

Quadratic Extended and Unscented Kalman Filter Updates

Simone Servadio, *Iowa State University, Ames, IA, USA*
 Chiran. B. Cherian, *Iowa State University, Ames, IA, USA*

Abstract—Common filters are usually based on the linear approximation of the optimal minimum mean square error estimator. The Extended and Unscented Kalman Filters handle nonlinearity through linearization and unscented transformation, respectively, but remain linear estimators, meaning that the state estimate is a linear function of the measurement. This paper proposes a quadratic approximation of the optimal estimator, creating the Quadratic Extended and Quadratic Unscented Kalman Filter. These retain the structure of their linear counterpart, but include information from the measurement square to obtain a more accurate estimate. Numerical results show the benefits in accuracy of the new technique, which can be generalized to upgrade other linear estimators to their quadratic versions.

Index Terms—Extended Kalman Filter, Unscented Kalman Filter, Quadratic Update, Polynomial Update, Clohessy–Wiltshire Equations, Minimum Mean Square Error Estimation

I. INTRODUCTION

For linear, Gaussian systems, the posterior remains Gaussian, and the Kalman Filter [1] provides an exact, recursive mechanism to compute its mean and covariance. However, most systems exhibit nonlinear dynamics and/or measurements, leading to non-Gaussian probability density functions (PDFs). A common way to handle such nonlinearities is to linearize the system about the current estimate. The Extended Kalman Filter (EKF) [2] applies the Kalman mechanization to this linearized model. Despite its simplicity, the EKF often struggles to converge when nonlinearities are severe [3].

An alternative is the Unscented Kalman Filter (UKF) [4], which uses a deterministic sampling to more accurately approximate the predicted mean and covariance. The unscented transformation approximates distributions as Gaussians and applies the nonlinear transformation of a set of well-selected sigma points. The UKF typically achieves higher accuracy and robustness than the EKF, since it accounts for higher-order terms in the nonlinear maps without explicitly computing Jacobians. Nevertheless, both EKF and UKF remain fundamentally linear estimators: at each step, the updated state is a linear function of the measurement. Both filters are derived from the linearized Minimum Mean Square Error (MMSE) principle, and they can be seen as different linear approximations of the conditional mean. Other “linear” filters approximate the nonlinear dynamics and measurement equation in different ways: Cubature Kalman filters (CKF) [5], Ensemble Kalman filters

(EnKF) [6], [7], Central Difference Kalman filters (CDKF) [8]. All of these still produce a linear measurement update, leaving the true MMSE estimator, which is generally nonlinear and intractable, unrealized. Even the use of State Transition Tensors (STT) [9] and, originally, Differential Algebra [10] have been used to improve the accuracy of the prediction step of the filters, on how they propagate and transform uncertainties, leaving the update structure untouched.

To go beyond linear estimators, one can consider Gaussian Sum Filters (GSF) [11], [12], based on the concept of representing nongaussian distributions as a sum of multiple Gaussian components, or particle filters (PF), such as the bootstrap PF (BPF) [13], the Gaussian PF (GPF) [14], and other PF based on Sequential Importance Sampling (SIS) [15]. However, these techniques leave the Kalman formulation to achieve an accurate estimate, based on the representation of PDFs via Gaussian Multiple Models (GMM) or as probability mass functions (PMF), yielding superior accuracy at the cost of significantly higher computational load.

A less explored route is to retain a linear update structure but replace the linear dependence on the measurement with a polynomial function. A nonlinear update within the Kalman structure has been offered in [16], augmenting the measurement with its square to form a quadratic update, and this idea was extended to arbitrary polynomial orders [17]. This formulation requires carrying estimates of the state squares, and it presents flaws when working with vector measurements [16]. Indeed, the polynomial update presented in [18] has been tested only with scalar measurements, avoiding the issue of matrix singularity. Moreover, its derivation requires lifting the nonlinear system via a Carleman approximation, creating large and cumbersome matrices.

In more recent studies, a different quadratic update (extendable to a polynomial update of any order) based on Taylor series expansions has been derived [19]. It requires knowledge of high-order central moments, calculated in the Differential Algebra (DA) framework, and carried at each time step, likewise the EKF carries mean and covariance. Further development [20] performs a polynomial update without carrying the higher-order central moments and, thus, reduces the overall computational cost. In [20], non-Gaussian distributions are approximated as a polynomial transformation of Gaussian random variables, where moments are evaluated efficiently in DA using Isserlis’ formulation [21]. Lastly, the polynomial update has been expanded to the estimation of the conditional covariance, applying a double update sequentially for mean and covariance [22].

Dr. Simone Servadio, Assistant Professor, Department of Aerospace Engineering, servadio@iastate.edu

Mr. Chiran Cherian, Graduate Student, Department of Aerospace Engineering, cbckbc@iastate.edu

This paper offers the derivation of two estimators based on the quadratic approximation of the MMSE, achieved via linearization and through the unscented transformation. The proposed Quadratic Extended Kalman Filter (QEKF) and Quadratic Unscented Kalman Filter (QUKF) closely resemble the EKF and the UKF structure, but they apply augmented covariances to achieve a nonlinear measurement update, where the state estimate is a parabolic function of the measurement. Numerical results show how the newly proposed nonlinear estimators outperform their linear counterparts.

II. THE OPTIMAL QUADRATIC ESTIMATOR

Previous works [19], [20], [22] derived a new series of polynomial estimators where the state estimate is a polynomial function of the given measurement. They make use of Differential Algebra techniques to calculate high-order central moments. In this work, we obtain the quadratic estimator from the definition of the orthogonality principle, but apply linearization and the unscented transformation to obtain the high-order gains and terms, so that the final results resemble the structure of the Extended and Unscented Kalman Filters, but with the quadratic term.

Minimum Mean Square Error (MMSE) estimators aim at minimizing the expected value of the error square; thus, their formulation leads to unbiased filters where the mean of the error square coincides with the error covariance. Whenever a linear function in the measurement is used, as a family of estimators, we obtain a linear representation of the MMSE, the LMMSE. These estimators have the form of

$$g_L(\mathbf{y}) = A + \mathbb{E}[\mathbf{x}] + B(\mathbf{y} - \mathbb{E}[\mathbf{y}]) \quad (1)$$

and are very common. Equation (1) is used by the EKF, the UKF, and other linear estimators in their state update step. Their estimate changes according to the technique used to evaluate the constants A and B , dealing with the nonlinearities of the measurement model.

This work applies a quadratic approximation to the MMSE, the QMMSE. Thus, the generic quadratic estimator family is

$$g_Q(\mathbf{y}) = A + \mathbb{E}[\mathbf{x}] + B\delta\mathbf{y} + C\delta\mathbf{y}^{[2]} \quad (2)$$

where

$$\delta\mathbf{y} = \mathbf{y} - \mathbb{E}[\mathbf{y}] \quad (3)$$

$$\delta\mathbf{y}^{[2]} = \delta\mathbf{y} \otimes \delta\mathbf{y} = (\mathbf{y} - \mathbb{E}[\mathbf{y}]) \otimes (\mathbf{y} - \mathbb{E}[\mathbf{y}]) \quad (4)$$

indicates the residual (or measurement deviation vector) and its square, obtained via the Kronecker product \otimes for the square of vectors.

The orthogonality principle states that, given the optimal estimator $g_Q^*(\mathbf{y})$, with optimal constants A^* , B^* , and C^* , the equation

$$\mathbb{E}[(\mathbf{x} - g_Q^*(\mathbf{y})) g_Q(\mathbf{y})^T] = 0 \quad (5)$$

holds for every other generic estimator. Therefore, picking

$$A = -\mathbb{E}[\mathbf{x}] + \mathbf{I}, \quad B = \mathbf{0}, \quad C = \mathbf{0} \quad (6)$$

where \mathbf{I} is the identity matrix, leads to

$$\begin{aligned} \mathbb{E}[\delta\mathbf{x} - A^* - B^*\delta\mathbf{y} - C^*\delta\mathbf{y}^{[2]}] &= \\ A^* &= -C^* \mathbf{v}(\mathbf{P}_{yy}). \end{aligned} \quad (7)$$

where $\mathbf{v}(\mathbf{P}_{yy})$ indicates the stack operator for matrix \mathbf{P}_{yy} . In a similar way, selecting the two sets

$$A = -\mathbb{E}[\mathbf{x}], \quad B = \mathbf{I}, \quad C = \mathbf{0} \quad (8)$$

$$A = -\mathbb{E}[\mathbf{x}], \quad B = \mathbf{0}, \quad C = \mathbf{I} \quad (9)$$

leads to

$$\begin{aligned} \mathbb{E}[(\delta\mathbf{x} - A^* - B^*\delta\mathbf{y} - C^*\delta\mathbf{y}^{[2]}) \delta\mathbf{y}^T] &= \\ \mathbf{P}_{xy} - B^*\mathbf{P}_{yy} - C^*\mathbf{P}_{y^{[2]}y} &= 0 \end{aligned} \quad (10)$$

and

$$\begin{aligned} \mathbb{E}[(\delta\mathbf{x} - A^* - B^*\delta\mathbf{y} - C^*\delta\mathbf{y}^{[2]}) \delta\mathbf{y}^{[2]T}] &= \\ \mathbf{P}_{xy^{[2]}} - B^*\mathbf{P}_{yy^{[2]}} + C^*(\mathbf{v}(\mathbf{P}_{yy})\mathbf{v}(\mathbf{P}_{yy})^T - \mathbf{P}_{y^{[2]}y^{[2]}}) &= 0 \end{aligned} \quad (11)$$

where we have defined covariances as

$$\mathbf{P}_{xy} = \mathbb{E}[\delta\mathbf{x} \delta\mathbf{y}^T], \quad (12)$$

$$\mathbf{P}_{xy^{[2]}} = \mathbb{E}[\delta\mathbf{x} \delta\mathbf{y}^{[2]T}], \quad (13)$$

$$\mathbf{P}_{yy} = \mathbb{E}[\delta\mathbf{y} \delta\mathbf{y}^T], \quad (14)$$

$$\mathbf{P}_{yy^{[2]}} = \mathbb{E}[\delta\mathbf{y} \delta\mathbf{y}^{[2]T}], \quad (15)$$

$$\mathbf{P}_{y^{[2]}y} = \mathbb{E}[\delta\mathbf{y}^{[2]} \delta\mathbf{y}^T], \quad (16)$$

$$\mathbf{P}_{y^{[2]}y^{[2]}} = \mathbb{E}[\delta\mathbf{y}^{[2]} \delta\mathbf{y}^{[2]T}]. \quad (17)$$

with

$$\delta\mathbf{x} = \mathbf{x} - \mathbb{E}[\mathbf{x}] \quad (18)$$

After substituting in Eq. (7), Eqs (10) and (11) can be rewritten as a set of coupled equations.

$$\begin{bmatrix} B^* & C^* \end{bmatrix} \begin{bmatrix} \mathbf{P}_{yy} & \mathbf{P}_{yy^{[2]}} \\ \mathbf{P}_{y^{[2]}y} & \mathbf{P}_{y^{[2]}y^{[2]}} - \mathbf{v}(\mathbf{P}_{yy})\mathbf{v}(\mathbf{P}_{yy})^T \end{bmatrix} = \begin{bmatrix} \mathbf{P}_{xy} \\ \mathbf{P}_{xy^{[2]}} \end{bmatrix}^T \quad (19)$$

which brings to the solution for the optimal constants

$$\mathbf{K} = \begin{bmatrix} B^* & C^* \end{bmatrix} = \mathbf{P}_{xy} \mathbf{P}_{yy}^{-1} \quad (20)$$

having defined

$$\mathbf{P}_{xy} = \begin{bmatrix} \mathbf{P}_{xy} & \mathbf{P}_{xy^{[2]}} \end{bmatrix} \quad (21)$$

$$\mathbf{P}_{yy} = \begin{bmatrix} \mathbf{P}_{yy} & \mathbf{P}_{yy^{[2]}} \\ \mathbf{P}_{y^{[2]}y} & \mathbf{P}_{y^{[2]}y^{[2]}} - \mathbf{v}(\mathbf{P}_{yy})\mathbf{v}(\mathbf{P}_{yy})^T \end{bmatrix} \quad (22)$$

The optimal constants B^* and C^* behave likewise the classic Kalman gain, but evaluated in its augmented form that considers the covariances and cross covariances with the square of the measurement deviation vector.

Substituting the derived constants from Eq. (7) and (20) back into the definition of the quadratic estimator, it creates the optimal quadratic estimator, which is the parabolic approximation of the true MMSE.

$$g_Q^*(\mathbf{y}) = \mathbb{E}[\mathbf{x}] + \mathbf{P}_{xy} \mathbf{P}_{yy}^{-1} \begin{bmatrix} \delta\mathbf{y} \\ \delta\mathbf{y}^{[2]} - \mathbf{v}(\mathbf{P}_{yy}) \end{bmatrix}. \quad (23)$$

The derived estimator is applied in the update step of the classic Extended Kalman Filter (EKF) and Unscented Kalman Filter (UKF) to obtain the quadratic version of their formulation, which is linear. Therefore, using linearization to evaluate the set of central moments in Eqs (12) to (17), we propose the Quadratic Extended Kalman Filter (QEKF) and, using the unscented transformation to evaluate moments, the Quadratic Unscented Kalman Filter (UKF).

III. HIGH ORDER CENTRAL MOMENTS

Linear estimators require the first two central moments, i.e., mean and covariance, to operate. On the contrary, quadratic estimators, due to their higher order, require information up to the fourth central moment to achieve a correct and consistent estimate. Therefore, we define with

$$\mathbf{S}_{xxx} = \mathbb{E}[(x - \hat{x}) \otimes ((x - \hat{x})(x - \hat{x})^T)] \quad (24)$$

the skewness of a distribution, expressed as a three-dimensional tensor, and with

$$\mathbf{K}_{xxxx} = \mathbb{E}[((x - \hat{x})(x - \hat{x})^T)^T \otimes ((x - \hat{x})(x - \hat{x})^T)] \quad (25)$$

the kurtosis, expressed as a forth dimensional tensor. It can be shown [20], [22] that, for a null mean distribution, where central moments and raw moments match, $\mathbf{S}_{xxx} = \mathbf{P}_{x[2]x} = \mathbf{P}_{xx[2]}^T$ and $\mathbf{K}_{xxxx} = \mathbf{P}_{x[2]x[2]}$.

IV. THE QUADRATIC EXTENDED KALMAN FILTER (QEKF)

The formulation of the EKF is well-known: it linearizes the dynamic and measurement model at the current estimate and it applies the Kalman Filter algorithm *as-if* everything is linear and Gaussian. It represents distributions as Gaussians, propagating and updating, iteratively, the state mean, $\hat{\mathbf{x}}$, and covariance, \mathbf{P}_{xx} .

Consider the nonlinear equation of motion, with dynamics $f(\cdot)$, affected by noise

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}) + \boldsymbol{\mu}_{k-1} \quad (26)$$

where $\boldsymbol{\mu}_{k-1}$ is a zero-mean Gaussian process noise with known covariance matrix $\mathbf{P}_{\mu\mu}$. According to the EKF implementation, the prediction step propagates mean and covariance according to

$$\hat{\mathbf{x}}_k^- = f(\hat{\mathbf{x}}_{k-1}^+) \quad (27)$$

$$\mathbf{P}_{xx,k}^- = \mathbf{F}_{k-1} \mathbf{P}_{xx,k-1}^+ \mathbf{F}_{k-1}^T + \mathbf{P}_{\mu\mu} \quad (28)$$

$$\mathbf{F}_{k-1} = \left. \frac{\partial f}{\partial \mathbf{x}} \right|_{\hat{\mathbf{x}}_{k-1}^+} \quad (29)$$

where $\hat{\mathbf{x}}_{k-1}^+$ is the state mean at the previous time step, $\hat{\mathbf{x}}_k^-$ is the current predicted mean of the distribution, \mathbf{P}_{xx}^- indicates the predicted (prior) covariance, and \mathbf{F}_k is the gradient of the dynamics.

The state PDF is now approximated as a Gaussian with known mean and covariance, which constitutes the prior for the Bayesian update. An observation is provided according to the measurement model $h(\cdot)$,

$$\mathbf{y} = h(\mathbf{x}) + \boldsymbol{\eta} \quad (30)$$

affected by zero-mean measurement noise with known covariance $\mathbf{P}_{\eta\eta}$, skewness $\mathbf{S}_{\eta\eta\eta}$, and kurtosis $\mathbf{K}_{\eta\eta\eta\eta}$. The predicted measurement mean is

$$\hat{\mathbf{y}}_k = h(\hat{\mathbf{x}}_k^-) \quad (31)$$

Given the measurement outcome $\tilde{\mathbf{y}}$, the actual numerical value of the random value from the sensors, the residual and its square are evaluated as

$$\delta \tilde{\mathbf{y}} = \mathbf{y} - \hat{\mathbf{y}}_k \quad (32)$$

$$\delta \tilde{\mathbf{y}}^{[2]} = \delta \tilde{\mathbf{y}} \otimes \delta \tilde{\mathbf{y}} \quad (33)$$

so that the update step of the QEKF can be performed according to

$$\mathbf{K} = \mathbf{P}_{xy} \mathbf{P}_{yy}^{-1} \quad (34)$$

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K} \begin{bmatrix} \delta \mathbf{y} \\ \delta \mathbf{y}^{[2]} - \mathbf{v}(\mathbf{P}_{yy}) \end{bmatrix} \quad (35)$$

$$\mathbf{P}_{xx,k}^+ = \mathbf{P}_{xx,k}^- - \mathbf{K} \mathbf{P}_{yy} \mathbf{K}^T \quad (36)$$

where \mathbf{K} is the Kalman gain, in its augmented form, and $\hat{\mathbf{x}}_k^+$ and $\mathbf{P}_{xx,k}^+$ are the updated mean and covariance of the state distribution.

The reader can notice how Eqs (35) and (36) resemble the structure of the normal Kalman update, but expanded to operate with the augmented covariances that include information of the square of the measurement vector. Indeed, the state-measurement cross-covariance is evaluated block-wise as in

$$\mathbf{P}_{xy} = \begin{bmatrix} \mathbf{P}_{xy} & \mathbf{P}_{xy[2]} \end{bmatrix} \quad (37)$$

with

$$\mathbf{P}_{xy} = \mathbf{P}_{xx,k}^- \mathbf{H}^T \quad (38)$$

$$\mathbf{P}_{xy[2]} = \mathbf{S}_{xxx} \mathbf{H}^{[2]T} \quad (39)$$

$$\mathbf{H}_k = \left. \frac{\partial h}{\partial \mathbf{x}} \right|_{\hat{\mathbf{x}}_k^-} \quad (40)$$

where $\mathbf{H}^{[2]} = \mathbf{H} \otimes \mathbf{H}$ squares the measurement model. In a similar pattern, the augmented measurement covariance is evaluated block-wise as in

$$\mathbf{P}_{yy} = \begin{bmatrix} \mathbf{P}_{yy} & \mathbf{P}_{yy[2]} \\ \mathbf{P}_{yy[2]}^T & \mathbf{P}_{yy[2]y[2]} - \mathbf{v}(\mathbf{P}_{yy})\mathbf{v}(\mathbf{P}_{yy})^T \end{bmatrix} \quad (41)$$

with

$$\overline{\mathbf{P}}_{yy} = \mathbf{H} \mathbf{P}_{xx,k}^- \mathbf{H}^T \quad (42)$$

$$\mathbf{P}_{yy} = \overline{\mathbf{P}}_{yy} + \mathbf{P}_{\eta\eta} \quad (42)$$

$$\mathbf{P}_{yy[2]} = \mathbf{H} \mathbf{S}_{xxx} \mathbf{H}^{[2]T} + \mathbf{S}_{\eta\eta\eta} \quad (43)$$

$$\begin{aligned} \mathbf{P}_{yy[2]} &= \mathbf{H}^{[2]} \mathbf{K}_{xxxx} \mathbf{H}^{[2]T} + \mathbf{K}_{\eta\eta\eta\eta} + \\ &+ \overline{\mathbf{P}}_{yy} \otimes \mathbf{P}_{\eta\eta} + \mathbf{m}(\mathbf{v}(\overline{\mathbf{P}}_{yy}) \otimes \mathbf{I}) (\mathbf{P}_{\eta\eta}^T \otimes \mathbf{I}) + \\ &+ \mathbf{P}_{\eta\eta} \otimes \overline{\mathbf{P}}_{yy} + \mathbf{m}(\mathbf{v}(\mathbf{P}_{\eta\eta}) \otimes \mathbf{I}) (\overline{\mathbf{P}}_{yy}^T \otimes \mathbf{I}) + \\ &+ \mathbf{H}^{[2]} \mathbf{v}(\mathbf{P}_{xx}) \mathbf{v}(\mathbf{P}_{\eta\eta})^T + \mathbf{v}(\mathbf{P}_{\eta\eta}) \mathbf{v}(\mathbf{P}_{xx})^T \mathbf{H}^{[2]T} \end{aligned} \quad (44)$$

where $\mathbf{m}(\cdot)$ is the matrix operator, inverse of the stack operator $\mathbf{v}(\cdot)$, and where $\overline{\mathbf{P}}_{yy}$ indicates the measurement covariance without any influence from the noise, the actual uncertainty transformation.

These expressions of the covariances are obtained by applying the expected value operator as described in Eqs (12) to (17) with the linearized model of the measurement equation. Full knowledge of the moments of the noise is assumed, while the state prior skewness and kurtosis can be approximated in a way that better fits the application. For a simple Gaussian assumption, each entry of the tensors can follow the Isserlis' formulation [21]:

$$\mathbf{S}_{ijk} = 0 \quad (45)$$

$$\mathbf{K}_{ijkm} = \mathbf{P}_{ij}\mathbf{P}_{km} + \mathbf{P}_{ik}\mathbf{P}_{jm} + \mathbf{P}_{im}\mathbf{P}_{jk} \quad (46)$$

The state has been updated according to Eqs (35) and (36) and the mean and covariance can undergo a new propagation normally. If propagation is of importance, a higher accuracy prediction step such as DA or STT can be applied to obtain the prior state distribution, followed by the quadratic update proposed here.

V. THE QUADRATIC UNSCENTED KALMAN FILTER (QUKF)

The quadratic estimator derived in Section II is generic and can be applied to multiple linear filters, to expand the technique that deals with the system and equation nonlinearity in the prediction to achieve a nonlinear update. In this Section, we exploit the unscented transformation to obtain the QUKF. The unscented transformation (UT) is applied to obtain the mean and covariance of a random variable, \mathbf{y} , which has a known relation, $h(\cdot)$, with respect to the given initial random variable \mathbf{x} , of dimension n . After defining the spread parameter α , the scaling parameter κ , and the constant β , a set of $2n + 1$ sigma points is created from the current state mean and covariance:

$$\lambda = \alpha^2(n + \kappa) - n \quad (47)$$

$$\mathbf{C}_{k-1}\mathbf{C}_{k-1}^T = (n + \lambda)\mathbf{P}_{xx,k-1}^+ \quad (48)$$

$$\mathbf{x}_{k-1} = [\hat{\mathbf{x}}_{k-1}^+ \quad \hat{\mathbf{x}}_{k-1}^+ + \mathbf{C}_{k-1} \quad \hat{\mathbf{x}}_k^+ - \mathbf{C}_{k-1}] \quad (49)$$

Each sigma point is associated with a weight, given by

$$w_0^{(m)} = \frac{\lambda}{n + \lambda} \quad (50)$$

$$w_0^{(c)} = \frac{\lambda}{n + \lambda} + (1 - \alpha^2 + \beta) \quad (51)$$

$$w_i^{(m)} = w_i^{(c)} = \frac{1}{2(n + \lambda)}, \quad i = 1, \dots, 2n \quad (52)$$

for the calculation of means, $w_i^{(m)}$, and covariances, $w_i^{(c)}$.

The prediction step is performed by propagating, separately, each sigma point and obtaining their weighted means for the predicted estimate and covariance.

$$\mathbf{x}_k^{(i)} = f(\mathbf{x}_{k-1}^{(i)}) \quad \forall i = 0, \dots, 2n \quad (53)$$

$$\hat{\mathbf{x}}_k^- = \sum_{i=0}^{2n} w_m^{(i)} \mathbf{x}_k^{(i)} \quad (54)$$

$$\mathbf{P}_{xx,k}^- = \sum_{i=0}^{2n} w_c^{(i)} (\mathbf{x}_k^{(i)} - \hat{\mathbf{x}}_k^-)(\mathbf{x}_k^{(i)} - \hat{\mathbf{x}}_k^-)^T + \mathbf{P}_{\mu\mu} \quad (55)$$

The quadratic update for the QUKF follows the same formulation as in Eqs (35) and (36), with the augmented covariances reported in Eqs (37) and (41), evaluated block-wise. However, each entry is computed via the sigma points rather than with linearization. Therefore, after obtaining the transformed sigma points in the measurement space and their predicted mean

$$\mathbf{y}_k^{(i)} = h(\mathbf{x}_k^{(i)}) \quad \forall i = 0, \dots, 2n \quad (56)$$

$$\hat{\mathbf{y}}_k = \sum_{i=0}^{2n} w_m^{(i)} \mathbf{y}_k^{(i)} \quad (57)$$

each sigma point is augmented with its Kronecker square. Thus, the deviation vectors of the state and the measurement are computed as

$$\delta \mathbf{x}_k^{(i)} = \mathbf{x}_k^{(i)} - \hat{\mathbf{x}}_k^- \quad \forall i = 0, \dots, 2n \quad (58)$$

$$\delta \mathbf{y}_k^{(i)} = \mathbf{y}_k^{(i)} - \hat{\mathbf{y}}_k \quad \forall i = 0, \dots, 2n \quad (59)$$

$$\delta \mathbf{y}_k^{(i)} = \delta \mathbf{y}_k^{(i)} \otimes \delta \mathbf{y}_k^{(i)} \quad \forall i = 0, \dots, 2n \quad (60)$$

This definition eases the evaluation of the covariances, which are obtained as simple summation. The state-measurement cross covariances components are

$$\mathbf{P}_{xy} = \sum_{i=0}^{2n} w_c^{(i)} \delta \mathbf{x}_k^{(i)} \delta \mathbf{y}_k^{(i)T} \quad (61)$$

$$\mathbf{P}_{xy^{[2]}} = \sum_{i=0}^{2n} w_c^{(i)} \delta \mathbf{x}_k^{(i)} \delta \mathbf{y}_k^{[2](i)T} \quad (62)$$

The measurement covariance is evaluated in two separate steps: first, the transformed covariances in the measurement state are computed, then the influence of the noise is added.

$$\overline{\mathbf{P}}_{yy} = \sum_{i=0}^{2n} w_c^{(i)} \delta \mathbf{y}_k^{(i)} \delta \mathbf{y}_k^{(i)T} \quad (63)$$

$$\overline{\mathbf{P}}_{yy^{[2]}} = \sum_{i=0}^{2n} w_c^{(i)} \delta \mathbf{y}_k^{(i)} \delta \mathbf{y}_k^{[2](i)T} \quad (64)$$

$$\overline{\mathbf{P}}_{y^{[2]}y^{[2]}} = \sum_{i=0}^{2n} w_c^{(i)} \delta \mathbf{y}_k^{[2](i)} \delta \mathbf{y}_k^{[2](i)T} \quad (65)$$

where additive noise leads to a simple noise addition for the contribution:

$$\mathbf{P}_{yy} = \overline{\mathbf{P}}_{yy} + \mathbf{P}_{\eta\eta} \quad (66)$$

$$\mathbf{P}_{yy^{[2]}} = \overline{\mathbf{P}}_{yy^{[2]}} + \mathbf{S}_{\eta\eta\eta} \quad (67)$$

$$\begin{aligned} \mathbf{P}_{y^{[2]}y^{[2]}} &= \overline{\mathbf{P}}_{y^{[2]}y^{[2]}} + \mathbf{K}_{\eta\eta\eta\eta} + \\ &+ \overline{\mathbf{P}}_{yy} \otimes \mathbf{P}_{\eta\eta} + m(\mathbf{v}(\overline{\mathbf{P}}_{yy}) \otimes \mathbf{I}) (\mathbf{P}_{\eta\eta}^T \otimes \mathbf{I}) + \\ &+ \mathbf{P}_{\eta\eta} \otimes \overline{\mathbf{P}}_{yy} + m(\mathbf{v}(\mathbf{P}_{\eta\eta}) \otimes \mathbf{I}) (\overline{\mathbf{P}}_{yy}^T \otimes \mathbf{I}) + \\ &+ \mathbf{v}(\overline{\mathbf{P}}_{yy}) \mathbf{v}(\mathbf{P}_{\eta\eta})^T + \mathbf{v}(\mathbf{P}_{\eta\eta}) \mathbf{v}(\overline{\mathbf{P}}_{yy})^T \end{aligned} \quad (68)$$

The QUKF algorithm is complete, as it has the same structure as the QEKf, where the difference lies in how the central moments (mean and covariances) are evaluated. The higher-order moments provide additional information to the filter with respect to its linear version, drastically improving the accuracy of the measurement update.

VI. GENERALIZATION TO QUADRATIC UPDATES

A quick consideration about similarities and differences between the QEKF and the QUKF is due. The two filters have the same structure, similarly to how both the EKF and UKF follow the Kalman formulation. They create an augmented system that follows the same procedure regardless of the uncertainty transformation technique selected: linearization for the QEKF and unscented transformation for the QUKF.

Therefore, the proposed quadratic update fits any technique, and it can be applied to other linear estimators, such as the Cubature Kalman Filter or the Ensemble Kalman Filter, where it changes how Eqs (61) to (68) are evaluated, while keeping the same structure for the augmented squaring of the measurement update.

VII. NUMERICAL APPLICATIONS

The proposed algorithms have been applied to two different applications: first, the newly developed update techniques have been used in a simple scalar toy problem to visually highlight the benefits of the quadratic approach; secondly, a more complex relative navigation in space application with non-Gaussian measurement noises has been tested.

A. Scalar Problem

A scalar problem is formulated to verify the improvements of the quadratic update, enabling nonlinear estimation, over any linear filtering techniques, such as the EKF and the UKF. The problem will show how the quadratic representation (QMMSE) better approximates the true MMSE function, especially when compared to its linear counterpart, the LMMSE.

Define a prior state $x \sim \mathcal{N}(1, 0.05)$ and a measurement

$$y = \arctan(x) + \eta \quad (69)$$

where $\eta \sim \mathcal{N}(0, 0.01)$ is the measurement noise.

We then simulate the true joint distribution of x and y using 10^5 samples, reported in Fig. 1. The optimal (nonlinear) MMSE is the conditional mean, which visually is the curved line dividing in half the distribution of y (horizontal spread of grey points) for each value of x . Starting from a Gaussian distribution, the spread of the points follows the joint PDF (thus, the posterior) shape due to the nonlinearity of the measurement equation. The figure also shows different estimators in various colors: the EKF, UKF, QEKF, and QUKF. The EKF (black) and the UKF (red) are straight lines, whose slope is the Kalman gain. A more accurate transformation of central moments leads to a better prediction in the measurement space and, therefore, to a more reliable Kalman gain. Conceptually, the optimal Kalman gain obtainable is represented by the brown line, which represents the true LMMSE evaluated directly from the particles. Linear estimators, regardless of the technique that can be used to approximate moments, asymptotically aim at reaching this line. On the contrary, the quadratic estimators are given the capability of obtaining a curved function. In the figure, the QEKF overlaps the EKF. This is due to the Gaussian prior assumption, which leads to $P_{xy^2} = 0$, meaning the square term in the estimator function, ay^2 , has $a = 0$, and it reduces to the EKF. This happens because a

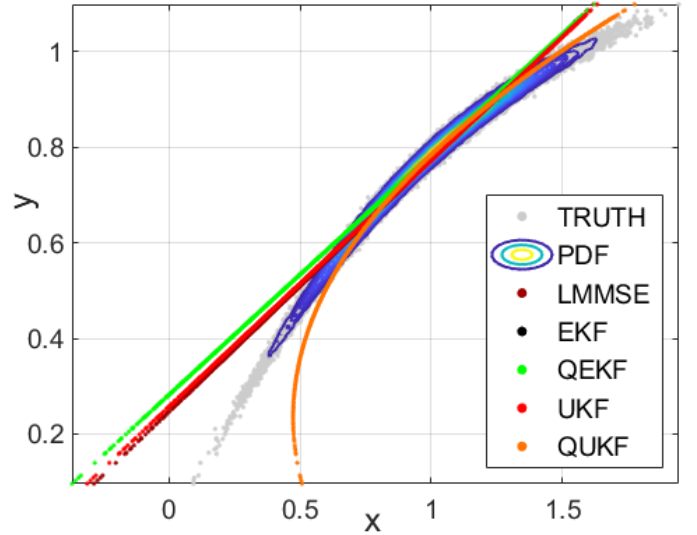


Fig. 1. Representation of the state-measurement joint distribution and of the relative estimates from different estimators.

full Gaussian assumption is applied, supported by the Isserlis' formula. On the contrary, the QUKF calculates P_{xy^2} directly from the sigma points, giving a non-null skewed behavior. The result is an estimator line that follows a parabolic trend in the variable y , as shown by the orange line. The QUKF estimate turns, and it follows the curved shape of the posterior distribution, achieving a more accurate estimate. The QUKF curved behavior improves both the representation of the tails of the distribution and the region around the mode of the posterior PDF.

The advantages of the quadratic approximation of the MMSE over the linear counterpart can be appreciated in Fig. 2, where the root mean square error has been evaluated as

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{x}_i - x_T)^2} \quad (70)$$

where x_T is the true value of x and N is the number of points representing the distribution. As previously mentioned, due to the Gaussianity assumption of the skewness, the QEKF reduces to the EKF, and they share the same error level. The UKF is more accurate than the EKF, improving accuracy drastically thanks to the unscented transformation being more reliable than mere linearization for the transformation of central moments. However, the UKF cannot surpass the limit given by the optimal LMMSE, which represents the most accurate linear estimator achievable. The LMMSE bar represents the best possible linear fit of the particles that compose the posterior distribution. On the other hand, the QUKF is able to better follow the shape of the distribution, and its error is much lower than the UKF, with an improvement of around 25% in accuracy. Ref [20] showed that, conceptually, by increasing the order of the polynomial update to cubic and higher, the accuracy keeps improving, until reaching, asymptotically, the true MMSE. The QUKF derived in this paper enhance to use of the unscented transformation to obtain a quadratic polynomial representation of the true MMSE.

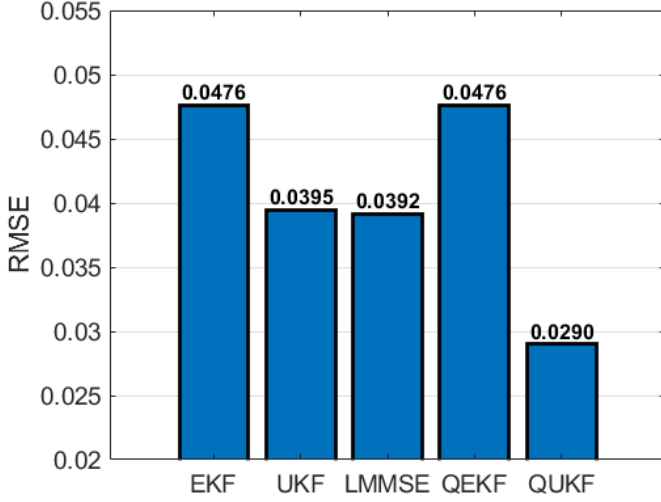


Fig. 2. RMSE among selected estimators

B. Clohessy–Wiltshire Relative Motion

The Clohessy–Wiltshire (CW) equations, also known as the Hill’s equations, describe the relative motion of a chaser spacecraft with respect to a target spacecraft in a circular orbit. Assuming the target is in a circular orbit and a local-vertical-local-horizontal (LVLH) coordinate frame centered on the target, the linearized equations of motion for the relative position vector $\mathbf{r} = [x, y, z]^T$ are given by:

$$\ddot{x} = 2n\dot{y} + 3n^2x \quad (71)$$

$$\ddot{y} = -2n\dot{x} \quad (72)$$

$$\ddot{z} = -n^2z \quad (73)$$

where n is the mean motion of the chief’s circular orbit, defined as $n = \sqrt{\mu/a^3}$, with μ the Earth’s standard gravitational parameter, and a the semi-major axis of the chief’s orbit, assumed to be 7000 km. These equations assume linearized dynamics, valid for small separations relative to the orbital radius. The initial relative PDF is assumed to be a Gaussian distribution with mean $\mathbf{x}_0 = [\mathbf{r}_0 \ \mathbf{v}_0]^T$ given by

$$\mathbf{r}_0 = [2 \ 10 \ -3.5]^T \text{ km} \quad (74)$$

$$\mathbf{v}_0 = [0.01 \ -0.005 \ 0.0005]^T \text{ km/s} \quad (75)$$

and covariance \mathbf{P}_0 defined as:

$$\mathbf{P}_0 = \text{blkdiag}(10^{-4}\mathbf{I}_3\text{km}^2, 10^{-9}\mathbf{I}_3\text{km/s}^2) \quad (76)$$

where $\text{blkdiag}()$ indicates the block diagonal operator and \mathbf{I}_3 a 3x3 identity matrix.

Relative angle measurements are acquired by the chief every minute for 3 hours according to

$$y_1 = \arctan(y/x) + \eta_1 \quad (77)$$

$$y_2 = \arcsin\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right) + \eta_2 \quad (78)$$

where η_i is a zero-mean non-Gaussian noise, whose distribution is described in Table I [17], [19].

TABLE I
MEASUREMENT NOISE DISTRIBUTION (RAD)

η_k	$1e-3$	$-3e-3$	$-9e-3$
$P(\eta_k)$	$\frac{15}{18}$	$\frac{2}{18}$	$\frac{1}{18}$

This application has been chosen since it presents a linear dynamic with nonlinear measurements. The linear system guarantees the correct uncertainty propagation of the state distribution, since a Gaussian PDF remains Gaussian after a linear transformation. Therefore, the difference and improvement in accuracy are due to the different order of the measurement update. Figure 3 shows the target’s relative motion with respect to the chief, creating a helical pattern around the origin whose amplitude is connected to the mean motion of the reference orbit.

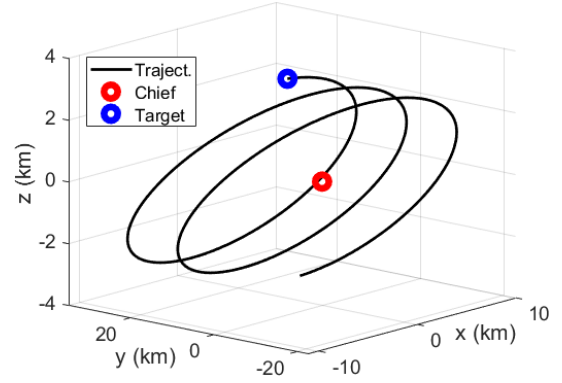


Fig. 3. Target’s relative motion with respect to the chief satellite.

The accuracy level and consistency of the filter are evaluated via a Monte Carlo analysis, reported in Fig. 4 for the QEKF. In the figure, the position and velocity state errors for each run are evaluated at each time step, as

$$\epsilon_i = \hat{x}_i - x_{true} \quad (79)$$

Each gray line is a separate simulation. The *estimated* covariance, in terms of 3σ boundary, is calculated directly from the diagonal entries of the updated covariance matrix of the filter, and reported in red. These lines indicate the filter’s own estimation of its error spread and indicate how confident it is in its estimate. On the contrary, the *effective* covariance level, again in terms of the 3σ boundary, represents how the filter accuracy actually performs, as these levels of standard deviation are calculated directly from the MC runs. A consistent filter has an overlapping between the estimated and the effective error covariance levels, meaning that it is able to correctly predict its own uncertainties in the estimates that it provides. Thus, the QEKF is a consistent estimator that achieves the correct tracking of the target motion. The QUKF behaves similarly to the QEKF, and, thus, the MC analysis has not been reported.

Now that the consistency of the QEKF and QUKF has been assessed, it is of interest to compare their accuracy to their linear counterparts. Figure 5 shows the advantages in accuracy

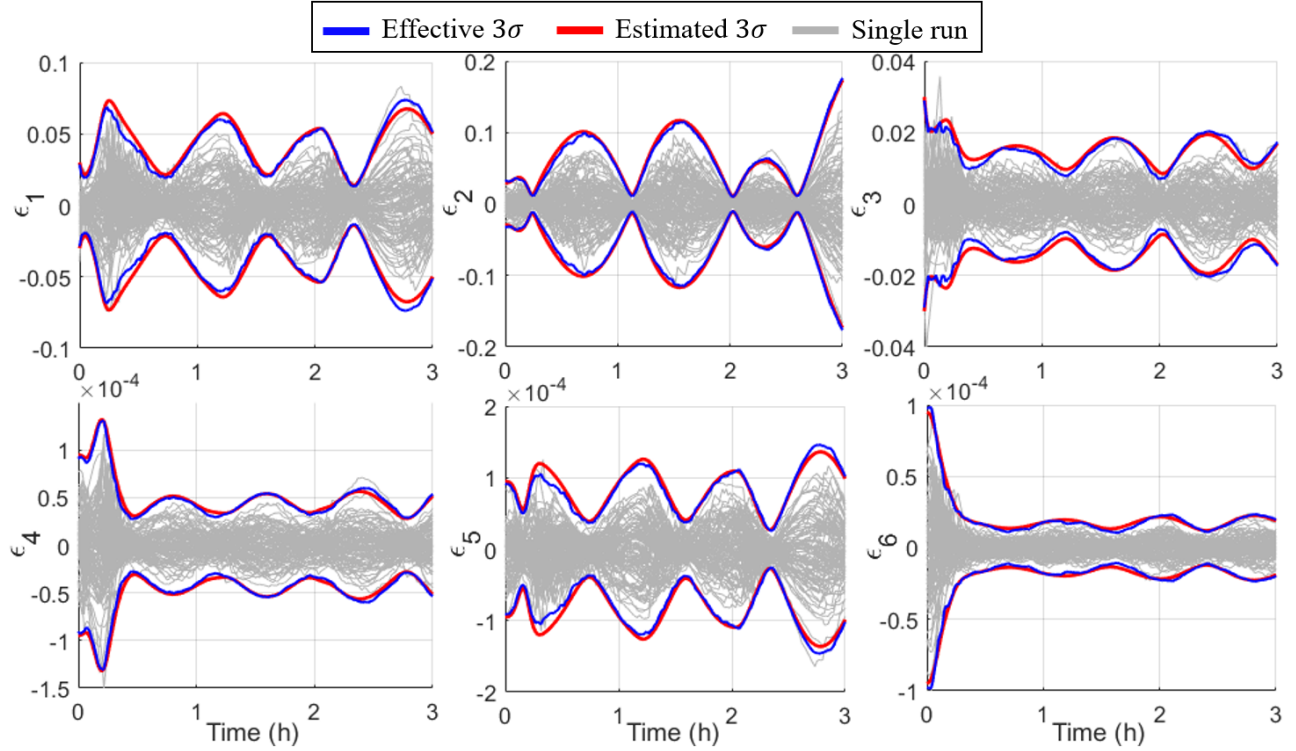


Fig. 4. Monte Carlo Consistency Analysis for the Convergence of the QEKF.

obtained by a QMMSE filter over an LMMSE one. The figure reports the error standard deviation levels both for position and velocity. Estimated covariances are evaluated directly from the updated step of the filter as

$$\sigma_{pos,EST} = \sqrt{P_{xx} + P_{yy} + P_{zz}} \quad (80)$$

$$\sigma_{vel,EST} = \sqrt{P_{v_x v_x} + P_{v_y v_y} + P_{v_z v_z}} \quad (81)$$

while the effective error standard deviation, coming from the Monte Carlo analysis with multiple runs, is evaluated as

$$\sigma_{pos,EFF} = \sqrt{\sum_{j=\{x,y,z\}} \left(\sum_{i=1}^{N_{MC}} (\epsilon_{j,i} - \hat{\epsilon}_j)^2 \right)} \quad (82)$$

$$\sigma_{vel,EFF} = \sqrt{\sum_{j=\{v_x,v_y,v_z\}} \left(\sum_{i=1}^{N_{MC}} (\epsilon_{j,i} - \hat{\epsilon}_j)^2 \right)} \quad (83)$$

for each time step of the simulation. Once again, a consistent filter is assessed by the overlapping of the two standard deviations, reported as continuous and dashed lines in the figure, respectfully. At first sight, it is noted that the EKF (in green) and the UKF (in orange) have a similar accuracy level and performance. The QEKF (in blue) and the QUKF (in black) show evident benefits in accuracy, with error levels well below the other filters. The two quadratic update estimator better accounts for the non-Gaussianity of the noise distribution thanks to the introduction of the high-order central moments. Between the two QMMSE estimators, the QUKF is more accurate than the QEKF due to the use of the unscented transformation over mere linearization, as expected, and shown

in the transient section of the simulation in the first hour of measurement acquisition.

The filters proposed in this paper provide a natural improvement of the classic estimators in the sense that a parabolic fitting always provides a more accurate representation than linear regression. We were able to obtain a quadratic approximation of the MMSE using well-known uncertainty propagation techniques, such that the resulting QEKF and QUKF are easily accessible to the general reader as new filter benchmarks. Moreover, the quadratic formulation of the update, which is the core of the QMMSE, can be expanded to any uncertainty transformation technique, as the Kronecker formulations provided work regardless of how \mathbf{P}_{yy} has been evaluated.

VIII. CONCLUSION

The paper reported two new filtering techniques based on the QMMSE, promoting nonlinear estimators. While the QEKF and QUKF can be seen as the natural improvement of EKF and UKF, the formulation and derivations reported in the paper are expandable to any type of uncertainty quantification technique. That is, the quadratic update approximates the true MMSE more accurately than the linear update, especially when influenced by non-Gaussian noise that requires information of high-order central moments, such as skewness and kurtosis. The QUKF and QEKF obtain more information in the evaluation of the augmented Kalman gain to provide a more accurate estimate, as they know more information regarding the true shape of the posterior distribution.

Conceptually, the technique can be expanded to include third and higher-order polynomial approximations of the up-

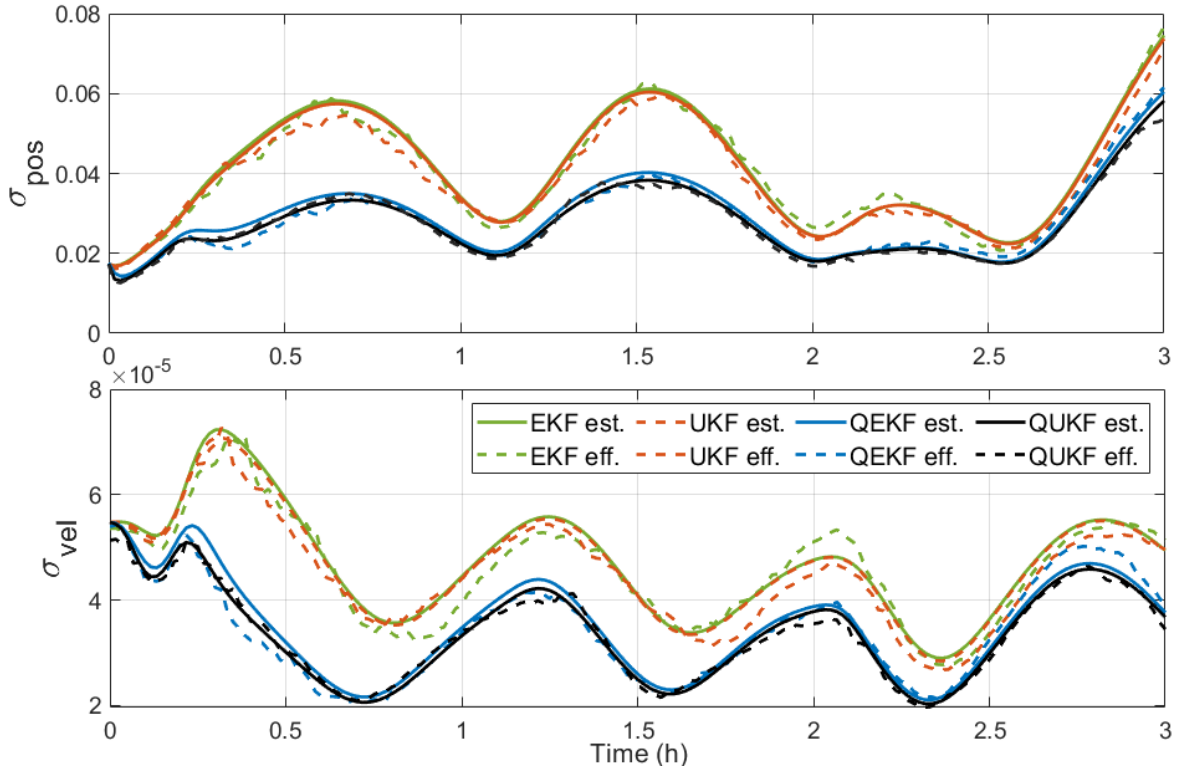


Fig. 5. Error Standard Deviation Comparison between the linear estimators (EKF, UKF) and the Quadratic Estimators (QEKF, QUKF)

date, but they require the precise and attentive evaluation of the influence of the noise on the evaluation of expected values whenever dealing with the residual. In fact, as the order of the update increases, the knowledge of higher-order central moments of distributions must be provided or approximated to guarantee the correct functionality of the filter.

Regarding the quadratic approximations provided here, the QEKF and QUKF propose significant improvements and benefits using well-known techniques and minor changes to the original EKF and UKF algorithms, so that the benefits in accuracy come with limited added computational cost.

REFERENCES

- [1] R. E. Kalman, "A new approach to linear filtering and prediction problems," 1960.
- [2] A. Gelb *et al.*, *Applied optimal estimation*. MIT press, 1974.
- [3] J. L. Junkins and P. Singla, "How nonlinear is it? a tutorial on nonlinearity of orbit and attitude dynamics," *The Journal of the Astronautical Sciences*, vol. 52, pp. 7–60, 2004.
- [4] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.
- [5] I. Arasaratnam and S. Haykin, "Cubature kalman filters," *IEEE Transactions on automatic control*, vol. 54, no. 6, pp. 1254–1269, 2009.
- [6] S. Servadio and R. Zanetti, "Differential algebra-based multiple gaussian particle filter for orbit determination," *Journal of Optimization Theory and Applications*, vol. 191, no. 2, pp. 459–485, 2021.
- [7] M. Valli, R. Armellini, P. Di Lizia, and M. Lavagna, "A gaussian particle filter based on differential algebra for spacecraft navigation," 2012.
- [8] T. S. Schei, "A finite-difference method for linearization in nonlinear estimation algorithms," *Automatica*, vol. 33, no. 11, pp. 2053–2058, 1997.
- [9] M. Majji, J. L. Junkins, and J. D. Turner, "A high order method for estimation of dynamic systems," *The Journal of the Astronautical Sciences*, vol. 56, no. 3, pp. 401–440, 2008.
- [10] M. Valli, R. Armellini, P. Di Lizia, and M. R. Lavagna, "Nonlinear mapping of uncertainties in celestial mechanics," *Journal of Guidance, Control, and Dynamics*, vol. 36, no. 1, pp. 48–63, 2013.
- [11] H. W. Sorenson and D. L. Alspach, "Recursive bayesian estimation using gaussian sums," *Automatica*, vol. 7, no. 4, pp. 465–479, 1971.
- [12] D. Alspach and H. Sorenson, "Nonlinear bayesian estimation using gaussian sum approximations," *IEEE transactions on automatic control*, vol. 17, no. 4, pp. 439–448, 2003.
- [13] N. J. Gordon, D. J. Salmond, and A. F. Smith, "Novel approach to nonlinear/non-gaussian bayesian state estimation," in *IEE proceedings F (radar and signal processing)*, vol. 140, pp. 107–113, IET, 1993.
- [14] F. Hutter and R. Dearden, "The gaussian particle filter for diagnosis of non-linear systems," *IFAC Proceedings Volumes*, vol. 36, no. 5, pp. 909–914, 2003.
- [15] S. Servadio, "Likelihood scouting via map inversion for a posterior-sampled particle filter," *IEEE Transactions on Aerospace and Electronic Systems*, 2024.
- [16] A. De Santis, A. Germani, and M. Raimondi, "Optimal quadratic filtering of linear discrete-time non-gaussian systems," *IEEE Transactions on Automatic Control*, vol. 40, no. 7, pp. 1274–1278, 1995.
- [17] F. Carravetta, A. Germani, and N. Raimondi, "Polynomial filtering of discrete-time stochastic linear systems with multiplicative state noise," *IEEE Transactions on Automatic Control*, vol. 42, no. 8, pp. 1106–1126, 1997.
- [18] A. Germani, C. Manes, and P. Palumbo, "Polynomial extended kalman filter," *IEEE Transactions on Automatic Control*, vol. 50, no. 12, pp. 2059–2064, 2005.
- [19] S. Servadio and R. Zanetti, "Recursive polynomial minimum mean-square error estimation with applications to orbit determination," *Journal of Guidance, Control, and Dynamics*, vol. 43, no. 5, pp. 939–954, 2020.
- [20] S. Servadio, R. Zanetti, and B. A. Jones, "Nonlinear filtering with a polynomial series of gaussian random variables," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 1, pp. 647–658, 2020.
- [21] L. Isserlis, "On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables," *Biometrika*, vol. 12, no. 1/2, pp. 134–139, 1918.
- [22] S. Servadio and R. Zanetti, "Estimation of the conditional state and covariance with taylor polynomials," *Journal of Advances in Information Fusion*, vol. 16, no. 2, p. 126, 2021.