A covariance correction step for Kalman filtering with an attitude

Mark W. Mueller
Mechanical Engineering Department, UC Berkeley, Berkeley CA 94720

Markus Hehn, Raffaello D'Andrea
Institute for Dynamic Systems and Control, ETH Zurich, Switzerland

Redundant attitude representations are often used in Kalman filters used for estimating dynamic states which include an attitude. A minimal, three element attitude deviation is combined with a reference attitude, where the deviation is included in the filter state and has an associated covariance estimate. This paper derives a reset step which adjusts the covariance matrix when information is moved from the attitude deviation to the reference attitude. When combined with the extended or unscented Kalman filter prediction and measurement steps, the reset allows one to easily construct a Kalman filter for a system whose state includes an attitude. This algorithm is closely related to (and a correction to) the Multiplicative Extended Kalman Filter (MEKF) or the Unscented Quaternion Estimator (USQUE), depending on whether the reset is combined with an extended or unscented Kalman filter. In comparison to the MEKF it is more general and includes a reset after the measurement update, and a reset after both the prediction and measurement update steps of the USQUE. This reset step is derived by tracking mean and covariance through a linearization, similarly to an extended Kalman filter prediction step. The reset step is validated using Monte Carlo sampling.

I. Introduction

The dynamic state for many systems of engineering interest include an attitude or orientation of one frame with respect to another. The estimation for such systems is often done with a Kalman filter, especially with an extended Kalman filter (EKF) or an unscented Kalman filter (UKF). The EKF applies the Kalman filter to a first-order approximation of the underlying nonlinear system, with the approximation evaluated at the current state estimate. The UKF instead uses a set of deterministically chosen points which approximate the underlying distribution, and these points are then transformed through the nonlinear equations [1]. Extensions of the EKF exist which include higher order effects, e.g. [2]. For strongly nonlinear systems neither the EKF nor UKF may provide satisfactory results – an example involving attitudes is given in [3].

A common engineering approach to estimating a dynamic state including an attitude is to use a redundant attitude representation, where both a reference attitude and an attitude error are used and their composition represents the attitude estimate: the attitude error is encoded with a minimal representation in the stochastic state (i.e. it has an associated covariance), and the reference attitude is updated to keep the attitude error small (but has no associated covariance). Alternative methods, such as the use of more than three dynamic states combined with constraints incur additional theoretical and computational complexity, see [4, 5] for examples.

Some recent work has focused on exploiting differential geometrical properties of systems with attitude dynamics. For example, in [6] an observer is developed which exploits symmetries in the system dynamics, which is shown to have favourable convergence properties. The theory allows, for example, to develop observers for rigid bodies moving in space, with vector measurements in a body-fixed frame. The ideas of [6] are further developed in [7], where the resulting Invariant EKF is compared also to the MEKF and USQUE. Strong statements about the properties of these filters may be made, but under restrictive assumptions (e.g. isotropic process noise and specific properties of the measurement equation).

In [8] a framework is given for global propagation of uncertainty of an attitude, using symplectic integration and a Fourier-spectrum-like representation of the attitude uncertainty – this may however be difficult to extend to systems with additional states. A globally exponentially stable observer, not based on Kalman filtering, is presented in [9], which estimates the rotation matrix as though it has nine degrees of freedom.

An early example of using a minimal attitude representation in a Kalman filter state to represent an attitude error in addition to having a reference attitude is given in [10], where the goal is to estimate the attitude and angular velocity of a spacecraft using an extended Kalman filter, where the filter state includes three Euler angles which are reset to zero after each measurement update and thus kept far from their singularities. A very influential example is the Multiplicative Extended Kalman Filter (MEKF) of [11, 12]. In [12] the problem of estimating the attitude and rate gyroscope bias of a spacecraft is addressed, and the presented algorithm uses a minimal three dimensional attitude error in a Kalman filter, in addition to a reference attitude. Information from the attitude error is moved to the reference attitude after each measurement update, in a so-called reset step. The authors claim that the reset step does not affect the filter covariance, specifically [12]: "The reset does not modify the covariance because it neither increases nor decreases the total information content of the estimate; it merely moves this information from one part of the attitude representation to another." Similar algorithms (with similar reset steps) are presented in [13–17], where similar claims are made that the reset does not modify the estimate covariance.

Unfortunately, this statement that the covariance is unchanged during the reset is incorrect, even to first order in the attitude error. This work derives a correction for the attitude error mean and covariance during the reset, based on a first-order approximation thereof. This first-order approximation is in line with the general philosophy behind the extended Kalman filter, of using first-order approximations of the system equations to approximate the conditional probability distributions of the state estimate, and the resulting covariance correction is of a similar form to the familiar Kalman filter prediction step. When combined with the usual Extended or Unscented Kalman filter, this allows to straight-forwardly create an estimator for a system with an attitude.

That something is amiss with the MEKF of [12] has been noted previously in the literature, specifically in [5, 18], and [19]. A filter for estimating an attitude based on bias-free rate gyroscope measurements and unit vector measurements is presented in [18]; and specifically a covariance correction is suggested similar to the one presented herein (see (28) and (29) of [18] – note however, that the assumptions made for the derivation are somewhat different). A second-order covariance correction is proposed in [5], but it is however still stated that no correction to the covariance is necessary to first order. Furthermore, in [16, p. 243] the reset of [12] is stated as being correct, with the caveat that "Not everyone agrees with this statement".

The approach of [12] has inspired many other works, and is widely used in practise: see for example [20–24]. The addition of the proposed reset step may be expected to improve estimation performance in such systems.

The contribution of this paper is as follows:

1. To propose an algorithm which allows estimation of the dynamic state of a system (where the state includes an attitude), by making use of a reference attitude and an attitude error. When based on the extended Kalman filter, the result is an extension to and correction of the MEKF; and when based on the Unscented Kalman filter, it may be seen as a correction of the USQUE. The algorithm may be directly applied to systems with arbitrarily complicated dynamics, including, for example, angular velocity dynamics dependent on other states.

2. To show that the widely used method presented in e.g. [12–16] does not keep track of the estimate statistics correctly, even to first order.

3. To demystify the attitude reset step, deriving the first-order correction using a 3×3 rotation matrix.

II. Attitude representations

For all derivations in this paper the fundamental representation of attitude is the 3×3 rotation matrix, so that the transformation of a vector by an attitude from one reference frame to another is a matrix multiplication of the attitudes. Furthermore, the composition of consecutive attitude transformations is their matrix product.

The skew-symmetric matrix form of the cross product $[\![v]\!]$ of a vector $v = (v_1, v_2, v_3)$ is defined such that $[\![v]\!]b = -v \times b$, i.e.

$$[\![v]\!] := \begin{bmatrix} 0 & v_3 & -v_2 \\ -v_3 & 0 & v_1 \\ v_2 & -v_1 & 0 \end{bmatrix}. \tag{1}$$

This follows the sign convention of [25], note however the sign difference to e.g. the operator $S(\cdot)$ of [9]. The notation (v_1, v_2, v_3) is used to compactly denote the elements of a column vector.

An attitude may also be expressed using a rotation vector $\delta \in \mathbb{R}^3$, where the unit vector in the direction of δ represents the axis of rotation, and the magnitude of δ represents the angle of rotation. The rotation matrix corresponding to the rotation vector may be computed as $\exp([\![\delta]\!])$, where $\exp(M)$ is the matrix exponential of a square matrix M, defined as

$$\exp(M) := \sum_{k=0}^{\infty} \frac{1}{k!} M^k. \tag{2}$$

Note that this may be readily computed in closed form using Euler's formula, see [25, eq. (96)-(99)]. It follows that $\exp([\![\delta]\!])^{-1} = \exp([\![\delta]\!])^{T} = \exp([\![-\delta]\!])$.

The inverse mapping, $\mathrm{rot}^{-1}(R)$, from rotation matrix to rotation vector is given by [25, (102) - (103)].

A. Kinematics of rotation

If two reference frames are moving with respect to one another, with the time-varying rotation matrix $R(t)$ representing their relative orientation and the vector $\omega(t)$ representing their relative angular velocity (the "body-referenced angular velocity"), the differential equation governing $R(t)$ is as below [25, eq. (261)]:

$$\frac{d}{dt} R(t) = [\![\omega(t)]\!]\, R(t). \tag{3}$$

These continuous-time equations are introduced here as they are central to the analysis of the covariance in the reset procedure.

If the angular velocity ω is constant over some time period $t \in [t_0, t_1]$, the above is a linear, time-invariant differential equation whose solution is

$$R(t_1) = \exp([\![(t_1 - t_0)\omega]\!])\; R(t_0). \tag{4}$$

The differential equation of the corresponding rotation vector is as below [25, eqs. (276) and (428)], where the time dependence is neglected for the sake of compactness.

$$\frac{d}{dt}\delta = \omega - \tfrac{1}{2}[\![\delta]\!]\omega + \frac{2 - |\delta| \cot\left(\tfrac{1}{2}|\delta|\right)}{2|\delta|^2}[\![\delta]\!]^2\omega \tag{5a}$$

$$= \omega - \tfrac{1}{2}[\![\delta]\!]\omega + o(|\delta|) \tag{5b}$$

where $|\cdot|$ represents the Euclidean norm and the Landau symbol $o(x)$ is used to represent higher-order terms, i.e. a quantity for which the following holds:

$$\lim_{x \to 0} \frac{o(x)}{x} = 0. \tag{6}$$

B. Alternative attitude representations

Instead of the rotation matrix, different representations may be used in the implementation of the resulting algorithm, whilst preserving the results presented herein. These representations, especially the Euler symmetric parameters (unit quaternion), may offer practical benefits such as improved numerical stability or computational speed. The matrix multiplications are then replaced by the composition rule for the chosen representation.

Alternative three element parametrizations may be used instead of the rotation vector, also without affecting the fundamental results derived in the paper. Examples of alternative parametrizations include the vector part of the Euler symmetric parameters, or the Rodrigues parameters [25] – to first order, these are equivalent to the rotation vector (up to a constant), and thus the first order analysis remains unchanged.

III. Problem statement and solution approach

The problem considered is that of estimating the dynamic state of a system using measurements, a model for the state dynamics, a model of the measurement system, and information about the dynamic states' initial probability distributions. The system's dynamic state includes an attitude R and other states collected into the vector $\xi \in \mathbb{R}^{n_\xi}$, where both R and ξ are random variables. The attitude R is taken to be a rotation matrix, although alternative representations may be used in an implementation, specifically the unit quaternion. The state evolves in discrete time steps k according to the following dynamic equations, with the random variable $\eta[k]$ representing process noise (assumed white, independent of the initial condition, and zero-mean):

$$\xi[k] = \bar{f}_1\left(k{-}1, \xi[k{-}1], R[k{-}1], \eta[k{-}1]\right) \tag{7}$$

$$R[k] = \bar{f}_2\left(k{-}1, \xi[k{-}1], R[k{-}1], \eta[k{-}1]\right). \tag{8}$$

If the underlying system dynamics are instead in continuous time, the above equations may be taken as the integrals of the continuous equations over one sampling interval, e.g. approximated by an Euler discretisation. The function $\bar{f}_2(\cdot)$ outputs a rotation matrix.

Measurements $z[k]$ are available at discrete times, as a function of the state and the measurement noise random variable $\zeta[k]$ (also assumed white, zero mean, and independent of both the initial condition and η):

$$z[k] = \bar{h}\left(k, \xi[k], R[k], \zeta[k]\right). \tag{9}$$

The goal is to estimate the system's state $\xi[k]$ and $R[k]$ recursively from the measurement sequence z, information about their initial conditions, and the dynamic and measurement models.

This is done by introducing the stochastic state $x[k] \in \mathbb{R}^{n_\xi+3}$, partitioned such that $x[k] = (\xi[k], \delta[k])$ where the random variable $\delta[k] \in \mathbb{R}^3$ represents a (small) attitude error parametrised through a rotation vector. It is defined as

$$\delta[k] := \mathrm{rot}^{-1}\big(R[k]\, R_{\mathrm{ref}}[k]^{-1}\big) \tag{10}$$

with $R_{\mathrm{ref}}[k]$ a deterministic attitude, so that

$$R[k] = \exp([\![\delta[k]]\!])\; R_{\mathrm{ref}}[k]. \tag{11}$$

This representation introduces a redundant attitude, which is exploited to avoid issues relating to singularities and constraints in attitude representations. This redundant formulation is not novel, and can be found in e.g. [12], or in a somewhat different form using Euler angles in [10].

The system dynamics (7)-(8) are now combined and rewritten to use the stochastic state $x[k]$ and the reference attitude $R_{\mathrm{ref}}[k]$ so that

$$x[k] = f(k{-}1, x[k{-}1], R_{\mathrm{ref}}[k{-}1], \eta[k{-}1]) \tag{12}$$

wherein $R_{\mathrm{ref}}[k-1]$ is a constant, and the change in attitude of (8) affects the components $\delta[k]$ of $x[k]$. Rewriting these equations requires the composition of attitudes using the rotation vector, such as those in Section II.

The measurement equation is likewise rewritten as a function of the new state variables:

$$z[k] = h(k, x[k], R_{\mathrm{ref}}[k], \zeta[k]). \tag{13}$$

The proposed algorithm uses these rewritten equations and is based on the extended Kalman filter (see e.g. [1, 26]), and introduces two additional steps to correct the attitude error statistics, so that the recursive estimation strategy then consists of four steps.

1. A Kalman prediction step, that uses the process equation (12) to propagate the estimate through the dynamics. During this step, the reference attitude is unchanged.

2. A prediction reset step, where the reference attitude is changed such that the estimate of the post-reset attitude error equals zero, i.e. it is maximally far from its singularities.

3. A Kalman measurement update, that uses the measurement model (13) to correct the estimate with a given measurement. During this step, the reference attitude is again unchanged.

4. A measurement reset step, where again the reference attitude is adapted such that the estimate of the post-reset attitude error is reset to zero.

The algorithm may be adapted straight-forwardly for use with the Unscented Kalman Filter (UKF) (see e.g. [27]), by replacing steps 1 and 3 with the corresponding steps from the UKF. In this case, the resulting algorithm would be a correction of the USQUE of [13].

The derivation and proof of necessity of the reset step are novel compared to the methods of [12–16], as is the generalisation of the method to systems of any dynamics that can be expressed in the form (7)-(9).

It should be noted that the covariance, as used in the EKF, is only an approximation of the covariance of the quantities to be estimated – for example, in practical situations the true distributions of the noise may be unknown, the system dynamics may be subject to various approximations, and the noise sequences may not be white. Furthermore, except in some special situations (e.g. [8])

it may be computationally intractable to accurately compute non-Gaussian statistical properties being transformed by nonlinear dynamics/measurements.

Remark. The proposed structure differs somewhat from that of the MEKF, where the attitude error is maintained at zero throughout the prediction step by varying the reference attitude. This is discussed in more detail in Section V A.

A. Attitude error reset

The reset step does not change the actual attitude in the estimate, but modifies the reference attitude R_{ref} so that the post-reset estimate of the attitude random variable δ is zero, i.e. is maximally far away from its singularities.

Problem 1. Let the pre-reset reference attitude be $R_{\mathrm{ref,pre}}$, and the pre-reset attitude error be δ_{pre} with associated mean and covariance:

$$\mu_{\mathrm{pre}} := \mathrm{E}(\delta_{\mathrm{pre}}) \tag{14}$$

$$\Sigma_{\mathrm{pre}} := \mathrm{Var}(\delta_{\mathrm{pre}}) \tag{15}$$

The corresponding post-reset random variable δ_{post} and reference attitude $R_{\mathrm{ref,post}}$ are introduced, which must satisfy the following two equalities:

$$\exp([\![\delta_{\mathrm{post}}]\!])\ R_{\mathrm{ref,post}} = \exp([\![\delta_{\mathrm{pre}}]\!])\ R_{\mathrm{ref,pre}} \tag{16}$$

$$\mathrm{E}(\delta_{\mathrm{post}}) = 0. \tag{17}$$

IV. First order attitude reset

In this section an approximate solution to Problem 1 is derived based on studying the reset operation as a continuous rotation, whose effects are analysed to first order in the attitude error.

Theorem 1. *To first order in the attitude error δ_{pre}, the solution to Problem 1 is:*

$$R_{\mathrm{ref,post}} = \exp([\![\mu_{\mathrm{pre}}]\!])\ R_{\mathrm{ref,pre}} \tag{18}$$

$$\Sigma_{\mathrm{post}} = \exp\left([\![\tfrac{1}{2}\mu_{\mathrm{pre}}]\!]\right) \Sigma_{\mathrm{pre}} \exp\left([\![\tfrac{1}{2}\mu_{\mathrm{pre}}]\!]\right)^{T}. \tag{19}$$

Proof. The pseudo-time $t \in [0,1]$ is introduced, and the reset is considered as a continuous operation starting at $t=0$ and ending at $t=1$. The time-varying reference attitude $R_{\text{ref}}(t)$ and $\delta(t)$ are introduced, which have to satisfy the boundary conditions $\delta(0)=\delta_{\text{pre}}$, $R_{\text{ref}}(0)=R_{\text{ref,pre}}$, $\delta(1)=\delta_{\text{post}}$, and $R_{\text{ref}}(1)=R_{\text{ref,post}}$. Given these boundary conditions, a sufficient condition for satisfying (16) is

$$\frac{d}{dt}\left(\exp(\llbracket\delta(t)\rrbracket)\ R_{\text{ref}}(t)\right)=0. \tag{20}$$

Applying the derivative product rule, and substituting the kinematic equation for the rotation matrix (3) yields

$$\llbracket\omega_\delta(t)\rrbracket\exp(\llbracket\delta(t)\rrbracket)\ R_{\text{ref}}(t)+\exp(\llbracket\delta(t)\rrbracket)\ \llbracket\omega_{\text{ref}}(t)\rrbracket R_{\text{ref}}(t)=0 \tag{21}$$

where $\omega_{\text{ref}}(t)$ and $\omega_\delta(t)$ are angular velocities, with specifically $\omega_{\text{ref}}(t)=\omega_{\text{ref}}$ taken as a deterministic constant, to be computed. This then yields

$$\llbracket\omega_\delta(t)\rrbracket=-\exp(\llbracket\delta(t)\rrbracket)\ \llbracket\omega_{\text{ref}}\rrbracket\exp(\llbracket\delta(t)\rrbracket)^{\ -1} \tag{22}$$

or, by simplifying (see (80) of [25]),

$$\omega_\delta(t)=-\exp(\llbracket\delta(t)\rrbracket)\,\omega_{\text{ref}}. \tag{23}$$

Substituting the definition of the matrix exponential (2), and then substituting into the kinematic equation (5b) gives

$$\frac{d}{dt}\delta(t)=-\omega_{\text{ref}}+\frac{1}{2}\llbracket\omega_{\text{ref}}\rrbracket\delta(t)+o(|\delta(t)|)\,. \tag{24}$$

Neglecting the higher order terms, (24) represents an affine, time-invariant, differential equation in $\delta(t)$, with ω_{ref} a deterministic constant, the solution to which is

$$\delta(1)=\exp\!\left(\llbracket\tfrac{1}{2}\omega_{\text{ref}}\rrbracket\right)\,\delta(0)-\omega_{\text{ref}}+o(|\delta(0)|)\,. \tag{25}$$

The simple closed-form solution follows from the fact that $\llbracket\omega\rrbracket\omega=0$ for all ω.

Neglecting again the higher order terms, and noting that $\delta_{\text{post}}=\delta(1)$, and $\delta_{\text{pre}}=\delta(0)$, it follows that

$$\mathrm{E}(\delta_{\text{post}})\approx\exp\!\left(\llbracket\tfrac{1}{2}\omega_{\text{ref}}\rrbracket\right)\,\mathrm{E}(\delta_{\text{pre}})-\omega_{\text{ref}} \tag{26}$$

$$\mathrm{Var}(\delta_{\text{post}})\approx\exp\!\left(\llbracket\tfrac{1}{2}\omega_{\text{ref}}\rrbracket\right)\,\mathrm{Var}(\delta_{\text{pre}})\exp\!\left(\llbracket\tfrac{1}{2}\omega_{\text{ref}}\rrbracket\right)^{\ T}. \tag{27}$$

Rearranging (26), exploiting the fact that $\exp([\![v]\!])\, v = v$ for all vectors v, and enforcing the requirement that $\mathrm{E}(\delta_{\mathrm{post}}) = 0$ yields

$$\omega_{\mathrm{ref}} = \mathrm{E}(\delta_{\mathrm{pre}}) \tag{28}$$

which yields (19) when substituted into (27).

The post-reset reference $R_{\mathrm{ref,post}}$ can be computed from (28), by noting that during the reset

$$\frac{d}{dt} R_{\mathrm{ref}}(t) = [\![\omega_{\mathrm{ref}}]\!]\, R_{\mathrm{ref}}(t). \tag{29}$$

Substituting (4), (18) follows. □

Remark. The assumption that the higher-order terms contribute only negligibly to the mean and variance in (26)-(27) is strong, and may be a source of substantial error. However, it is the same assumption that is typically made when using the EKF, and as such should not place any additional restrictions on an EKF that incorporates this reset step.

Remark. Note that Theorem 1 contradicts the assertions made in [12–16] that a reset step does not affect the estimate covariance. Without a transformation of the form (19) any Kalman filter implementation will fail to estimate the error statistics correctly, even to first order.

Remark. Although the derivation approach is different, the result of Theorem 1 is related to the result of [18, eq. (28)-(29)]. For Theorem 1, however, the correction to the reference attitude (18) is shown to follow as a consequence of the requirement that $\mathrm{E}(\delta_{\mathrm{post}}) = 0$, rather than being taken as an assumption.

A. Validation

The accuracy of the reset step may be quantified by Monte Carlo sampling, where samples from an initial attitude distribution are transformed through the proposed reset step. The sample mean and covariance of the transformed samples are then compared to the mean and covariance as computed in Theorem 1. First, a single example is analysed, which is then followed by an ensemble of Monte Carlo tests to allow for statistical evaluation of performance.

1. Single example

Consider the following example, where the pre-reset variables are taken as below:

$$R_{\text{ref,pre}} = I, \quad \mu_{\text{pre}} = (0.1, 0, 0) \tag{30}$$

$$\Sigma_{\text{pre}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 0 \end{bmatrix} \times 10^{-2}. \tag{31}$$

There are thus two directions with zero uncertainty: this was chosen because the effects of the reset are made more obvious.

Pre-reset particles $\delta_{\text{mc,pre}}[i]$ are sampled independently from a normal distribution $\mathcal{N}(\mu_{\text{pre}}, \Sigma_{\text{pre}})$, with mean μ_{pre} and covariance Σ_{pre}. They are transformed to post-reset particles $\delta_{\text{mc,post}}[i]$, with

$$\delta_{\text{mc,post}}[i] := \text{rot}^{-1}\Big(\exp([\![\delta_{\text{mc,pre}}[i]]\!])\ R_{\text{ref,post}}^{-1}\Big) \tag{32}$$

where $R_{\text{ref,post}}$ is computed as in Theorem 1 using the mean μ_{pre}. This implies that the pre- and post-reset particles, with the given $R_{\text{ref,post}}$, satisfy (16).

A set of 10^9 pre-reset samples $\delta_{\text{mc,pre}}[i]$ were transformed, and the sample mean and covariance of $\delta_{\text{mc,post}}[\cdot]$ were computed as:

$$\text{E}(\delta_{\text{mc,post}}[\cdot]) \approx (8.37, -0.05, 0.00) \times 10^{-4} \tag{33}$$

$$\text{Var}(\delta_{\text{mc,post}}[\cdot]) \approx \begin{bmatrix} 0 & 0 & 0 \\ 0 & 9.967 & 0.499 \\ 0 & 0.499 & 0.025 \end{bmatrix} \times 10^{-2}. \tag{34}$$

This may be compared to the post-reset estimated as computed with Theorem 1:

$$\mu_{\text{post}} = (0, 0, 0) \tag{35}$$

$$\Sigma_{\text{post}} \approx \begin{bmatrix} 0 & 0 & 0 \\ 0 & 9.975 & 0.499 \\ 0 & 0.499 & 0.025 \end{bmatrix} \times 10^{-2}. \tag{36}$$

Two dimensionless scalar error measures are defined for comparing the post-reset sample mean and covariance to the predicted mean and covariance: ϵ_μ is a normalised indication of the error in the mean, and ϵ_Σ is a normalised indication of the error in the covariance:

$$\epsilon_\mu(\bar{\mu}) := \frac{|\mathrm{E}(\delta_{\mathrm{mc,post}}[\cdot]) - \bar{\mu}|}{|\mu_{\mathrm{pre}}|} \tag{37}$$

$$\epsilon_\Sigma(\bar{\Sigma}) := \frac{\bar{\sigma}\left(\mathrm{Var}(\delta_{\mathrm{mc,post}}[\cdot]) - \bar{\Sigma}\right)}{\bar{\sigma}\left(\Sigma_{\mathrm{pre}}\right)} \tag{38}$$

where $\bar{\sigma}\left(\cdot\right)$ represents the maximum singular value of its matrix argument.

For the post-reset estimates as computed in (35)-(36), the mean and covariance errors are $\epsilon_\mu \approx 8.4 \times 10^{-3}$ and $\epsilon_\Sigma \approx 8.0 \times 10^{-4}$, respectively. These normalised error metrics will be used again during the ensemble comparison, below.

Remark. If the covariance is left unchanged in the reset, as is claimed correct in [12–16], the covariance error is $\epsilon_\Sigma \approx 50 \times 10^{-3}$, or more than sixty times larger than with the proposed method. Furthermore, the post-reset covariances in the third row and column would be estimated as zero.

2. *Ensemble validation*

The preceding Monte Carlo analysis may be extended by computing the normalised error statistics $\epsilon_\mu[j]$ and $\epsilon_\Sigma[j]$ over an ensemble of initial attitude distributions, where each instance j in the ensemble consists of a large number of individual Monte Carlo samples $\delta_{\mathrm{mc,pre}}[j,i]$.

Let the variable ρ represent a magnitude, which will be used to quantify the magnitude of an ensemble: for each instance j in the ensemble, a pre-reset mean $\mu_{\mathrm{pre}}[j]$ is sampled from $\mathcal{N}\left(0, (\rho\,\pi/180)\,I\right)$, and a pre-reset covariance $\Sigma_{\mathrm{pre}}[j]$ is generated as

$$\Sigma_{\mathrm{pre}}[j] = \sum_{k=1}^{3} s_k[j] s_k[j]^T \tag{39}$$

where the $s_k[j]$ are also independently sampled from $\mathcal{N}\left(0, (\rho\,\pi/180)\,I\right)$. The pre-reset reference attitude is taken as identity in all instances. A million samples $\delta_{\mathrm{mc,pre}}[j,i]$ are generated from $\mathcal{N}\left(\mu_{\mathrm{pre}}[j], \Sigma_{\mathrm{pre}}[j]\right)$, similarly to in Section IV A 1, and are subsequently transformed to post-reset particles $\delta_{\mathrm{mc,post}}[j,i]$ analogously to (32).

Ten thousand instances j are generated in each ensemble, and for each instance the sample mean and covariance of the transformed particles are compared to the values computed with Theorem 1, similar to what was done for the single example. The distribution of the resulting errors ϵ_μ and ϵ_Σ

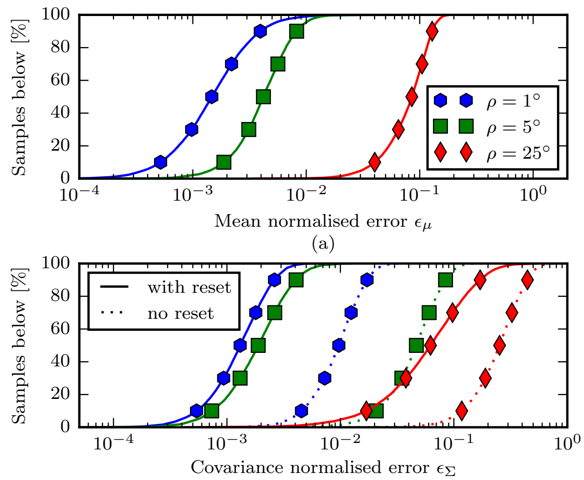

Fig. 1 The normalised mean and covariance relative errors for ensembles of different magnitudes ρ, when comparing the post-reset sample mean and covariance to the predicted mean and covariance computed with the first order attitude reset. The icons are used to indicate the magnitude ρ across both sub-plots. Sub-plot (a) shows the distribution of the normalised mean error ϵ_μ as defined in (37). The normalised covariance error ϵ_Σ as defined in (38) is shown in sub-plot (b) for the proposed method as solid lines, and with the reset step as in [12] as dotted lines (where the covariance is left unchanged during the reset step). For an ensemble with $\rho = 1°$, for example, for 95% of the instances the error ϵ_Σ was below 0.003 with the proposed method (as compared to 0.020 when not changing the covariance during the reset step).

are shown in Fig. 1. The figure shows the results for ensemble magnitudes $\rho \in \{1°, 5°, 25°\}$. The estimation errors are shown to increase as the ensemble magnitude ρ is increased: this is to be expected as for larger vales of ρ the higher-order terms neglected in the derivation of Theorem 1 have a larger influence.

Remark. Fig. 1(b) compares the performance of the reset of Theorem 1 to that proposed in [12–16] where the covariance is left unchanged. The covariance error is shown to be significantly lower with the proposed method than if the covariance is unchanged during the reset.

B. Extension to full state

The reset of Theorem 1 may be straight-forwardly extended to apply to the full estimator state $x = (\xi, \delta)$ and R_{ref}. During the reset, the states ξ remain unchanged and the attitude states

δ are transformed according to Theorem 1:

$$\xi_{\text{post}} = \xi_{\text{pre}} \tag{40}$$

$$\delta_{\text{post}} = \exp\left(\left[\!\left[\tfrac{1}{2}\mathrm{E}(\delta_{\text{pre}})\right]\!\right]\right) \delta_{\text{pre}} - \mathrm{E}(\delta_{\text{pre}})\,. \tag{41}$$

From this follows

$$x_{\text{post}} = T_{\text{reset}}(\mathrm{E}(\delta_{\text{pre}}))x_{\text{pre}} - (0, \mathrm{E}(\delta_{\text{pre}})) \tag{42}$$

where the extended reset transformation matrix is given by

$$T_{\text{reset}}\,(\delta) := \mathrm{diag}\left(I, \exp\left(\left[\!\left[\tfrac{1}{2}\delta\right]\!\right]\right)\right) \tag{43}$$

with $\mathrm{diag}\,(\cdot)$ returning a block diagonal matrix. Thus,

$$\mathrm{E}(x_{\text{post}}) = \left(\mathrm{E}(\xi_{\text{pre}})\,, 0\right) \tag{44}$$

$$\mathrm{Var}(x_{\text{post}}) = T_{\text{reset}}(\mathrm{E}(\delta_{\text{pre}}))\mathrm{Var}(x_{\text{pre}})\, T_{\text{reset}}(\mathrm{E}(\delta_{\text{pre}}))^T \tag{45}$$

$$R_{\text{ref,post}} = \exp([\![\mathrm{E}(\delta_{\text{pre}})]\!])\; R_{\text{ref,pre}}. \tag{46}$$

All covariances related to the attitude error are affected through (45), i.e. the last three rows and columns of the covariance matrix. Because T_{reset} is an orthogonal matrix, the correction does not change any eigenvalues in the covariance matrix.

Remark. The rotation applied to the covariance through (45) is *half* the rotation applied to the reference attitude R_{ref} in (46). Furthermore, the change of the covariances in (45) should not be thought of as a coordinate transformation due to the change of R_{ref}, as no quantities in the state ξ are expressed in the reference frame represented by R_{ref}. Instead, components of ξ may be expressed in the attitude defined by the composition of R_{ref} and δ, which remains unchanged during the reset operation (to first order in the variable δ).

V. Resulting algorithm

One may now construct an Extended/Unscented Kalman filter, similarly to the MEKF or USQUE, but incorporating the attitude reset steps. This proceeds as follows: The usual EKF (or UKF) prediction step is applied to compute the *a priori* mean and covariance of the state variable

x using (12) (and the linearization thereof, for an EKF). Importantly, the reference attitude is left unchanged. A first attitude reset is now performed, using Theorem 1, so that the attitude component of the estimator state equals zero, and the covariance is adjusted.

Measurement information is then incorporated using the (potentially linearized) measurement equation (13), where again the reference attitude is held constant. Finally, another attitude reset is performed according to Theorem 1.

Note that for an EKF, for certain systems (e.g. if the state to be estimated is limited to an attitude and a rate gyro bias), the first attitude reset step may be analytically combined with the *a-priori* prediction step, reducing the number of required matrix multiplications in an implementation.

Remark. Compared to the Invariant Extended Kalman Filter (IEKF) of [7] fewer stochastic properties of the filter can be asserted (and indeed, the proposed filter maintains all caveats typically associated with the EKF/UKF, including the possibility of estimator divergence). However, the proposed filter does not require the additional assumptions required for the IEKF, and retains the extended/unscented Kalman filters' ease of implementation and extensibility.

A. Comparison to the MEKF

The presented algorithm changes the three-element attitude states δ during the dynamic prediction and measurement update steps, and changes the reference attitude only during the reset steps. The MEKF, on the other hand, varies the reference attitude during the prediction step whilst keeping the three-element attitude states zero. It can be shown that it thereby implicitly encodes a reset during its prediction, however it lacks the reset step required after the measurement update. The MEKF's implicit prediction reset may be hard to generalise to systems with more complicated dynamics (compared to the straight-forward partial derivatives required for the presented algorithm).

Despite the lack of measurement reset, the MEKF has been successfully used in multiple applications (e.g. [20–24]). For systems with predominantly isotropic noise sources, dynamics, and measurements, the attitude covariance is likely to be approximately diagonal with equal eigenvalues, such that the covariance reset (19) has little effect. Furthermore, for a linear Kalman filter, the innovation sequence is white and zero mean [28]. For a well-tuned MEKF, for a system with "weak" nonlinearities, the covariance resets will thus be approximately zero mean, reducing the effect of the omission of the reset after the measurement update.

VI. Conclusion

The presented algorithm is a generic framework for applying the extended Kalman filter to systems whose dynamic state includes an attitude. The algorithm contains a correction to the Multiplicative Extended Kalman Filter (MEKF) of [12] and the Unscented Quaternion Estimator (USQUE) of [13] (in the form of the reset step), and generalises the MEKF to a broader class of systems. A three element attitude error is used in the estimate state vector (which has an associated covariance), in addition to a reference attitude (which does not have associated covariance). Monte Carlo sampling is used to validate the reset step. An extended/unscented Kalman filter may then easily be constructed for a system whose state includes an attitude, by refactoring the dynamics and measurement equations to be in terms of the attitude deviation, and applying a reset after the usual prediction and measurement steps. The reset step may also be extended straight-forwardly to problems containing multiple attitudes, e.g. a robotic arm with multiple serial joints.

Acknowledgments

The authors would like to thank Mark Psiaki and Landis Markley for their feedback and helpful comments to this paper. This research was supported by the Swiss National Science Foundation (SNSF), under grant application 138112.

References

[1] Simon, D., *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*, John Wiley & Sons, 2006. doi:10.1002/0470045345, pp. 395–458.

[2] Daum, F., "Nonlinear filters: beyond the Kalman filter," *IEEE Aerospace and Electronic Systems Magazine*, Vol. 20, No. 8, 2005, pp. 57–69. doi:10.1109/MAES.2005.1499276.

[3] Psiaki, M. L., "Backward-smoothing extended Kalman filter," *Journal of guidance, control, and dynamics*, Vol. 28, No. 5, 2005, pp. 885–894. doi:10.2514/1.12108.

[4] Julier, S. J. and LaViola Jr, J. J., "On Kalman filtering with nonlinear equality constraints," *IEEE Transactions on Signal Processing*, Vol. 55, No. 6, 2007, pp. 2774–2784. doi:10.1109/TSP.2007.893949.

[5] Zanetti, R., Majji, M., Bishop, R. H., and Mortari, D., "Norm-constrained Kalman filtering," *Journal of Guidance, Control, and Dynamics*, Vol. 32, No. 5, 2009, pp. 1458–1465. doi:10.2514/1.43119.

[6] Bonnabel, S., Martin, P., and Rouchon, P., "Non-linear symmetry-preserving observers on Lie groups," *IEEE Transactions on Automatic Control*, Vol. 54, No. 7, 2009, pp. 1709–1713. doi: 10.1109/TAC.2009.2020646.

[7] Barrau, A. and Bonnabel, S., "Intrinsic filtering on Lie groups with applications to attitude estimation," *IEEE Transactions on Automatic Control*, Vol. 60, No. 2, 2015, pp. 436–449. doi: 10.1109/TAC.2014.2342911.

[8] Lee, T., Leok, M., and McClamroch, N. H., "Global symplectic uncertainty propagation on SO (3)," *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, IEEE, 2008, pp. 61–66. doi: 10.1109/CDC.2008.4739058.

[9] Grip, H. F., Fossen, T. I., Johansen, T. A., and Saberi, A., "Globally exponentially stable attitude and gyro bias estimation with application to GNSS/INS integration," *Automatica*, Vol. 51, 2015, pp. 158–166. doi:10.1016/j.automatica.2014.10.076.

[10] Farrell, J. L., "Attitude determination by Kalman filtering," *Automatica*, Vol. 6, No. 3, 1970, pp. 419–430. doi:10.1016/0005-1098(70)90057-9.

[11] Lefferts, E. J., Markley, F. L., and Shuster, M. D., "Kalman filtering for spacecraft attitude estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 5, No. 5, 1982, pp. 417–429. doi:10.2514/6.1982-70.

[12] Markley, F. L., "Attitude error representations for Kalman filtering," *Journal of Guidance, Control, and Dynamics*, Vol. 26, No. 2, 2003, pp. 311–317. doi:10.2514/2.5048.

[13] Crassidis, J. L. and Markley, F. L., "Unscented filtering for spacecraft attitude estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 26, No. 4, 2003, pp. 536–542. doi:10.2514/6.2003-5484.

[14] Markley, F. L., Crassidis, J. L., and Cheng, Y., "Nonlinear attitude filtering methods," *AIAA Guidance, Navigation, and Control Conference*, 2005, pp. 15–18. doi:10.2514/6.2005-5927.

[15] Crassidis, J. L., Markley, F. L., and Cheng, Y., "Survey of nonlinear attitude estimation methods," *Journal of Guidance, Control, and Dynamics*, Vol. 30, No. 1, 2007, pp. 12–28. doi:10.2514/1.22452.

[16] Markley, F. L. and Crassidis, J. L., "Filtering for Attitude Estimation and Calibration," *Fundamentals of Spacecraft Attitude Determination and Control*, Springer, 2014, pp. 235–285. doi:10.1007/978-1-4939-0802-8.

[17] Steffes, S., Steinbach, J. P., and Theil, S., "Investigation of the Attitude Error Vector Reference Frame in the INS EKF," *Advances in Aerospace Guidance, Navigation and Control*, Springer, 2011, pp. 345–357. doi:10.1007/978-3-642-19817-5_27.

[18] Reynolds, R. G., "Asymptotically optimal attitude filtering with guaranteed convergence," *Journal of Guidance, Control, and Dynamics*, Vol. 31, No. 1, 2008, pp. 114–122. doi:10.2514/1.30381.

[19] Markley, F. L., "Lessons Learned," *The Journal of the Astronautical Sciences*, Vol. 57, No. 1-2, 2009, pp. 3–29. doi:10.1007/BF03321491.

[20] Trawny, N. and Roumeliotis, S. I., "Indirect Kalman filter for 3D attitude estimation," *University of Minnesota, Deptartment of Computer Science & Engineering, Report 2005-002*, January 2005.

[21] Hall, J. K., Knoebel, N. B., and McLain, T. W., "Quaternion attitude estimation for miniature air vehicles using a multiplicative extended Kalman filter," *IEEE/ION Position, Location and Navigation Symposium*, IEEE, 2008, pp. 1230–1237. doi:10.1109/PLANS.2008.4570043.

[22] Truch, M. D., Ade, P. A., Bock, J. J., Chapin, E. L., Devlin, M. J., Dicker, S. R., Griffin, M., Gundersen, J. O., Halpern, M., Hargrave, P. C., et al., "The Balloon-borne Large Aperture Submillimeter Telescope (BLAST) 2006: Calibration and Flight Performance," *The Astrophysical Journal*, Vol. 707, No. 2, 2009, pp. 1723–1728. doi:10.1088/0004-637X/707/2/1723.

[23] Springmann, J. C., Sloboda, A. J., Klesh, A. T., Bennett, M. W., and Cutler, J. W., "The attitude determination system of the RAX satellite," *Acta Astronautica*, Vol. 75, 2012, pp. 120–135. doi: 10.1016/j.actaastro.2012.02.001.

[24] Filipe, N., Kontitsis, M., and Tsiotras, P., "Extended Kalman Filter for Spacecraft Pose Estimation Using Dual Quaternions," *Journal of Guidance, Control, and Dynamics*, Vol. 38, 2015, pp. 1625–1641. doi:10.2514/1.G000977.

[25] Shuster, M. D., "A survey of attitude representations," *Journal of the Astronautical Sciences*, Vol. 41, No. 4, 1993, pp. 439–517.

[26] Anderson, B. D. O. and Moore, J. B., *Optimal filtering*, Prentice-Hall, 1979, pp. 195–205.

[27] Julier, S. J. and Uhlmann, J. K., "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, Vol. 92, No. 3, 2004, pp. 401–422. doi:10.1109/JPROC.2003.823141.

[28] Bar-Shalom, Y., Li, X. R., and Kirubarajan, T., *Estimation with applications to tracking and navigation*, John Wiley & Sons, 2004. doi:10.1002/0471221279, pp. 213.