attitude estimation is presented in Algorithm 1.

---

**Algorithm 1.** The filtering steps of one time instant in the proposed approach to attitude estimation.

---

**Inputs:**  $\hat{R}_{k-1|k-1}, \tilde{\Sigma}_{k|k-1} = \Sigma_{k-1|k-2}, \Omega_{k-1}, H, Y_k, \Sigma_w, d = 3, b', b''$

**Time update:**

- 1:  $\hat{R}_{k|k-1} = \hat{R}_{k-1|k-1} \Omega_{k-1}$
- 2:  $z_k = \hat{R}_{k|k-1} Y_k - \begin{pmatrix} b' \\ b'' \end{pmatrix}$

**Measurement update:**

- 3: Initialization:  $\hat{R}_{k|k}^{(0)} = \hat{R}_{k|k-1}, \Sigma_{k|k}^{(0)} = \tilde{\Sigma}_{k|k-1}, \hat{\Psi}_{k|k-1} = k \tilde{\Sigma}_{k|k-1}, \hat{\lambda}_{k|k-1} = k + d + 1, \Delta_{k|k}^{(0)} = \mathbf{0}_{n \times 1}$
- for**  $i$  from 0 to  $N-1$

- update  $q^{(i+1)}(\Sigma_{k|k-1}) = \text{IW}(\Sigma_{k|k-1}; \hat{\lambda}_k^{(i+1)}, \hat{\Psi}_k^{(i+1)})$  given  $q^{(i)}(R_k)$ :
- 4:  $\Pi_k^{(i)} = \Sigma_{k|k}^{(i)} + \Delta_{k|k}^{(i)} \left( \Delta_{k|k}^{(i)} \right)^T, \hat{\lambda}_k^{(i+1)} = \hat{\lambda}_{k|k-1} + 1, \hat{\Psi}_k^{(i+1)} = \Pi_k^{(i)} + \hat{\Psi}_{k|k-1}$   
 update  $q^{(i+1)}(R_k) = \mathcal{G}(\hat{R}_{k|k}^{(i+1)}, \Sigma_{k|k}^{(i+1)})$  given  $q^{(i+1)}(\Sigma_{k|k-1})$ :
- 5:  $E^{(i+1)}[\Sigma_{k|k-1}^{-1}] = \left( \hat{\lambda}_k^{(i+1)} - d - 1 \right) \left( \hat{\Psi}_k^{(i+1)} \right)^{-1}, \Sigma_{k|k-1}^{(i+1)} = \left\{ E^{(i+1)}[\Sigma_{k|k-1}^{-1}] \right\}^{-1}$
- 6:  $K_k^{(i+1)} = \Sigma_{k|k-1}^{(i+1)} H^T \left( H \Sigma_{k|k-1}^{(i+1)} H^T + \Sigma_V \right)^{-1}$
- 7:  $\hat{R}_{k|k}^{(i+1)} = \exp\left( \left( \Delta_{k|k}^{(i)} \right)^\wedge \right) \hat{R}_{k|k-1}, \Delta_{k|k}^{(i)} = K_k^{(i+1)} z_k$
- 8:  $\Sigma_{k|k}^{(i+1)} = \Sigma_{k|k-1}^{(i+1)} - K_k^{(i+1)} H \Sigma_{k|k-1}^{(i+1)}$

**end for**

- 9:  $\hat{R}_{k|k} = \hat{R}_{k|k}^{(N)}, \Sigma_{k|k} = \Sigma_{k|k}^{(N)}, \Sigma_{k|k-1} = \Sigma_{k|k-1}^{(N)}, \hat{\lambda}_{k|k} = \hat{\lambda}_k^{(N)}, \hat{\Psi}_{k|k} = \hat{\Psi}_k^{(N)}$

**Outputs:**  $\hat{R}_{k|k}, \Sigma_{k|k-1}, \Sigma_{k|k}, \hat{\Psi}_{k|k}$

---

### 3.4. Parameter Selection for the Proposed Approach to Attitude Estimation on SO(3)

For the attitude estimation problem, the matrix Lie group system (6) and (7) for rotation  $R_k$  is converted to the Euclidean space system (11) and (14) for error  $\zeta_k$ ; therefore, according to the equivalence of  $R_k$  and  $\zeta_k$  in the probability distribution, the propagation of the filtering parameters for rotation  $R_k$  mimics that of the classical Kalman filter for error  $\zeta_k$ . Note that the projected Euclidean space system (3) and (7) for error  $\zeta_k$  is actually a linear time-invariant system, for which the optimal Kalman filtering parameters, including the prior error covariance  $\Sigma_{k|k-1}$  and the Kalman gain, would converge to constants [14,22,24]. Similarly, for attitude estimation the parameters  $\Sigma_{k|k-1}$  and the Kalman gain of the invariant Kalman filter will also converge to their optimal values, which has been validated in [14] and can be used to help simplify the parameter selection for the proposed approach. Therefore, to initialize the prior error covariance  $\Sigma_{k|k-1}$ , the covariance  $\Sigma_{k-1|k-2}$  of the last time instant can be used as the initial estimate  $\tilde{\Sigma}_{k|k-1} = \Sigma_{k-1|k-2}$  before the iteration so that the negative influence caused by the inaccurate covariance  $\Sigma_w$  can be considerably reduced.

In the conventional invariant Kalman filter, the prior error estimate  $\hat{R}_{k|k-1}$  is actually based on all the historical observations  $Y_{1:k-1}$  before time instant  $k$ . In this work, the covariance parameter  $\Sigma_{k|k-1}$  is also inferred using the historical observations  $Y_{1:k-1}$  until time instant  $k$ . As to the inverse Wishart distribution of  $\Sigma_{k|k-1}$ , at time instant  $k$  the dof parameter is set as  $\hat{\lambda}_{k|k-1} = k + d + 1$  and the inverse scale matrix parameter is set as  $\hat{\Psi}_{k|k-1} = k\tilde{\Sigma}_{k|k-1}$  because the following conditions should be satisfied according to (19) and (20):

$$\hat{\lambda}_{k|k-1} > d + 1, \quad d = p = 3 \quad (49)$$

$$\mathbb{E}[\tilde{\Sigma}_{k|k-1}^{-1}] = (\hat{\lambda}_{k|k-1} - d - 1)\hat{\Psi}_{k|k-1}^{-1}. \quad (50)$$

The number of iterations  $N$  is also a crucial parameter for the proposed approach and, generally, it should be set to a value larger than  $d$  to guarantee the convergence of variational Bayesian iterations; nevertheless, a value that is too large is sure to increase the computational cost of the algorithm's implementation and so a balance between precision and cost should be considered according to the particular application.

Note that, for invariant Kalman filtering without an accurate process noise covariance, it is impossible to directly obtain the optimal estimates. In the related work [21–27], some suboptimal approximations and assumptions were used to asymptotically approach the optimal results. In this sense, the proposed variational iteration-based invariant Kalman filter is actually a suboptimal approach for the following reasons:

- (1) In the parameter setting shown in Algorithm 1, the initialization of the parameter  $\tilde{\Sigma}_{k|k-1}$  at the  $k$ th time instant is based on the estimate  $\Sigma_{k-1|k-2}$  of the last time instant, i.e.,  $\tilde{\Sigma}_{k|k-1} = \Sigma_{k-1|k-2}$ ; the advantage of this setting is that usage of an inaccurate  $\Sigma_w$  can be avoided, but the validity of the parameter  $\tilde{\Sigma}_{k|k-1} = \Sigma_{k-1|k-2}$  actually assumes that the filtering remains around its steady state. A similar usage can be found in [21,23–25].
- (2) The variational Bayesian iteration method is based on fixed-point iterations that are only guaranteed to converge to a local optimum [28], and iteratively updating steps are employed to reduce the negative influence caused by an inaccurate covariance parameter. A similar usage can be found in [26,27].
- (3) The precision and convergence performance of the proposed approach can be further improved by regulating the filtering process into a steady state; for example, using a larger  $\Sigma_w$  for the first few time instants of the filtering process to initialize the  $\tilde{\Sigma}_{k|k-1} = \Sigma_{k-1|k-2}$  of the proposed approach will contribute to better results.

#### 4. Numerical Simulations

To further demonstrate the performance of the variational Bayesian iteration-based invariant Kaman filter, the attitude estimation system (6) and (7) was simulated with  $\Sigma_{0|0} = 0.5236^2 I_{3 \times 3}$ ,  $\Sigma_w = 0.01745^2 I_{3 \times 3}$ ,  $b' = [1, 0, 0]^T$ ,  $b'' = [0, 1, 0]^T$ ,  $\Sigma_{v'} = 0.0873^2 I_{3 \times 3}$ ,  $\Sigma_{v''} = 0.0873^2 I_{3 \times 3}$ . The real attitude trajectories were generated with the true  $\Sigma_w$  and the number of total time steps was set to 10,000 s. Note that, according to (11), the control parameter  $\Omega_k$  is independent of the error propagation process; in this work, it was set to a time-dependent random rotation matrix. To study the filtering adaptability to  $\Sigma_w$  of various qualities, the filtering of the proposed approach (denoted VBAIKF), the IKF [14,16], and the QeIKF [23,24] was conducted using an inaccurate  $\hat{\Sigma}_w = \Sigma_w \times \text{diag}([\alpha, 1/\alpha, 1])$  with  $\alpha$  ranging from 1 to 10. To implement the VBIKF and the QeIKF, the filtering instants before  $k = 8$  were initialized using the IKF with an inaccurate  $\hat{\Sigma}_w$  and, for this attitude model, at time step 8 the covariance parameters of the IKF gradually converge. The error variable in the Lie algebra of 5000 random simulations was employed to evaluate the root mean square error  $RMSE_k$  during the filtering processes and the average root mean square error  $ARMSE$ , i.e.,

$$RMSE_k \triangleq \sqrt{\frac{1}{5000} \sum_{l=1}^{5000} \|\zeta_{k,l} - \zeta_{k,l}^*\|^2} \quad (51)$$

$$ARMSE \triangleq \sqrt{\frac{1}{5000 \times 5000} \sum_{k=1}^{5000} \sum_{l=1}^{5000} \|\hat{\zeta}_{k,l} - \zeta_{k,l}\|^2} \quad (52)$$

where  $\zeta_{k,l}$  denotes the true estimate at the  $k$ -th time instant of the  $l$ -th simulation run and  $\hat{\zeta}_{k,l}$  is the corresponding estimate for the true  $\zeta_{k,l}$ ; and  $\|\cdot\|$  denotes the Euclidean vector norm.

As to the performance of the proposed VBIKF, the following conclusions can be drawn according to the results shown in Figures 1–8 and Table 1:

- (1) Although for  $\hat{\Sigma}_w = \Sigma_w$  (i.e.,  $\alpha = 1$ ) the precision of the VBIKF is not optimal, the  $ARMSE$  value of the VBIKF (0.0356) is only slightly inferior to that of the IKF (0.0353) and better than that of the QeIKF (0.0359) as shown in Figure 1 and Table 1;
- (2) For all cases of the biased  $\hat{\Sigma}_w$  with  $\alpha \neq 1$ , the presented  $ARMSE$  and  $RMSE_k$  data clearly demonstrate that the proposed VBIKF not only shows better filtering precision than the QeIKF but its filtering stability is obviously superior to that of QeIKF;
- (3) Note that, for the biased  $\hat{\Sigma}_w$  with different  $\alpha$ , although the  $ARMSE$  of the proposed VBIKF is still influenced to some extent (i.e., the higher  $ARMSE$  value 0.0376 for  $\alpha = 10$ ), the negative influence caused by the inaccurate  $\hat{\Sigma}_w$  is significantly reduced compared with and smaller than the 0.0422 of the QeIKF and the 0.443 of the IKF;
- (4) The respective errors of three elements of  $\zeta_k$  with the corresponding  $3\sigma$  boundary for the VBIKF are presented in Figure 8 using the scaled  $\hat{\Sigma}_w$  with  $\alpha = 1, 2, 4, 6, 8$ , and 10, which clearly shows that most of the time estimation errors would fall within the  $3\sigma$  boundary.
- (5) As to the computational cost, the usage of extra fixed-point iterations introduces a longer running time than that of the conventional methods. For example, in this work the iteration number  $N$  was set to 8 and the average running time was about 6 times that of the conventional IKF. Obviously, an  $N$  that is too large is sure to increase the computational cost of the algorithm's implementation and so a balance between precision and cost should be considered according to the particular application.

As to the  $ARMSE$  result displayed in Figure 1 and Table 1, if the scale parameter  $\alpha$  is close or equal to 1 ( $\alpha = 1$ ), the employed covariance  $\hat{\Sigma}_w$  is close to its true  $\Sigma_w$ . Then, the  $ARMSE$  value of the standard IKF is the lowest among the three methods and its performance is optimal in the least square sense, which is also certified by the data on  $RMSE_k$  displayed in Figure 2. However, when  $\alpha$  becomes larger (for instance,  $\alpha \geq 2$ ), the accuracy of the employed  $\hat{\Sigma}_w$  is biased by the scaling parameter  $\alpha$  and the  $ARMSE$  value of

the IKF increases significantly, which means that the optimal estimation performance of the IKF is destroyed by the inaccurate  $\hat{\Sigma}_w$ . We can conclude that the performance of the IKF is rather sensitive to the accuracy of  $\hat{\Sigma}_w$  and, for cases of an inaccurate  $\hat{\Sigma}_w$ , adaptive methods are necessary to improve the performance of attitude estimation.

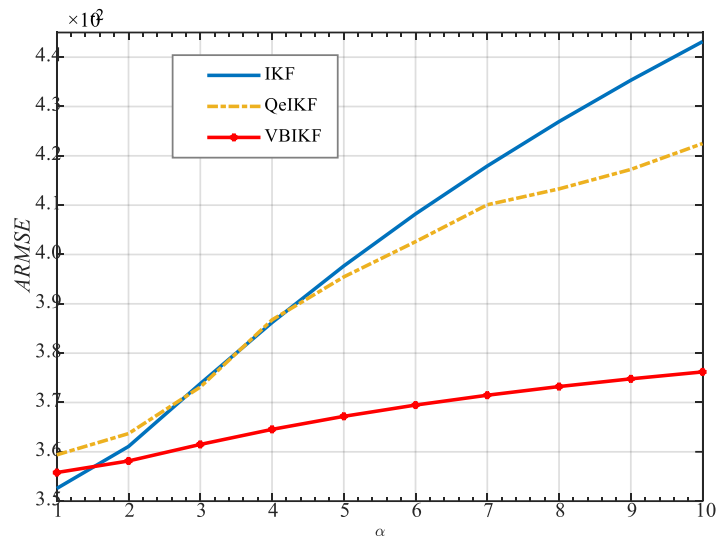


Figure 1. The ARMSE result of the IKF, the QeIKF, and the VBIKF for the scaled  $\hat{\Sigma}_w$  with  $1 \leq \alpha \leq 10$ .

Table 1. The ARMSE result of the different filtering methods with  $\hat{\Sigma}_w$  of different accuracies.

$\hat{\Sigma}_w = \Sigma_w \times \text{diag}([\alpha, 1/\alpha, 1])$	IKF	QeIKF	Proposed VBIKF
$\hat{\Sigma}_w = \Sigma_w$	0.0353	0.0359	0.0356
$\hat{\Sigma}_w = \Sigma_w \times \text{diag}([2, 1/2, 1])$	0.0361	0.0364	0.0358
$\hat{\Sigma}_w = \Sigma_w \times \text{diag}([4, 1/4, 1])$	0.0386	0.0387	0.0365
$\hat{\Sigma}_w = \Sigma_w \times \text{diag}([6, 1/6, 1])$	0.0408	0.0403	0.0369
$\hat{\Sigma}_w = \Sigma_w \times \text{diag}([8, 1/8, 1])$	0.0427	0.0413	0.0373
$\hat{\Sigma}_w = \Sigma_w \times \text{diag}([10, 1/10, 1])$	0.0443	0.0422	0.0376

Note that the ARMSE result of the QeIKF shows some adaptability to the inaccurate  $\hat{\Sigma}_w$  with  $\alpha > 4$ , which can also be verified by the data on  $RMSE_k$  in Figures 4–7; however, the estimation precision is still degraded by the inaccurate parameter and its effective range is limited. For the cases of  $\alpha > 6$ , the estimation precision is still degraded by the inaccurate parameter and its effective range is rather small and limited. Moreover, as the  $RMSE_k$  data show in Figures 5–7, the process data on the QeIKF show some instability during the initializing stage; in fact, the critical issue with the QeIKF is that its filtering is not stable and rather prone to divergence [28]. In conclusion, the filtering precision and stability of the available QeIKF could not effectively reduce the influence of the inaccurate  $\hat{\Sigma}_w$ .

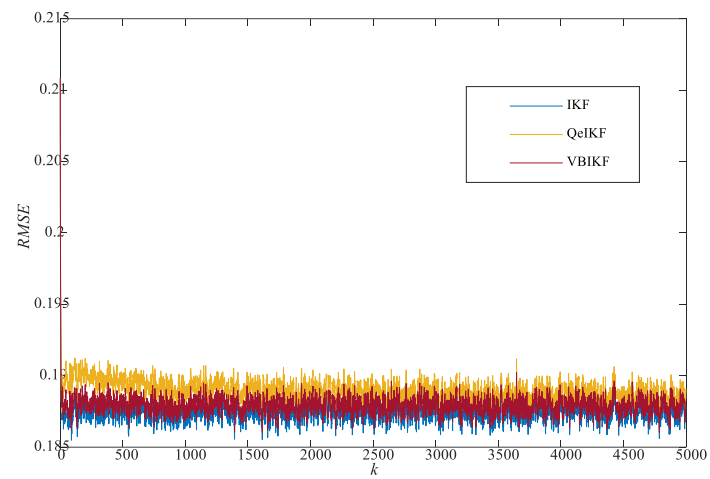


Figure 2. The  $RMSE_k$  result of the IKF, the QeIKF, and the VBIKF for the scaled  $\hat{\Sigma}_w$  with  $\alpha = 1$ .

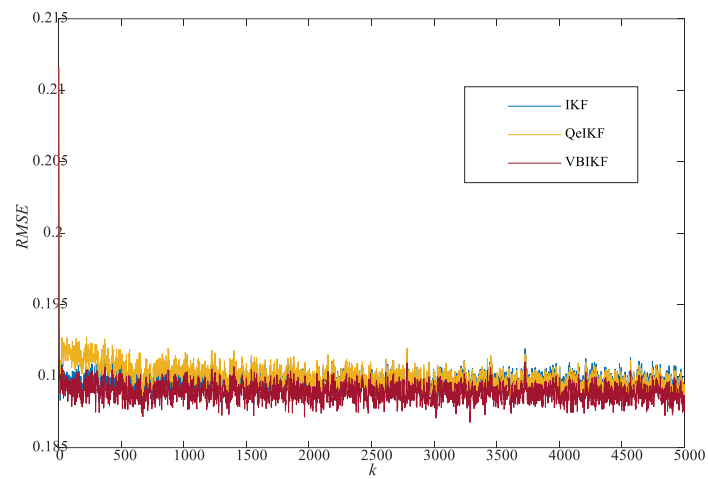


Figure 3. The  $RMSE_k$  result of the IKF, the QeIKF, and the VBIKF for the scaled  $\hat{\Sigma}_w$  with  $\alpha = 2$ .

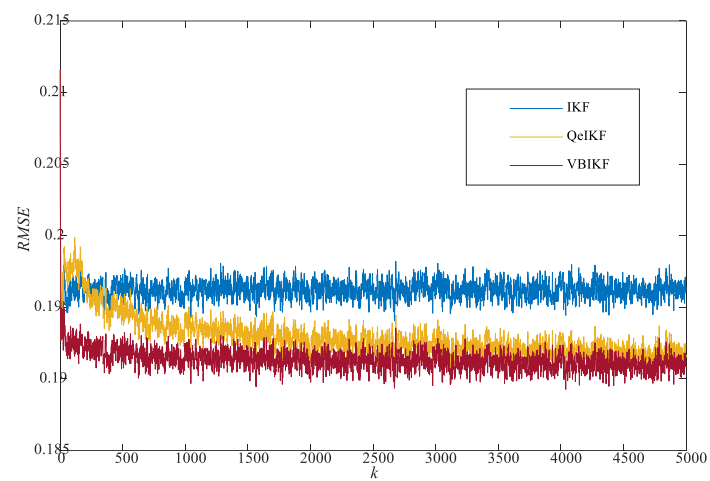


Figure 4. The  $RMSE_k$  result of the IKF, the QeIKF, and the VBIKF for the scaled  $\hat{\Sigma}_w$  with  $\alpha = 4$ .

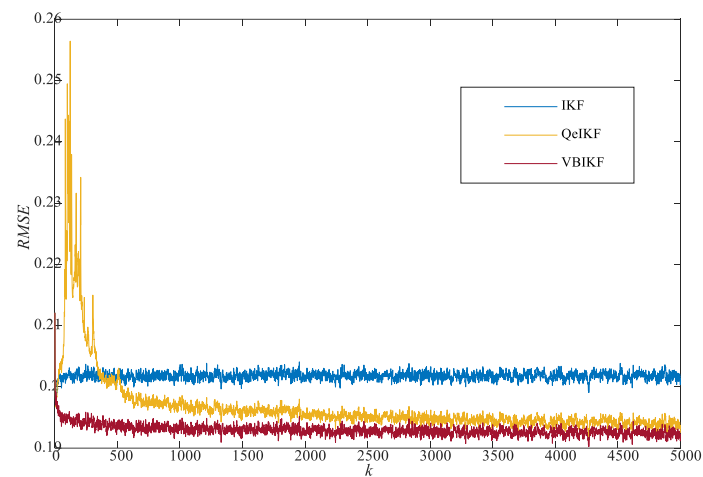


Figure 5. The  $RMSE_k$  result of the IKF, the QeIKF, and the VBIKF for the scaled  $\hat{\Sigma}_w$  with  $\alpha = 6$ .

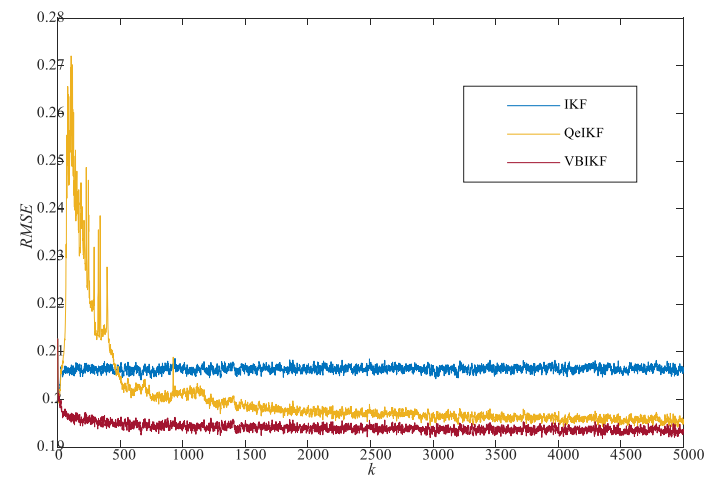


Figure 6. The  $RMSE_k$  result of the IKF, the QeIKF, and the VBIKF for the scaled  $\hat{\Sigma}_w$  with  $\alpha = 8$ .

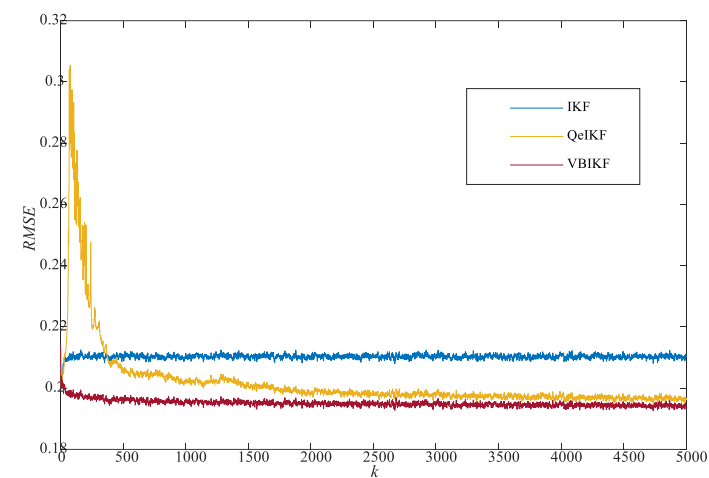
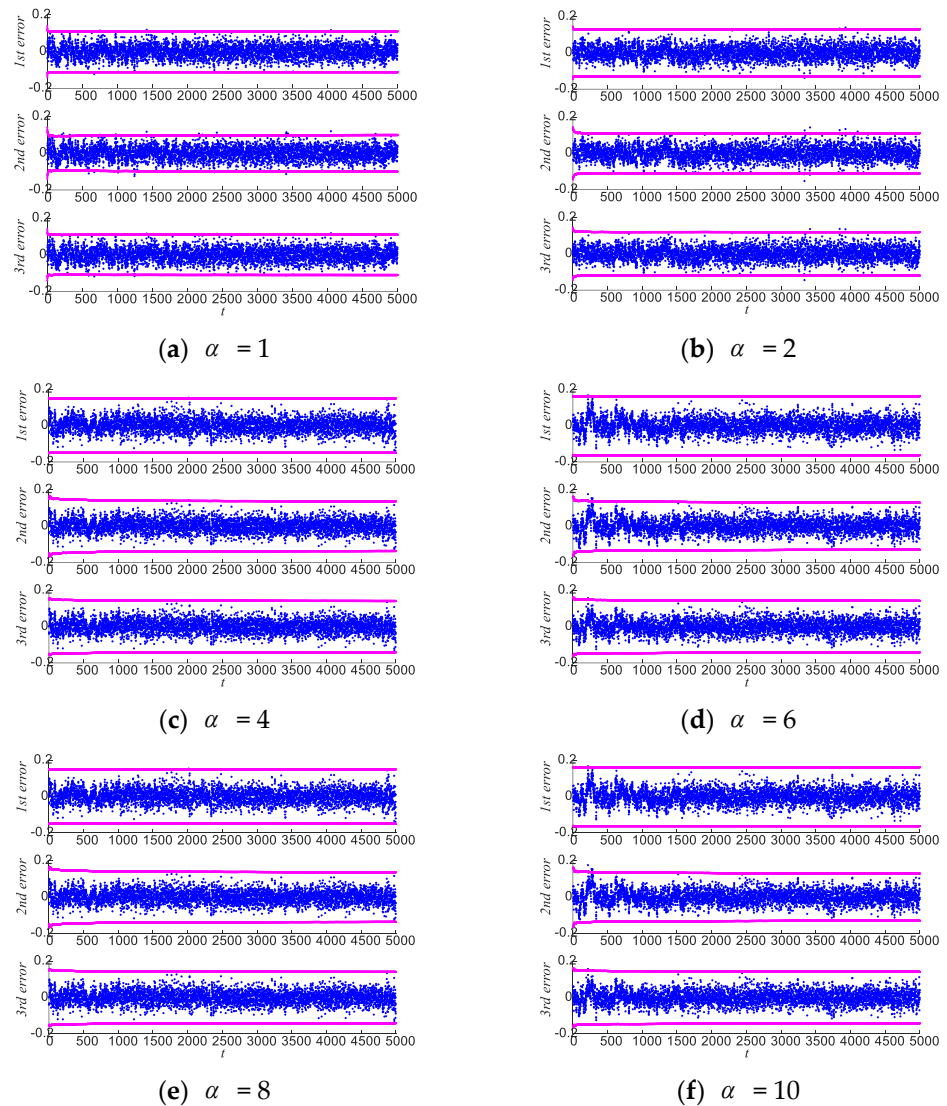


Figure 7. The  $RMSE_k$  result of the IKF, the QeIKF, and the VBIKF for the scaled  $\hat{\Sigma}_w$  with  $\alpha = 10$ .



**Figure 8.** The error data with the  $3\sigma$  boundary on the VBIKF for the scaled  $\hat{\Sigma}_w$  with  $\alpha = 1, 2, 4, 6, 8,$  and  $10$ .

## 5. Conclusions

For aerospace, satellite, and robotics engineering, the matrix Lie group attitude estimation problem with an inaccurate process noise covariance was investigated and a variational Bayesian iteration-based adaptive invariant Kalman filter (VBIKF) was proposed. In the VBIKF, the a priori error covariance is not propagated by the conventional steps but directly calibrated in an iterative manner based on the posterior sequences. The main advantage is that the statistics parameter of the system process noise is no longer required such that the IKF's hard dependency on accurate process noise statistics can be reduced significantly. The numerical simulation results presented demonstrate the superior performance of the proposed VBIKF in terms of filtering adaptability and simplicity.

**Author Contributions:** Data curation, conceptualization, methodology, writing—review and editing, software and funding, J.W.; writing—review and editing, software and validation, Z.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research is funded by the Fundamental Research Funds for Central Universities with number JUSRP121022 and National Natural Science Foundation of China with number 62003112.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data are not publicly available as the data also forms part of an ongoing study.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## References

1. Wu, J.; Shan, S. Dot Product Equality Constrained Attitude Determination from Two Vector Observations: Theory and Astronautical Applications. *Aerospace* **2019**, *6*, 102. [[CrossRef](#)]
2. Phisannupawong, T.; Kamsing, P.; Torteeka, P.; Channumsin, S.; Sawangwit, U.; Hematulin, W.; Jarawan, T.; Somjit, T.; Yooyen, S.; Delahaye, D.; et al. Vision-Based Spacecraft Pose Estimation via a Deep Convolutional Neural Network for Noncooperative Docking Operations. *Aerospace* **2020**, *7*, 126. [[CrossRef](#)]
3. Louédec, M.; Jaulin, L. Interval Extended Kalman Filter-Application to Underwater Localization and Control. *Algorithms* **2021**, *14*, 142. [[CrossRef](#)]
4. Soken, H.E.; Sakai, S.-I.; Asamura, K.; Nakamura, Y.; Takashima, T.; Shinohara, I. Filtering-Based Three-Axis Attitude Determination Package for Spinning Spacecraft: Preliminary Results with Arase. *Aerospace* **2020**, *7*, 97. [[CrossRef](#)]
5. Li, J.; Wei, X.; Zhang, G. An Extended Kalman Filter-Based Attitude Tracking Algorithm for Star Sensors. *Sensors* **2017**, *17*, 1921. [[CrossRef](#)]
6. Pan, C.; Qian, N.; Li, Z.; Gao, J.; Liu, Z.; Shao, K. A Robust Adaptive Cubature Kalman Filter Based on SVD for Dual-Antenna GNSS/MIMU Tightly Coupled Integration. *Remote Sens.* **2021**, *13*, 1943. [[CrossRef](#)]
7. Zheng, L.; Zhan, X.; Zhang, X. Nonlinear Complementary Filter for Attitude Estimation by Fusing Inertial Sensors and a Camera. *Sensors* **2020**, *20*, 6752. [[CrossRef](#)]
8. Ayala, V.; Román-Flores, H.; Torreblanca Todco, M.; Zapana, E. Observability and Symmetries of Linear Control Systems. *Symmetry* **2020**, *12*, 953. [[CrossRef](#)]
9. Deibe, Á.; Antón Nacimiento, J.A.; Cardenal, J.; López Peña, F. A Kalman Filter for Nonlinear Attitude Estimation Using Time Variable Matrices and Quaternions. *Sensors* **2020**, *20*, 6731. [[CrossRef](#)]
10. Guo, H.; Liu, H.; Hu, X.; Zhou, Y. A Global Interconnected Observer for Attitude and Gyro Bias Estimation with Vector Measurements. *Sensors* **2020**, *20*, 6514. [[CrossRef](#)]
11. Chaturvedi, N.; Sanyal, A.; McClamroch, A. Rigid-body attitude control using rotation matrices for continuous singularity-free control laws. *IEEE Control Syst. Mag.* **2011**, *31*, 30–51.
12. Bonnabel, S.; Martin, P.; Salaun, E. Invariant Extended Kalman Filter: Theory and application to a velocity-aided attitude estimation problem. In Proceedings of the 48th IEEE Conference on Decision and Control (CDC) Held jointly with 2009 28th Chinese Control Conference, Shanghai, China, 15–18 December 2009. [[CrossRef](#)]
13. Vasconcelos, J.; Cunha, R.; Silvestre, C.; Oliveira, P. A nonlinear position and attitude observer on SE(3) using landmark measurements. *Syst. Control Lett.* **2010**, *59*, 155–166. [[CrossRef](#)]
14. Barrau, A.; Bonnabel, S. Intrinsic filtering on Lie groups with applications to attitude estimation. *IEEE Trans. Autom. Contr.* **2014**, *60*, 436–449. [[CrossRef](#)]
15. Barrau, A.; Bonnabel, S. The invariant extended Kalman filter as a stable observer. *IEEE Trans. Autom. Contr.* **2017**, *62*, 1797–1812. [[CrossRef](#)]
16. Barrau, A.; Bonnabel, S. Invariant Kalman filtering. *Annu. Rev. Control Robot. Auton. Syst.* **2018**, *1*, 237–257. [[CrossRef](#)]
17. Batista, P.; Silvestre, C.; Oliveira, P. A GES attitude observer with single vector observations. *Automatica* **2012**, *49*, 388–395. [[CrossRef](#)]
18. Chirikjian, G.; Kobilarov, M. Gaussian approximation of non-linear measurement models on lie groups. In Proceedings of the IEEE Conference on Decision and Control, Osaka, Japan, 15–18 December 2015.
19. Barfoot, T.; Furgale, P. Associating uncertainty with three-dimensional poses for use in estimation problems. *IEEE Trans. Robot.* **2014**, *30*, 679–693. [[CrossRef](#)]
20. Said, S.; Manton, J. Extrinsic mean of Brownian distributions on compact lie groups. *IEEE Trans. Inf. Theory* **2012**, *58*, 3521–3535. [[CrossRef](#)]
21. Karasalo, M.; Hu, X. An optimization approach to adaptive Kalman filtering. *Automatica* **2011**, *47*, 1785–1793. [[CrossRef](#)]
22. Wang, J.; Wang, J.; Zhang, D.; Shao, X.; Chen, G. Kalman filtering through the feedback adaption of prior error covariance. *Signal Process.* **2018**, *152*, 47–53. [[CrossRef](#)]
23. Feng, B.; Fu, M.; Ma, H.; Xia, Y.; Wang, B. Kalman filter with recursive covariance Estimation-sequentially estimating process noise covariance. *IEEE Trans. Ind. Electron.* **2014**, *61*, 6253–6263. [[CrossRef](#)]



24. Zanni, L.; Le Boudec, J.; Cherkaoui, R.; Paolone, M. A prediction-error covariance estimator for adaptive Kalman filtering in step-varying processes: Application to power-system state estimation. *IEEE Trans. Contr. Syst. Technol.* **2017**, *25*, 1683–1697. [[CrossRef](#)]
25. Mohamed, A.; Schwarz, K. Adaptive Kalman Filtering for INS/GPS. *J. Geod.* **1999**, *73*, 193–203. [[CrossRef](#)]
26. Ardeshiri, T.; Özkan, E.; Orguner, U.; Gustafsson, F. Approximate Bayesian smoothing with unknown process and measurement noise covariance. *IEEE Signal Process. Lett.* **2015**, *22*, 2450–2454. [[CrossRef](#)]
27. Assa, A.; Plataniotinos, K. Adaptive Kalman filtering by covariance sampling. *IEEE Signal Process. Lett.* **2017**, *24*, 1288–1292. [[CrossRef](#)]
28. Huang, Y.; Zhang, Y.; Wu, Z.; Li, N.; Chambers, J. A novel adaptive Kalman filter with inaccurate process and measurement noise covariance matrices. *IEEE Trans. Autom. Contr.* **2018**, *63*, 594–601. [[CrossRef](#)]
29. Ćesić, J.; Markovi, I.; Petrovi, I. Mixture Reduction on Matrix Lie Groups. *IEEE Signal Process. Lett.* **2017**, *24*, 1719–1723. [[CrossRef](#)]
30. Ćesić, J.; Markovi, I.; Bukal, M.; Petrović, I. Extended Information Filter on Matrix Lie Groups. *Automatica* **2017**, *82*, 226–234. [[CrossRef](#)]
31. Kang, D.; Jang, C.; Park, F. Unscented Kalman Filtering for Simultaneous Estimation of Attitude and Gyroscope Bias. *IEEE/ASME Trans. Mechatron.* **2019**, *24*, 350–360. [[CrossRef](#)]
32. Bourmaud, G.; Mégret, R.; Arnaudon, M.; Giremus, A. Continuous-Discrete Extended Kalman Filter on Matrix Lie Groups Using Concentrated Gaussian Distributions. *J. Math. Imaging Vis.* **2015**, *51*, 209–228. [[CrossRef](#)]
33. Tzikas, D.; Likas, A.; Galatsanos, N. The Variational Approximation for Bayesian Inference. *IEEE Signal Process. Mag.* **2008**, *25*, 131–146. [[CrossRef](#)]