2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

Enhanced Quadratic Extended Kalman Filter

Giovanni Palombo, Massimiliano d'Angelo*, Federico Papa, Valerio Cusimano, Alessandro Borri

***Abstract*— In this note, we present a new solution to the filtering problem for stochastic discrete-time nonlinear systems, which we refer to as the Enhanced Quadratic Extended Kalman Filter (eQEKF). Starting from the concept underlying the existing formulation of the Quadratic Extended Kalman Filter (QEKF), based on the definition of an augmented output through Kronecker powers, we propose a different method that enables us to overcome certain inevitable standard approximation issues, reducing the computational workload. Also, we show the effectiveness of the proposed approach with respect to the QEKF and with respect to the classical Extended Kalman Filter, as highlighted by two numerical examples, in the case of Gaussian and non-Gaussian noises.**

I. INTRODUCTION

The relevance of optimal estimation in engineering applications is notorious nowadays [1]. The problem is usually modeled as a nonlinear programming problem with the goal of finding the minimum of a performance criterion with dynamical constraints, *i.e.* an underlying dynamical system. A typical case is when the system's state to be estimated/controlled is linear and some performance criterion, or cost function, is a quadratic form with respect to the state and control variables. In the stochastic case, *i.e.* when the state and/or the output measurements are affected by stochastic noise, the celebrated Linear Quadratic Gaussian (LQG) Regulator solves the optimal control problem when the noise sequences are jointly Gaussian [2]. A well-known property of the LQG Regulator problem with partial state information is that the optimal regulator, synthesized by the LQ optimal technique, is generated from the optimal linear estimate of the state, namely the Kalman Filter (KF). In fact, for linear Gaussian systems, the KF is the optimal recursive estimator in the minimum mean-square error sense. On the other hand, for linear non-Gaussian systems, the KF is the best affine estimator but it is yet possible to develop estimators that are more accurate. Furthermore, in the last decades increasing attention has been paid to non-Gaussian systems in control engineering [3]–[8]. On the other hand, when the system linearity hypothesis id dropped, the estimation task becomes even more challenging. Indeed, in the domain of nonlinear systems, numerous estimation methods have been embraced to tackle the challenges posed by the nonlinearity of dynamics and the inevitable presence of non-Gaussian noises. The well-known Extended Kalman Filter (EKF) is a direct extension of the KF, essentially involving the application of a KF to the corresponding linearized nonlinear systems. Despite its widespread use across various applications and successful outcomes, there are no guarantees regarding the numerical stability of the algorithm or bounds on the estimation error. Consequently, the field remains highly active. In this regards, a significant contribution has been given in [9] where the polynomial version of the well-known EKF is presented making use of the Carleman approximation for nonlinear systems. An effective method to cope with both non-Gaussian systems and nonlinearities of the system dynamics is through the use of polynomial methods involving the Kronecker powers of the systems. In [10], a quadratic filter, namely an estimator of the state which use the second power of the measurement, is applied to discrete-time linear system with additive non-Gaussian noise. In [11], this method has been extended to a generic polynomial estimate of order μ. Other interesting extensions are the works [12], where the polynomial filtering technique has been extended to system with multiplicative state noise, whilst applications to the case with unknown forth-order moments of the noises and to descriptor systems have been studied in [13] and [14], respectively. Also, systems with switching measurements and quantization effects have been studied in [15] and [16], respectively, whilst applications to target tracking problem and economic models have been given in [17], [18] and [19], respectively.

More recently, an important issue has been resolved in the papers [20], [21] where the authors proposed a quadratic filter for linear systems with non-Gaussian noise that improves the performance over the Kalman filter also in the case of non-asymptotically stable systems. Furthermore, the prediction provided by quadratic or polynomial predictors has been exploited in the optimal control problem in [22] and [23]. Extensions to the time-varying case have been provided in [24], systems with nonlinear measurements have been studied in [23], whilst packet dropping networks have been considered in [25] and [26].

In this paper, we concentrate on the framework of nonlinear discrete-time stochastic systems. Drawing inspiration from the concept of the Quadratic Extended Kalman Filter as introduced in [9], we present a novel filtering algorithm that merges the QEKF and the EKF in a unique manner. This integration yields an algorithm that shows to be more efficient both in terms of computation, performance and stability of both the QEKF and the EKF. The paper is organized as follows: Section II introduces the problem, while Section III covers some preliminaries on the Quadratic Extended Kalman Filter proposed by [9]. Section IV presents the rationale behind the new formulation and introduces

*Corresponding author.
All the authors are with the Istituto di Analisi dei Sistemi e Informatica, Italian National Research Council (IASI-CNR), Rome, 00185, Italy, {giovanni.palombo,massimiliano.dangelo, federico.papa,valerio.cusimano, alessandro.borri}@iasi.cnr.it.

the new filter named Enhanced Quadratic Extended Kalman Filter (eQEKF). Two illustrative examples are provided to highlight the performance of the proposed filter.

Notation If $A \in \mathbb{R}^{n\times n}$ then $A^\top$ denotes its transpose and $|A|$ denotes its determinant. If $v_1,\dots,v_n$ are column vectors in $\mathbb{R}^n$, then $v = \mathrm{col}(v_1,\dots,v_n)$ denotes the vector $v = [v_1^\top,\dots,v_n^\top]^\top$. Moreover, if $v \in \mathbb{R}^n$, then we denote with $\mathrm{diag}(v) \in \mathbb{R}^{n\times n}$ the diagonal matrix with entries the components of v. The symbol † denotes the pseudo-inverse of a matrix. If A is a squared matrix, then $\mathrm{st}(A)$ is the stack (or vectorization) operation and $\mathrm{st}^{-1}(\cdot)$ its inverse operation. If A and B are two matrices in $\mathbb{R}^{n\times n}$, then the Kronecker product is $A \otimes B$ and $A^{[i]}$ is the i-th Kronecker power of A. M is the commutation matrix for vectors in $\mathbb{R}^q$, namely if $v, w \in \mathbb{R}^q$ then $v \otimes w = M(w \otimes v)$. We indicate with I and $\underline{0}$ the identity matrix and the zero matrix, respectively, of appropriate dimension. The euclidean norm in $\mathbb{R}^{n\times n}$ is denoted with $\|\cdot\|$. If $f : \mathbb{R}^n \to \mathbb{R}^n$ then $\nabla f(x)|_{x=\bar{x}}$ is the Jacobian of f evaluated in $\bar{x}$, namely $(\nabla f)_{ij}(x) = \frac{\partial f_i}{\partial x_j}(x)$. Finally, given a random variable (r.v.) X in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, implicit in the rest of the paper, we denote with $\mathbb{E}[X]$ its expectation. We denote with $X \sim \mathcal{N}(\mu, \sigma^2)$ a Gaussian r.v. X with mean μ and variance σ^2. Given the stochastic noise sequence $\{w_k\}$, the second, third and fourth moment of $\{w_k\}$ are denoted as $\phi_w^{(j)} = \mathbb{E}[w(k)^{[j]}]$, with $j = 2, 3, 4$, while $\tilde{\phi}_w^{(3)} = \mathbb{E}[w_k w_k^{[2]}]$.

II. Problem Statement

This paper investigates the filtering problem for nonlinear affine discrete time systems in the following form:

$$x(k+1) = f(x(k)) + v(k) \tag{1}$$

$$y(k) = h(x(k)) + w(k), \tag{2}$$

with initial condition $x(0) = x_0 \in \mathbb{R}^n$, where $x(k) \in \mathbb{R}^n$ is the state vector and $y(k) \in \mathbb{R}^q$ is the measurement vector, $k \geq 0$ is the discrete time variable, f and h are vector fields, $f : \mathbb{R}^n \to \mathbb{R}^n$ and $h : \mathbb{R}^n \to \mathbb{R}^q$, differentiable up to some order μ. The terms $v(k) \in \mathbb{R}^n$ and $w(k) \in \mathbb{R}^q$ represent the zero-mean independent random noise sequences affecting the state and output equations, respectively. We note that assuming Gaussian distributions for the stochastic sequences $v(k)$ and $w(k)$ is not necessary. However, statistical information (and boundedness) up to the fourth order is required. For instance, in the case of a Gaussian distribution, knowledge of the first two moments (mean and covariance) is necessary. Additionally, we remark that the complexity of the proposed solution remains unaffected by the addition of a control input to the system and/or time dependencies of the vector fields f and h.

The primary motivation behind this paper draws inspiration from the research presented in [9], where the authors introduce the Polynomial Extended Kalman Filter (PEKF) for nonlinear stochastic systems, namely a filter that estimates the state of the system through a polynomial transformation of the state and measurement equations of the system. This result is achieved by using Kronecker algebra which allows one to perform operations on matrix and vectors in a easier and more fashionable way. In particular the Carleman approximation is used to approximate the nonlinear system as a truncated Taylor series. Then a KF is used to solve the filtering problem for the polynomial-linearized system.

Our filter adopts concepts of both the EKF and PEKF. We named this filter Enanched Quadratic Kalman Filter (eQEKF), since we consider the polynomial version of order two. We will show that eQEKF offers enhanced efficiency with respect to its predecessor, benefiting both computational efficiency and error performances.

III. Preliminaries

This section is mostly devoted to the description of the Polynomial Extended Kalman Filter (PEKF) introduced by [9].

Before delving into the approach, let us first provide a brief summary of the Extended Kalman Filter (EKF). This will help elucidate the motivations behind our work. It is worth noting that the EKF utilizes (1)–(2) as the prediction equations for the state and output, respectively:

$$\hat{x}(k+1|k) = f(\hat{x}(k)) \tag{3}$$

$$\hat{y}(k+1|k) = h(\hat{x}(k+1|k)). \tag{4}$$

The estimation in the EKF follows similarly as in the classical Kalman filter, where the gain is computed using the linearized equations of the state and output, specifically, $A_k = \nabla f(x)|_{x=\hat{x}(k)}$ and $C_k = \nabla h(x)|_{x=\hat{x}(k|k-1)}$.

The EKF provides an approximation of the best linear estimation of the state $x(k)$ given the output sequence $y(k)$ up to time k.

A. Polynomial Extended Kalman Filter (PEKF)

The fundamental concept behind the PEKF is to formulate an estimate $\hat{x}$ of the state x as a polynomial function of the output measurements y. This differs from the linear estimation employed by the Extended Kalman Filter (EKF). In essence, the PEKF leverages a polynomial model to capture the relationship between the state and the output.

To have a better grasp of theoretical foundation of the approach, let $(\Omega, \mathcal{F}, P)$ be a probability space, and let $\mathcal{G} \subset \mathcal{F}$ and $\mathcal{L}^2(\mathcal{G}, n)$ be the Hilbert space of the n-dimensional $\mathcal{G}-$measurable random variables with finite second moment. We write $\mathcal{L}^2(X, n)$ for the Hilbert space of the n-dimensional random variables with finite second moment, measurable with respect to the sigma-algebra generated by X. Moreover let $\Pi(\cdot|M)$ be the orthogonal projection onto a given Hilbert space M. Given system (1)–(2) and the vector of the output sequence as $Y_k = \mathrm{col}(y(0), ..., y(k))$, defining an auxiliary vector $Y'_k = \mathrm{col}(1, Y_k)$, the minimum variance estimate of $x(k)$ is the orthogonal projection of $x(k)$ onto the space $\mathcal{L}^2(Y'_k, n)$:

$$\hat{x}(k) = \mathbb{E}[x(k) \,|\, Y_k] = \Pi(x(k) \,|\, \mathcal{L}^2(Y'_k, n)) \tag{5}$$

It is well known that, in the jointly Gaussian linear case, this projection is equivalent to the projection on $\mathcal{L}_y^k$, the subspace of the linear transformations of Y'_k. Clearly $\mathcal{L}_y^k \subset$

$\mathcal{L}^2(Y'_k, n)$. The Kalman filter is an algorithm that recursively projects onto $\mathcal{L}_y^k$, which proves to be optimal in the minimum variance sense under the linear Gaussian assumption. However, when one or both assumptions are not met, the algorithm is limited to providing the best linear estimation of $x(k)$. Indeed better sub-optimal estimations can be sought projecting on larger sub-spaces than $\mathcal{L}_y^k$. For example, one may consider the space of quadratic transformations of Y_k, denoted by $\mathcal{Q}_y^k$. Indeed, since $\mathcal{L}_y^k \subset \mathcal{Q}_y^k \subset \mathcal{L}^2(Y'_k, n)$, projecting the state onto $\mathcal{Q}_y^k$ will return an error variance of the estimation, smaller or equal than that of the one made projecting on $\mathcal{L}_y^k$. The generalization to a power of generic polynomial degree being straightforward. The essence of the Polynomial Kalman Filter lies in crafting an augmented system to represent (1) and (2) and applying a Kalman filter to it. To achieve this, let μ be a positive integer, and consider the Kronecker powers of the state and the output

$$x^{[m]}(k+1) = \left(f(x(k)) + v(k)\right)^{[m]}, \tag{6}$$

$$y^{[m]}(k) = \left(h(x(k)) + w(k)\right)^{[m]}, \tag{7}$$

with $m = 1, \ldots, \mu$. We can thus introduce the extended state $X(k)$, the collection of the Kronecker powers of the state up to order μ, and the extended output $Y(k)$, the collection of the Kronecker powers of the output up to order μ:

$$X(k+1) = \begin{bmatrix} f(x(k)) + v(k) \\ (f(x(k)) + v(k))^{[2]} \\ \vdots \\ (f(x(k)) + v(k))^{[\mu]} \end{bmatrix}, \tag{8}$$

$$Y(k) = \begin{bmatrix} h(x(k)) + w(k) \\ (h(x(k)) + w(k))^{[2]} \\ \vdots \\ (h(x(k)) + w(k))^{[\mu]} \end{bmatrix}. \tag{9}$$

Using the Carleman approximation, a Taylor series expansion employing the Kronecker algebra, the extended vector can be represented as an infinite series of Kronecker polynomials [27], [28]. By the truncation of those polynomial series at degree μ, we can obtain the desired linear augmented system:

$$X^{\mu}(k+1) = \mathcal{A}^{x^\star} X^{\mu}(k) + \mathcal{U}^{x^\star} + V^{x^\star}(k), \tag{10}$$

$$Y^{\mu}(k) = \mathcal{C}^{x^\star} X^{\mu}(k) + \Gamma^{x^\star} + W^{x^\star}(k), \tag{11}$$

where $X^{\mu}(k) = \operatorname{col}(X_1(k), \ldots, X_{\mu}(k))$ and $Y^{\mu}(k) = \operatorname{col}(Y_1(k), \ldots, Y_{\mu}(k))$, whose components $X_m(k)$ and $Y_m(k)$ are the Carleman approximation of $x^{[m]}(k)$ and $y^{[m]}(k)$, respectively, $m = 1, ..., \mu$, up to the $\mu-$th degree, which are computed as

$$X_m(k+1) = \sum_{i=1}^{\mu} A_{m,i}^{x^\star} X_i(k) + u_i^{x^\star} + v_i^{x^\star}(k), \tag{12}$$

$$Y_m(k) = \sum_{i=1}^{\mu} C_{m,i}^{x^\star} X_i(k) + \gamma_i^{x^\star} + w_i^{x^\star}(k), \tag{13}$$

with $X_m(0) = x^{[m]}(0) \in \mathbb{R}^{n^m}$, while the matrices $\mathcal{A}^{x^\star}$, $A_{m,i}^{x^\star}$, $\mathcal{C}^{x^\star}$, $C_{m,i}^{x^\star}$ and the vectors $\mathcal{U}^{x^\star}$, $u_i^{x^\star}$, $V^{x^\star}(k)$, $v_i^{x^\star}(k)$, $\Gamma^{x^\star}$, $\gamma_i^{x^\star}$, $W^{x^\star}(k)$, $w_i^{x^\star}(k)$ are defined in [9] (see Section II and the Appendix). Notice that, all the matrices are evaluated at $x^\star$, which is the evaluation point of the Carleman approximation. We point out that the sequences $\mathcal{U}^{x^\star}$ and $\Gamma^{x^\star}$ are deterministic sequences, while $V^{x^\star}(k)$ and $W^{x^\star}(k)$ (bilinear functions of the extended state $X^{\mu}(k)$) are stochastic zero-mean white sequences, uncorrelated with $X^{\mu}(k)$ (a direct consequence of the fact that the original noise sequences $\{v(k)\}$ and $\{w(k)\}$ are independent, white and uncorrelated with the initial condition $x(0)$).

It is worth mentioning that (10)–(11) are approximations of the original extended system (8)–(9). Employing a standard Kalman filter algorithm for time-varying linear systems (10)-(11) yields a linear estimate of the state $X^{\mu}(k)$ from $Y(k)$, namely $\hat{X}(k)$. Since $Y(k)$ contains the Kronecker powers of the original output $y(k)$ up to the order μ, by definition, this linear estimation from $Y(k)$ approximates the best μ-th polynomial estimation from $y(k)$. Or, in other words, since the first n components of $X^{\mu}(k)$ approximate $x(k)$, the polynomial estimate of the original state $x(k)$ is simply obtained by taking the first n components of the augmented state estimate $\hat{X}(k)$.

The application of the Kalman filter to (10)–(11) achieving the μ-th polynomial estimation for (1)–(2), has the following equations

$$\hat{X}(k+1|k) = \mathcal{A}(k)\hat{X}(k) + \mathcal{U}(k) \tag{14}$$

$$P_p(k+1) = \mathcal{A}(k)P(k)\mathcal{A}^{\top}(k) + \Psi^{V}(k) \tag{15}$$

$$K(k+1) = P_p(k+1)\mathcal{C}^{\top}(k+1)\cdot \left(\mathcal{C}(k+1)P_p(k+1)\mathcal{C}^{\top}(k+1) + \Psi^{W}(k+1)\right)^{\dagger}, \tag{16}$$

$$\hat{Y}(k+1|k) = \mathcal{C}(k+1)\hat{X}(k+1|k) + \Gamma(k+1), \tag{17}$$

$$\hat{X}(k+1) = \hat{X}(k+1|k) + K(k+1)(Y(k+1) - \hat{Y}(k+1|k)), \tag{18}$$

$$P(k+1) = (I - K(k+1)\mathcal{C}(k+1))P_p(k+1), \tag{19}$$

$$\hat{x}(k+1) = [I_n, \quad \underline{0}]\hat{X}(k+1), \tag{20}$$

with $\hat{X}(0|-1) = \mathbb{E}[X(0)]$ and $P_p(0) = \operatorname{Cov}(X(0))$, where $\mathcal{A}(k) = \mathcal{A}^{\hat{x}(k)}$, $\mathcal{C}(k) = \mathcal{C}^{\hat{x}(k|k-1)}$, $\mathcal{U}(k) = \mathcal{U}^{\hat{x}(k)}$, $\Gamma(k) = \Gamma^{\hat{x}(k|k-1)}$, and $\Psi^{V}(k)$ and $\Psi^{W}(k)$ are the covariance matrices of the extended noise $V(k) = V^{\hat{x}(k)}(k)$ and $W(k) = W^{\hat{x}(k|k-1)}(k)$. Clearly, the Quadratic Extended Kalman Filter (QEKF) is obtained when the degree is $\mu = 2$. From now on, for the purpose of comparison with the current paper where we develop a quadratic version of the polynomial filter, we will refer to the Polynomial Extended Kalman Filter of [9] as the Quadratic Extended Kalman Filter (QEKF), namely the polynomial version of order 2.

IV. Enhanced Quadratic Extended Kalman Filter (eQEKF)

Our work originated from the insight that the estimation step of the QEKF, represented by (14), could be approached differently, aligning more closely with the EKF methodology, while preserving the essence of the polynomial-like projection introduced by the QEKF. To better understand this concept, let us further explore the approximations employed by the QEKF approach in order to better underscore the distinctions between the two approaches and provide a more comprehensive explanation of our work.

In the nonlinear and/or non-Gaussian case, it is generally not possible to evaluate the conditional expectation of the state at time $k+1$ with respect to the output sequence up to time k, namely $\mathbb{E}[x(k+1)\,|\,Y_k]$. To overcome this difficulty, in the PEKF, the Carleman approximation is employed: the above expectation can be approximated by

$$\begin{aligned}\hat{x}(k+1|k) &= \mathbb{E}[x(k+1)\,|\,Y_k] \\ &= \mathbb{E}[f(x(k))\,|\,Y_k] + \mathbb{E}[v(k)\,|\,Y_k] \\ &\approx \mathbb{E}\left[\sum_{i=0}^{\mu} \frac{\nabla^{[i]} \otimes f(\bar{x})}{i!}(x(k)-\bar{x})^{[i]}\,|\,Y_k\right] \\ &= \mathbb{E}\left[\sum_{i=0}^{\mu} A_i(\bar{x})x^{[i]}(k) + B(\bar{x})\,|\,Y_k\right]. \qquad (21)\end{aligned}$$

where the independence hypothesis between state noise $v(k)$ and the output sequence Y_k have been utilized. Equation (21) is true for any $\bar{x} \in \mathbb{R}^n$. Therefore, we can find $\bar{\mathcal{A}}^{\bar{x}}$ and $\bar{\mathcal{U}}^{\bar{x}}$ and write

$$\begin{aligned}\hat{x}(k+1|k) &\approx \sum_{i=0}^{\mu} A_i(\bar{x})\mathbb{E}\left[x^{[i]}(k)\,|\,Y_k\right] + B(\bar{x}) \\ &\approx \bar{\mathcal{A}}^{\bar{x}}\hat{X}(k) + \bar{\mathcal{U}}^{\bar{x}}, \qquad (22)\end{aligned}$$

where the last approximation sign in (22) is used because $\hat{X}(k)$ resulting from the QEKF does not correspond to $\mathbb{E}\left[\mathrm{col}(x^{[i]}(k))\,|\,Y_k\right]$, as the original system is also approximated in the output equation (meaning that while the filter uses the measurement data $Y(k)$ that comes from (9), the prediction of the approximated output $Y^{\mu}(k)$ use the linearized dynamics (11), as show in (17)). We note that (22) represents the first n components of equation (14). At this point it appears clear why the estimations of the extended state are needed, since the conditional expectation of all the powers of $x^{[i]}(k)$ are considered in the summation. By considering not only $x(k)$ but the extended state $X(k)$, the procedure can be repeated to obtain the filtering algorithm in closed form presented in the previous section.

Up to this point, two approximations are evident: the Carleman approximation up to order μ of the nonlinear function f (and of h as well), and the approximation of the conditional expected value in the second line of (22). However, there is a further approximation. When the filter is applied, the state equation matrices are evaluated in the state estimation at the preceding time (while the output equation matrices are evaluated in the state prediction at the actual time), as in (14). Let us retrieve it from (21) with this new substitution

$$\hat{x}(k+1|k) \approx \mathbb{E}\left[\sum_{i=0}^{\mu} A_i(\hat{x}(k))x^{[i]}(k) + B(\hat{x}(k))\,|\,Y_k\right]. \qquad (23)$$

The point we want to highlight here is that, unfortunately, $\hat{x}(k)$ and $x(k)$ are not independent. Indeed $\mathbb{E}\left[A_i(\hat{x}(k))x^{[i]}(k)\,|\,Y_k\right] \neq A_i(\hat{x}(k))\mathbb{E}\left[x^{[i]}(k)\,|\,Y_k\right]$ and thus a third approximation arises: $\mathbb{E}\left[A_i(\hat{x}(k))x^{[i]}(k)\,|\,Y_k\right] \approx A_i(\hat{x}(k))\mathbb{E}\left[x^{[i]}(k)\,|\,Y_k\right]$. Once this additional step is completed, the derivation becomes quite fluid and straightforward, enabling the prediction and estimation of x and all its powers up to order μ.

The eQEKF revolves around the idea that the filter prediction step can be performed in line with the style of the EKF, while retaining the idea of considering the powers of the output for the estimation step, proper of the QEKF. The prediction step of the eQEKF is thus achieved in the same way as the EKF as follows

$$\begin{aligned}\hat{x}(k+1|k) &= \mathbb{E}[x(k+1)\,|\,Y_k] \qquad (24)\\ &= \mathbb{E}[f(x(k))\,|\,Y_k] + \mathbb{E}[v(k)\,|\,Y_k] \qquad (25)\\ &\approx f(\mathbb{E}[x(k)\,|\,Y_k]) \qquad (26)\\ &\approx f(\hat{x}(k)) \qquad (27)\end{aligned}$$

In (26), we employ an EKF-like approximation, specifically $\mathbb{E}[f(x(k))\,|\,Y_k] \approx f(\mathbb{E}[x(k)\,|\,Y_k])$. Moving to (27), the approximation mirrors the one introduced in (22). This methodology can be readily extended to encompass any power of a nonlinear function. More precisely, in the sequel, for $h(x(k))^{[i]}$, we adopt $\mathbb{E}[h(x(k))^{[i]}\,|\,Y_k] \approx h(\mathbb{E}[x(k)\,|\,Y_k])^{[i]}$ for any i. It is essential to highlight that the prediction equation in (24) deviates from the EKF formulation, indeed it is performed on the polynomial extended output sequence, instead of the simple output sequence. The comprehensive exposition of the filter, coupled with the strategic choices made to tailor a fully functional algorithm, will be elucidated in the subsequent sections.

A. *eQEKF algorithm*

In this section, we delve into the details of the eQEKF algorithm. The idea of the eQEKF is to use the original n dimensional state equation jointly with a quadratic extended output equation, to define the equivalent model on which applying an EKF. This will allow us to achieve a polynomial-like estimation of the state, reducing the memory requirement with respect to the QEKF and allowing us to achieving better performances.

Since the state estimation part of the algorithm is quite similar to the one of the EKF, we will first address the output equation definition and characterization. It is worth noticing that we do not use the Carleman approximation to define the extended output equation.

Indeed, expanding the output equation given in (9), using (7), we obtain

$$\begin{aligned}Y(k) &= \begin{bmatrix} h(x(k)) + w(k) \\ h(x(k))^{[2]} + \phi_w^{(2)} + w^{(2)}(k) \end{bmatrix} \\ &= H(x(k)) + \Theta(k) + W(k), \qquad (28)\end{aligned}$$

where,

$$H(x(k)) = \begin{bmatrix} h(x(k)) \\ h(x(k))^{[2]} \end{bmatrix}, \quad \Theta(k) = \begin{bmatrix} 0 \\ \phi_w^{(2)} \end{bmatrix}, \tag{29}$$
$$W(k) = \begin{bmatrix} w(k) \\ w^{(2)}(k) \end{bmatrix},$$

with $w^{(2)}(k) = (I+M)(h(x(k)) \otimes w(k)) + w(k)^{[2]} - \phi_w^{(2)}$ zero-mean, white noise. We note that the term $\Theta(k)$ is a known deterministic vector, while the term $W(k)$ is the noise term of the extended output vector, white and uncorrelated with the initial condition $x(0)$.

The equivalent system on which apply the EKF has the following form:

$$x(k+1) = f(x(k)) + v(k), \tag{30}$$
$$Y(k) = H(x(k)) + \Theta(k) + W(k). \tag{31}$$

and the filtering algorithm we propose takes the following form

$$\hat{x}(k+1|k) = f(\hat{x}(k)) \tag{32}$$
$$P_p(k+1) = A(k)P(k)A^\top(k) + \Psi_v \tag{33}$$
$$K(k+1) = P_p(k+1)C^\top(k+1) \cdot (C(k+1)P_p(k+1)C^\top(k+1) + \Psi_W(k+1))^\dagger, \tag{34}$$
$$\hat{Y}(k+1|k) = H(\hat{x}(k+1|k)) + \Theta(k+1), \tag{35}$$
$$\hat{x}(k+1) = \hat{x}(k+1|k) + K(k+1)(Y(k+1) - \hat{Y}(k+1|k)), \tag{36}$$
$$P(k+1) = (I - K(k+1)C(k+1))P_p(k+1), \tag{37}$$

where $A(k) = \nabla f(x)|_{x=\hat{x}(k)} \in \mathbb{R}^{n\times n}$ and $C(k) = \nabla H(x)|_{x=\hat{x}(k|k-1)} \in \mathbb{R}^{(q+q^2)\times n}$.

Even though, at first glance, the equations may seem quite similar to those of the QEKF, they are not. First, in this part of the algorithm, there is no Carleman approximation. In fact, $A(k)$ and $C(k)$ are evaluated in the standard manner of the EKF. This leads to the consequence that (32) and (36) are n dimensional equations, and eqs. (33) and (37) are $n \times n$ dimensional. This is a significant implication with respect to the QEKF, since in the QEKF the same quantities have dimention $n+n^2$ and $(n+n^2)\times(n+n^2)$, respectively. On the other hand, the eQEKF still mimics the projection of the state onto the space of polynomial transformation of the output, showing promising performances as shown in the next section.

The last point to deal with is the characterization of the noise covariance matrices of the new extended system: (30)– (31). Even if Ψ_v is simply the covariance matrix of $v(k)$, the definition of $\Psi_w(k)$ is more trickier. Indeed, to compute covariance matrix of the extended output noise term $W(k)$, one have to deal with the approximation of the terms $E[H(x(k))]$ and $E[H(x(k))H(x(k))^\top]$. We choose to use the Carleman expansion for the term $H(x(k))$. In particular, given $\nu \in \mathbb{N}$ an $\bar{x} \in \mathbb{R}^n$, one can find $G \in \mathbb{R}^{n\times\bar{\nu}}$ and $\Gamma \in \mathbb{R}^n$, with $\bar{\nu} = n + \cdots + n^\nu$, such that[1]

$$H(x(k)) = \sum_{i=0}^{\nu} \frac{\nabla^{[i]} \otimes H(\bar{x})^{[i]}}{i!}(x(k) - \bar{x})^{[i]}, = GX^\nu(k) + \Gamma, \tag{40}$$

where $X^\nu(k) = \mathrm{col}(x(k), x(k)^{[2]}, \ldots, x(k)^{[\nu]})$. Detailed expressions of G and Γ can be found in [9]. We notice that the order of the approximation ν can be taken arbitrarily large.

Proposition 1: The random sequence $W(k)$ defined in (28) is zero-mean with covariance matrix $\Psi_W(k)$ given by

$$\Psi_W(k) = \begin{bmatrix} \Psi_W^{11} & \Psi_W^{12}(k) \\ \Psi_W^{12\,\top}(k) & \Psi_W^{22}(k) \end{bmatrix}, \tag{41}$$

with

$$\Psi_W^{11} = \mathrm{st}^{-1}(\phi_w^{(2)}), \tag{42}$$
$$\Psi_W^{12}(k) = \Big((GZ_k + \Gamma) \otimes \mathrm{st}^{-1}(\phi_w^{(2)})\Big)(I+M)^\top + \tilde{\phi}_w^{(3)}, \tag{43}$$
$$\Psi_W^{22}(k) = (I+M)\Big((G\Psi_k^{X^\nu}G^\top + GZ_k\Gamma^\top + \Gamma Z_k^\top G^\top \tag{44}$$
$$+ \Gamma\Gamma^\top) \otimes \mathrm{st}^{-1}(\phi_w^{(2)})\Big)(I+M)^\top + \tag{45}$$
$$+ (I+M)\Big((GZ_k + \Gamma) \otimes \tilde{\phi}_w^{(3)}\Big) \tag{46}$$
$$+ \Big((GZ_k + \Gamma) \otimes \tilde{\phi}_w^{(3)}\Big)^\top (I+M)^\top \tag{47}$$
$$+ \mathrm{st}^{-1}(\phi_w^{(4)}) - \phi_w^{(2)}\phi_w^{(2)\top}, \tag{48}$$

where $Z_k = \mathbb{E}[X^\nu(k)]$ and $\Psi_k^{X^\nu} = \mathbb{E}[X^\nu(k)X^\nu(k)^\top]$. The proof is achieved by using standard manipulation and properties of Kronecker algebra.

We note that Z_k and $\Psi_k^{X^\nu}$ can be easily computed through the recursive equations given in eqs. (42) and (43) of [9]. Here lies a delicate aspect of our approach. We still use the Carleman approximation, but only to evaluate the expected value of $h(x(k))$ and $h(x(k))h(x(k))^\top$ which arise in the computation of Ψ_W.

V. Illustrative examples

In this section, the performance of the proposed eQEKF are compared with the EKF and QEKF of [9] in two illustrative examples where we perform $N = 10^4$ Monte Carlo runs with a time horizon for each simulation $T = 500$. Every filter is implemented using the pseudoinverse in the calculation of the Kalman gain, to better achieve a fair comparison between the methods. The first example involves the Chialvo oscillator subjected to Gaussian noise, while the second example considers the Kazufumi model driven by

[1]The operator $\nabla^{[i]}\otimes$ applied to a function $\psi : \mathbb{R}^n \to \mathbb{R}^n$ is defined as

$$\nabla^{[0]} \otimes \psi = \psi \tag{38}$$
$$\nabla^{[i+1]} \otimes \psi = \nabla \otimes \nabla^{[i]} \otimes \psi, \quad i \geq 1 \tag{39}$$

with $\nabla = [\partial/\partial x_1, \ldots, \partial/\partial x_n]$.

Example	MSE_{EKF}	MSE_{QEKF}	MSE_{eQEKF}
Chialvo	55.81 92%	∞ 1%	40.65 97%
Kazufumi	5.36×10^{-2} 100%	5.13×10^{-2} 100%	4.72×10^{-2} 100%

TABLE I

MSE OVER THE MONTE CARLO RUNS OF THE EKF, QEKF AND THE PROPOSED EQEKF TOGETHER WITH THE PERCENTAGES OF STABLE TRAJECTORIES OF ESTIMATION ERROR.

discrete non-Gaussian sequences. The Mean Squared Error (MSE) computed over the N Monte Carlo runs is given by

$$\text{MSE} = \frac{1}{N}\frac{1}{T}\sum_{s=1}^{N}\sum_{k=1}^{T}\left\|x_{(s)}(k) - \hat{x}_{(s)}(k)\right\|^2, \tag{49}$$

where $x_{(s)}(k)$ and $\hat{x}_{(s)}(k)$ are the state and its estimate at time k of the s-th realization respectively[2].

Chialvo oscillator: the model proposed by Chialvo [29] describes excitable biological systems, including neuronal dynamics,

$$\begin{aligned}
x_1(k+1) &= x_1^2(k)e^{(x_2(k)-x_1(k))} + d + v_1(k),\\
x_2(k+1) &= ax_2(k) - bx_1(k)x_2(k) + c + v_2(k),\\
y_1(k) &= x_1x_2^2(k) + w_1(k)\\
y_2(k) &= \text{atan}(x_(k)) + w_2(k)
\end{aligned}$$

With the choice $a = 0.8$, $b = 0.2$, $c = 0.28$, $d = 0.105$ one can ensure periodic solutions. We chose this example because it is a challenging system to estimate, since it has a rich dynamics, which may vary from oscillatory to chaotic behavior, also showing non-trivial responses to small stochastic fluctuations. Even if in the non-Gaussian case the performance were even more promising, we decided to enhance the nonlinearity of the problem with a strongly nonlinear output and consider the Gaussian noise case. For this example, we assume the noise sequences v_1, v_2, w_1 and w_2 to be Gaussian, each with zero mean and variances set to 10^{-1}. The initial condition $x(0)$ is specified as Gaussian with mean $\text{col}(0.3, 1.3)$ and a variance of 1. It is noteworthy that in this example, the QEKF exhibits divergent behavior, rendering it incapable of estimating the system state. In contrast, both the EKF and the proposed eQEKF demonstrate stability under the same conditions. The obtained MSEs together with the percentage of stable trajectories of the estimation error are given in table I. We can see that the improvement in performance of the proposed eQEKF with respect to the EKF is about 27%, while the previous QEKF is not applicable. This substantial improvement underscores the efficacy of the proposed eQEKF algorithm in achieving more accurate state estimation, emphasizing its potential as a superior alternative in practical estimation tasks. For illustration purposes, Fig. 1 shows a sample of a trajectory of the state and of the state estimate of the EKF, eQEKF in a typical simulation of a realization of noise and initial condition.

[2]The realizations leading to the divergence of estimation errors were not taken into account in computing the MSE. We have detailed the percentage of stable trajectories of the estimation error in Table I.

The Kazufumi model: the second example is the one considered in [9] described by the equations

$$\begin{aligned}
x_1(k+1) &= 0.8x_1(k) + x_1(k)x_2(k) + 0.1 + v_1(k),\\
x_2(k+1) &= 1.5x_2(k) - x_1(k)x_2(k) + 0.1 + v_2(k),\\
y(k) &= x_2(k) + w(k).
\end{aligned}$$

As a comparison, here we have a nonlinear model with the partial knowledge of the system ($y(k) = x_2(k) + w(k)$) and with non-Gaussian noises.

We assume the noise sequences v_1, v_2 and w to be zero-mean discrete non-Gaussian sequences such that $\mathbb{P}(v_i = 10^{-2}) = 0.8$ and $\mathbb{P}(v_i = -4 \cdot 10^{-2}) = 0.2$ for $i = 1, 2$, and $\mathbb{P}(w = 10^{-1}) = 0.8$ and $\mathbb{P}(w = -4 \cdot 10^{-1}) = 0.2$. The initial condition $x(0)$ is specified as Gaussian with mean $\text{col}(0.8, 0.4)$ and a variance of 10^{-2}. In this case, all the filters, EKF, QEKF and the proposed eQEKF, demonstrate stability in all the realizations. The obtained MSEs are given in table I. It is evident that the proposed eQEKF enhances performance compared to both EKF and QEKF (12% and 8%, respectively). For illustration purposes, Fig. 2 shows a sample of a trajectory of the state and of the state estimate of the EKF, QEKF and eQEKF in a typical simulation of a realization of noise and initial condition.

To conclude, the simulations show that the QEKF estimate is often improved by the eQEKF which is also computationally cheaper, while EKF is still a good compromise since it provides relatively good estimates in spite of its minimal complexity compared to the quadratic filters.

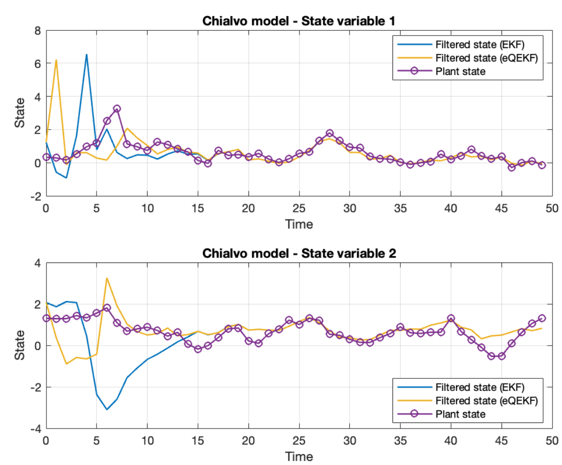

Fig. 1. Chialvo model: comparison between the filtered state of EKF (blue line), eQEKF (yellow line), and the true state of the plant system (purple line with circles), in a randomly sampled trajectory: state variable 1 (top panel) and state variable 2 (bottom panel).

VI. CONCLUSIONS

In this note, we have introduced the Enhanced Quadratic Extended Kalman Filter (eQEKF) as an alternative solution to the filtering problem for stochastic discrete-time nonlinear

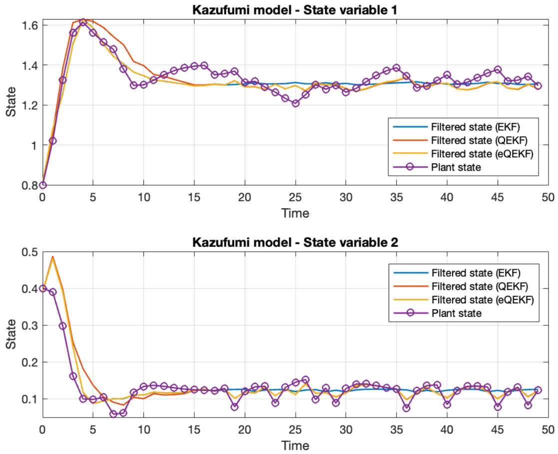

Fig. 2. Kazufumi model: comparison between the filtered state of EKF (blue line), QEKF (red line), eQEKF (yellow line), and the true state of the plant system (purple line with circles), in a randomly sampled trajectory: state variable 1 (top panel) and state variable 2 (bottom panel).

systems. Building upon the framework of [9], we proposed a new filter that deal with some inevitable approximation issues in a new way, reducing the computational workload while improving performance. Through two numerical examples, we have demonstrated the efficacy of the eQEKF in both cases of Gaussian and non-Gaussian noise, even with respect to the renowned EKF. The approach needs further investigations, from the extension of the algorithm to a generic polynomial order degree, to the study of the approximation of the expected value and the covariance of the state with alternatives methods.

References

[1] Derui Ding, Qing-Long Han, Zidong Wang, and Xiaohua Ge. A survey on model-based distributed control and filtering for industrial cyber-physical systems. *IEEE Transactions on Industrial Informatics*, 15(5):2483–2499, 2019.

[2] Huibert Kwakernaak and Raphael Sivan. *Linear optimal control systems*, volume 1072. Wiley-interscience, 1969.

[3] M. Sanjeev Arulampalam, Simon Maskell, Neil Gordon, and Tim Clapp. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188, 2002.

[4] John L. Maryak, James C. Spall, and Bryan D. Heydon. Use of the Kalman filter for inference in state-space models with unknown noise distributions. *IEEE Transactions on Automatic Control*, 49(1):87–90, 2004.

[5] Vladimir Stojanovic, Shuping He, and Baoyong Zhang. State and parameter joint estimation of linear stochastic systems in presence of faults and non-Gaussian noises. *International Journal of Robust and Nonlinear Control*, 30(16):6683–6700, 2020.

[6] Wenshuo Li, Zidong Wang, Yuan Yuan, and Lei Guo. Two-stage particle filtering for non-Gaussian state estimation with fading measurements. *Automatica*, 115:108882, 2020.

[7] Jingjing Wang, Jiaheng Li, Shefeng Yan, Wei Shi, Xinghai Yang, Ying Guo, and T. Aaron Gulliver. A novel underwater acoustic signal denoising algorithm for Gaussian/non-Gaussian impulsive noise. *IEEE Transactions on Vehicular Technology*, 70(1):429–445, 2021.

[8] Guoqing Wang, Yonggang Zhang, and Xiaodong Wang. Maximum correntropy Rauch–Tung–Striebel smoother for nonlinear and non-Gaussian systems. *IEEE Transactions on Automatic Control*, 66(3):1270–1277, 2021.

[9] Alfredo Germani, Costanzo Manes, and Pasquale Palumbo. Polynomial extended kalman filter. *IEEE Transactions on Automatic Control*, 50(12):2059–2064, 2005.

[10] Alberto De Santis, Alfredo Germani, and Massimo Raimondi. Optimal quadratic filtering of linear discrete-time non-Gaussian systems. *IEEE Transactions on Automatic Control*, 40(7):1274–1278, 1995.

[11] Francesco Carravetta, Alfredo Germani, and Massimo Raimondi. Polynomial filtering for linear discrete time non-Gaussian systems. *SIAM Journal on Control and Optimization*, 34(5):1666–1690, 1996.

[12] Francesco Carravetta, Alfredo Germani, and Massimo Raimondi. Polynomial filtering of discrete-time stochastic linear systems with multiplicative state noise. *IEEE Transactions on Automatic Control*, 42(8):1106–1126, 1997.

[13] Francesco Carravetta and Gabriella Mavelli. Minimax quadratic filtering of uncertain linear stochastic systems with partial fourth-order information. *IEEE transactions on automatic control*, 44(6):1287–1292, 1999.

[14] Alfredo Germani, Costanzo Manes, and Pasquale Palumbo. Polynomial filtering for stochastic non-Gaussian descriptor systems. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 51(8):1561–1576, 2004.

[15] Alfredo Germani, Costanzo Manes, and Pasquale Palumbo. State estimation of stochastic systems with switching measurements: a polynomial approach. *International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal*, 19(14):1632–1655, 2009.

[16] Qinyuan Liu, Zidong Wang, Qing-Long Han, and Changjun Jiang. Quadratic estimation for discrete time-varying non-Gaussian systems with multiplicative noises and quantization effects. *Automatica*, 113:108714, 2020.

[17] Filippo Cacace, Francesco Conte, Alfredo Germani, and Giovanni Palombo. Optimal linear and quadratic estimators for tracking from distance measurements. *Systems & Control Letters*, 139:104674, 2020.

[18] Stefano Battilotti, Filippo Cacace, Massimiliano d'Angelo, and Alfredo Germani. Cooperative filtering with absolute and relative measurements. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 7182–7187. IEEE, 2018.

[19] Filippo Cacace, Alfredo Germani, and Marco Papi. On parameter estimation of Heston's stochastic volatility model: a polynomial filtering method. *Decisions in Economics and Finance*, 42(2):503–525, 2019.

[20] Filippo Cacace, Francesco Conte, Alfredo Germani, and Giovanni Palombo. Quadratic filtering for non-gaussian and not asymptotically stable linear discrete-time systems. In *53rd IEEE Conference on Decision and Control*, pages 4995–5000. IEEE, 2014.

[21] Filippo Cacace, Francesco Conte, Alfredo Germani, and Giovanni Palombo. Feedback quadratic filtering. *Automatica*, 82:158–164, 2017.

[22] Stefano Battilotti, Filippo Cacace, Massimiliano d'Angelo, and Alfredo Germani. The polynomial approach to the LQ non-Gaussian regulator problem through output injection. *IEEE Transactions on Automatic Control*, 64(2):538–552, 2018.

[23] Filippo Cacace, Francesco Conte, Massimiliano d'Angelo, and Alfredo Germani. Filtering of systems with nonlinear measurements with an application to target tracking. *International Journal of Robust and Nonlinear Control*, 29(14):4956–4970, 2019.

[24] Filippo Cacace, Francesco Conte, Massimiliano d'Angelo, and Alfredo Germani. Feedback polynomial filtering and control of non-Gaussian linear time-varying systems. *Systems & Control Letters*, 123:108–115, 2019.

[25] Stefano Battilotti, Filippo Cacace, Massimiliano d'Angelo, Alfredo Germani, and Bruno Sinopoli. LQ non-Gaussian regulator with Markovian control. *IEEE Control Systems Letters*, 3(3):679–684, 2019.

[26] Stefano Battilotti, Filippo Cacace, Massimiliano d'Angelo, Alfredo Germani, and Bruno Sinopoli. LQ non-Gaussian control with I/O packet losses. In *2020 American Control Conference (ACC)*, pages 2802–2807. IEEE, 2020.

[27] Roger W Brockett. Volterra series and geometric control theory. *Automatica*, 12(2):167–176, 1976.

[28] Filippo Cacace, Valerio Cusimano, Alfredo Germani, Pasquale Palumbo, and Marco Papi. Optimal continuous-discrete linear filter and moment equations for nonlinear diffusions. *IEEE Transactions on Automatic Control*, 65(10):3961–3976, 2019.

[29] Dante R Chialvo and A Vania Apkarian. Modulated noisy biological dynamics: three examples. *Journal of Statistical Physics*, 70:375–391, 1993.