Accuracy Issues in Kalman Filtering State Estimation of Stiff Continuous-Discrete Stochastic Models Arisen in Engineering Research

Gennady Yu. Kulikov
Center for Computational and Stochastic Mathematics
Instituto Superior Técnico, Universidade de Lisboa
Lisbon, Portugal
gkulikov at math.ist.utl.pt

Maria V. Kulikova
Center for Computational and Stochastic Mathematics
Instituto Superior Técnico, Universidade de Lisboa
Lisbon, Portugal
maria.kulikova at ist.utl.pt

***Abstract*—This paper aims at exploring accuracy of Kalman-like filters. Its particular interest lies in estimation of stochastic systems whose drift coefficients expose a stiff behavior. The latter means that the Jacobian of the drift coefficient in such a continuous-discrete system, which is presented by an Itô-type stochastic differential equation (SDE) for modeling the plant's dynamic behavior and a discrete-time equation for simulating its measurement process, has large eigenvalues at the solution trajectory. Here, we employ the so-called "discrete-discrete" approach, which is grounded in SDE discretization schemes, and compare the outcome accuracy of EKF-, CKF- and UKF-type methods when these are based on the Euler-Maruyama and Itô-Taylor discretizations of the strong convergence orders 0.5 and 1.5 and applied for estimating the Van der Pol oscillator and Oregonator reaction models. We evidence that state estimation errors committed in our stiff stochastic scenarios are sensitive to both the type of Kalman filtering method utilized and the SDE discretization scheme implemented. So these must be chosen carefully in accurate and robust state estimation algorithms intended for treating stiff continuous-discrete stochastic systems.**

***Index Terms*—continuous-discrete stochastic system, extended Kalman filter, cubature Kalman filter, unscented Kalman filter**

I. INTRODUCTION

The Kalman filtering methods for treating stiff stochastic systems arisen in engineering research constitute a novel topic of state estimation theory discovered recently in [1]–[5]. The modern stochastic models admit often the following continuous-discrete formulation [6]:

$$dx(t) = F\big(t, x(t)\big)dt + G dw(t), \quad t > 0, \tag{1}$$

$$z_k = h(x_k) + v_k, \quad k \geq 1. \tag{2}$$

The process model in the stochastic system (1) and (2) is an Itô-type *Stochastic Differential Equation* (SDE). In this SDE, the drift coefficient $F : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ is time-variant and also depends on the state vector $x(t) \in \mathbb{R}^n$ evaluated at time t. Its driving noise term consists of the time-invariant diffusion matrix G of size $n \times q$ and the zero-mean white Gaussian process $\{w(t),\, t > 0\}$ with covariance $Q > 0$ of size $q \times q$, which is considered to be time-invariant, below. In this paper, the notation $Q > 0$ refers to the positive definiteness of the matrix Q. SDE (1) serves for simulating the plant's dynamic behavior. Its initial state x_0 is a random and normally distributed variable $x_0 \sim \mathcal{N}(\bar{x}_0, \Pi_0)$ with mean $\bar{x}_0$ and covariance $\Pi_0 > 0$.

The measurement model in the above stochastic system has the predefined discrete-time form (2), where k stands for a discrete time index (i.e. x_k means $x(t_k)$) and the differentiable measurement function $h : \mathbb{R}^n \to \mathbb{R}^m$ links the system's state x_k to the measurement information $z_k \in \mathbb{R}^m$ received at the sampling time t_k. We stress that the measurement vectors arrive noisy, i.e. these are corrupted by m-dimensional random vectors v_k whose distributions are assumed to be the Gaussian ones with zero mean and known covariance matrices $R_k > 0$. In what follows, we consider that the measurement information comes equidistantly and with the sampling rate (sampling period) $\delta = t_k - t_{k-1}$. In addition, all random processes in the stochastic system (1), (2) and its initial state are supposed to be mutually independent.

The main issue of our interest lies in studying the Kalman filtering' performance when the continuous-discrete stochastic system exposes a stiff behavior. In this case, the Jacobian of the drift function $F(t, x(t))$, which means its partial derivative with respect to the second argument, in SDE (1) has large eigenvalues at the state trajectory estimated [1], [3]–[5]. In particular, Kulikov and Kulikova [2], [3] discover that the traditional *Extended Kalman Filter* (EKF) used widely in applied science and engineering can outperform the modern and more accurate *Unscented Kalman Filter* (UKF) presented in [7]–[9] and the recently-designed *Cubature Kalman Filter* (CKF) [10], [11]. In the context of the continuous-discrete Kalman filtering, this counterintuitive phenomenon is justified from the stability point of view in [5].

Here, we focus upon the EKF-, CKF- and UKF-type methods built in a different way, which is referred to as the *discrete-discrete* approach. In contrast to the *continuous-discrete* Kalman filtering methods grounded in ODE solvers,

This work was supported in part by the Portuguese National Fund (*Fundação para a Ciência e a Tecnologia*) grant UID/Multi/04621/2013 and within the *Investigador FCT 2013* programme.

978-1-5386-4444-7/18/$31.00 ©2018 IEEE

the discrete-discrete ones are implemented by means of SDE discretization schemes, as that explained in [12], for instance.

We have seen that the EKF based on the Euler-Maruyama discretization of SDE (1) is more accurate and outperforms the CKF and UKF grounded in the Itô-Taylor expansion of order 1.5 in treating the stiff stochastic Van der Pol oscillator [2] and Oregonator reaction [3]. Besides, one might think that this counterintuitive result stems from the difference in the SDE discretization methods utilized, i.e. the Euler-Maruyama formula of order 0.5 (abbreviated to EM-0.5) and the Itô-Taylor one of order 1.5 (abbreviated to IT-1.5). It is known that IT-1.5 is theoretically more accurate than EM-0.5. In turn, the CKF and UKF outperform usually the traditional EKF in estimating nonstiff stochastic systems (see the above-cited literature). Thus, our purpose is to build the EKF with IT-1.5 discretization of SDE (1) and, in return, to implement the CKF and UKF with use of EM-0.5. We intend, first, for observing changes in the filters' performance which can stem from this alteration and, second, for making a comparison with the conclusions drawn in [2], [3] to see how the filters' accuracies depend on the SDE discretization methods applied.

II. THE NOVEL EKF-, CKF- AND UKF-BASED STATE ESTIMATION METHODS

In Sec. II, we describe the above-announced state estimators in detail and begin with our novel EKF method.

A. The Extended Kalman Filter Grounded in IT-1.5

As said in Sec. I, we restrict ourselves to the discrete-discrete fashion of the Kalman filtering under consideration. This entails that all discretization meshes used in our study should be of the equidistant form. In other words, for approximating SDE (1), any mesh consists of $m-1$ equally spaced subdivision nodes (with the user-supplied prefixed quantity m) in each sampling interval $[t_{k-1}, t_k]$ of size δ, i.e. its step size obeys the obvious formula $\tau = \delta/m$. We further apply the discretization method presented in [11] for converting the given SDE model to the discrete-time form as follows:

$$x_{k-1}^{l+1} = F_d\big(t_{k-1}^l, x_{k-1}^l\big) + \widetilde{G}w_1 + \mathbb{L}F\big(t_{k-1}^l, x_{k-1}^l\big)w_2. \quad (3)$$

In Eq. (3), the matrix $\widetilde{G} = GQ^{1/2}$ is calculated by a right-multiplication with $Q^{1/2}$ being the lower triangular factor in the Cholesky-type factorization of the covariance Q, which is termed the *square root*, below. In other words, we employ the covariance factorization $Q = Q^{1/2}Q^{\top/2}$ in which the notation $Q^{\top/2}$ refers to the transposed matrix of the square root $Q^{1/2}$. In addition, the discretized drift coefficient obeys the formula

$$\begin{aligned} F_d\big(t_{k-1}^l, x_{k-1}^l\big) &= x_{k-1}^l + \tau F\big(t_{k-1}^l, x_{k-1}^l\big) \\ &+ 0.5\tau^2 \mathbb{L}_0 F\big(t_{k-1}^l, x_{k-1}^l\big) \end{aligned} \quad (4)$$

in the outcome stochastic system. Here, the vector x_{k-1}^l refers to the IT-1.5-based approximation of the state $x(t_{k-1}^l)$ computed by method (3) and (4). Note it is evaluated at the discrete time $t_{k-1}^l = t_{k-1} + l\tau$, $l = 0, 1, \dots, m$. We recall that $F(\cdot)$ stands for the drift coefficient in the original process model (1) and $\tau = \delta/m$ is the step size of the user-supplied mesh, which underlies the m-step discretization scheme (3). The differential operators $\mathbb{L}_0$ and $\mathbb{L}_j$ appearing in formulas (3) and (4) are defined mathematically as follows:

$$\mathbb{L}_0 = \frac{\partial}{\partial t} + \sum_{i=1}^n F_i \frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{j,p,r=1}^n \widetilde{G}_{pj}\widetilde{G}_{rj}\frac{\partial^2}{\partial x_p \partial x_r}, \quad (5)$$

$$\mathbb{L}_j = \sum_{i=1}^n \widetilde{G}_{ij}\frac{\partial}{\partial x_i}, \quad j = 1, 2, \dots, n. \quad (6)$$

In Eqs (5) and (6), an (i,j)-entry in the above-introduced matrix $\widetilde{G}$ is denoted by $\widetilde{G}_{ij}$. The latter formulas imply that an (i,j)-entry in the matrix $\mathbb{L}F\big(t_{k-1}^l, x_{k-1}^l\big)$ of size $n \times n$ used in equation (3) is computed by the operator $\mathbb{L}_j F_i\big(t_{k-1}^l, x_{k-1}^l\big)$ from (6) and entries of the n-dimensional vector $\mathbb{L}_0 F\big(t_{k-1}^l, x_{k-1}^l\big)$ obey formula (5) in the discretized drift function (4). The pair of the correlated n-dimensional random Gaussian variables w_1 and w_2 is generated from the pair of the uncorrelated standard random Gaussian variables ν_1 and ν_2 by the rule: $w_1 = \sqrt{\tau}\nu_1$, $w_2 = \tau^{3/2}(\nu_1 + \nu_2/\sqrt{3})/2$. Therefore the covariance matrices of these random variables are calculated at once as follows:

$$\mathbf{E}\{w_1 w_1\} = \tau I_n,\ \mathbf{E}\{w_1 w_2\} = \frac{\tau^2}{2} I_n,\ \mathbf{E}\{w_2 w_2\} = \frac{\tau^3}{3} I_n \quad (7)$$

where the notation $\mathbf{E}\{\cdot\}$ stands for the expectation operator and I_n refers to the identity matrix of size n [11].

The first moment of the random variable x_{k-1}^{l+1} defined by the stochastic process (3) satisfies evidently the relation

$$\mathbf{E}\big\{x_{k-1}^{l+1}\big\} = \mathbf{E}\big\{F_d\big(t_{k-1}^l, x_{k-1}^l\big)\big\}, \quad (8)$$

where the drift coefficient $F_d\big(t_{k-1}^l, x_{k-1}^l\big)$ comes from formula (4), because both Gaussian variables w_1 and w_2 are zero-mean. For simplifying our notation, we set $\mathbf{E}\big\{x_{k-1}^{l+1}\big\} = \hat{x}_{k-1}^{l+1}$, $\mathbf{E}\big\{x_{k-1}^l\big\} = \hat{x}_{k-1}^l$ and so on, below.

Next, the state vector x_{k-1}^l is independent of the noises w_1 and w_2 in the stochastic system (3). Then, with use of covariances (7), Arasaratnam et al. [11] establish the following second moment recursion:

$$\begin{aligned} \mathbf{var}\big\{x_{k-1}^{l+1}\big\} = \mathbf{var}\Big\{ & x_{k-1}^l + \tau F\big(t_{k-1}^l, x_{k-1}^l\big) \\ & + \frac{\tau^2}{2}\mathbb{L}_0 F\big(t_{k-1}^l, x_{k-1}^l\big)\Big\} + \tau \widetilde{G}\widetilde{G}^\top \\ & + \frac{\tau^2}{2}\widetilde{G}\big[\mathbb{L}F\big(t_{k-1}^l, x_{k-1}^l\big)\big]^\top \\ & + \frac{\tau^2}{2}\big[\mathbb{L}F\big(t_{k-1}^l, x_{k-1}^l\big)\big]\widetilde{G}^\top \\ & + \frac{\tau^3}{3}\big[\mathbb{L}F\big(t_{k-1}^l, \hat{x}_{k-1}^l\big)\big]\big[\mathbb{L}F\big(t_{k-1}^l, \hat{x}_{k-1}^l\big)\big]^\top. \end{aligned} \quad (9)$$

The EKF technique implies that the moment equations (8) and (9) are solved approximately in each subdivision interval $[t_{k-1}^l, t_{k-1}^{l+1}]$ of the IT-1.5-based discretization (3). This is fulfilled by means of linearization of the drift coefficient

$F(t^l_{k-1}, x^l_{k-1})$ of SDE (1) around the state expectation $\hat{x}^l_{k-1}$ at each mesh's node t^l_{k-1}, i.e. one uses its Taylor expansion

$$\begin{aligned} F\big(t^l_{k-1}, x^l_{k-1}\big) &= F\big(t^l_{k-1}, \hat{x}^l_{k-1}\big) + \partial_x F\big(t^l_{k-1}, \hat{x}^l_{k-1}\big) \\ &\times \big(x^l_{k-1} - \hat{x}^l_{k-1}\big) + HOT \end{aligned} \quad (10)$$

where $\partial_x F(t^l_{k-1}, \hat{x}^l_{k-1}) = \partial F(t^l_{k-1}, \hat{x}^l_{k-1})/\partial x^l_{k-1}$ is the corresponding partial derivative (Jacobian) of the drift coefficient $F(t^l_{k-1}, x^l_{k-1})$ with respect to the variable x^l_{k-1} and evaluated at $(t^l_{k-1}, \hat{x}^l_{k-1})$ and HOT refers to higher-order terms in the above formula. Then, taking the linear part of expansion (10) into account and substituting it in the moment equations (8) and (9), we arrive at the following m-step IT-1.5-based EKF algorithm.

Algorithm 1. IT-1.5 EKF (*Conventional implementation*)

INITIALIZATION: $(k = 0)$ $\hat{x}_{0|0} := \bar{x}_0$, $P_{0|0} := \Pi_0$.
TIME UPDATE: $(k = \overline{1, K})$ $\triangleright$ PRIORI ESTIMATION

Set $\hat{x}^0_{k-1|k-1} := \hat{x}_{k-1|k-1}$ and $P^0_{k-1|k-1} := P_{k-1|k-1}$;
For $l = 0, 1, \ldots, m-1$ **do** (with $\tau := (t_k - t_{k-1})/m$)
 $\hat{x}^{l+1}_{k-1|k-1} := F_d\big(t^l_{k-1}, \hat{x}^l_{k-1}\big)$ at $t^l_{k-1} := t_{k-1} + l\tau$;
 $P^{l+1}_{k-1|k-1} := \big[M^l_{k-1}\big] P^l_{k-1|k-1} \big[M^l_{k-1}\big]^\top$
 $+\tau^2/2\Big(\widetilde{G}\big[\mathbb{L}F\big(t^l_{k-1}, \hat{x}^l_{k-1}\big)\big]^\top + \big[\mathbb{L}F\big(t^l_{k-1}, \hat{x}^l_{k-1}\big)\big]\widetilde{G}^\top\Big)$
 $+\tau\widetilde{G}\widetilde{G}^\top + \tau^3/3\big[\mathbb{L}F\big(t^l_{k-1}, x^l_{k-1}\big)\big]\big[\mathbb{L}F\big(t^l_{k-1}, x^l_{k-1}\big)\big]^\top$
 where $M^l_{k-1} := I_n + \big(\tau + \tau^2/2\big)\partial_x F\big(t^l_{k-1}, \hat{x}^l_{k-1}\big)$,
 $F_d\big(t^l_{k-1}, \hat{x}^l_{k-1}\big)$, $\mathbb{L}F\big(t^l_{k-1}, \hat{x}^l_{k-1}\big)$ are from (4), (6)
 and $\partial_x F\big(t^l_{k-1}, \hat{x}^l_{k-1}\big)$ is the Jacobian at $\hat{x}^l_{k-1|k-1}$;
Set $\hat{x}_{k|k-1} := \hat{x}^m_{k-1|k-1}$ and $P_{k|k-1} := P^m_{k-1|k-1}$.
MEASUREMENT UPDATE: $\triangleright$ POSTERIORI ESTIMATION
 $R_{e,k} := R_k + \big[\partial_x h(\hat{x}_{k|k-1})\big] P_{k|k-1} \big[\partial_x h(\hat{x}_{k|k-1})\big]^\top$;
 $K_k := P_{k|k-1}\big[\partial_x h(\hat{x}_{k|k-1})\big]^\top R^{-1}_{e,k}$;
 $\hat{x}_{k|k} := \hat{x}_{k|k-1} + K_k\big(z_k - h(\hat{x}_{k|k-1})\big)$;
 $P_{k|k} := P_{k|k-1} - K_k\big[\partial_x h(\hat{x}_{k|k-1})\big] P_{k|k-1}$
 where $\partial_x h(\hat{x}_{k|k-1})$ is the Jacobian at $\hat{x}_{k|k-1}$.

This algorithm computes the linear least-square estimate $\hat{x}_{k|k}$ of the state $x(t_k)$ subject to measurements $\{z_1, \ldots, z_k\}$.

B. The Cubature Kalman Filter Grounded in EM-0.5

In contrast to the CKF algorithm presented in [11], our new variant is grounded in the EM-0.5-based discretization. The latter method yields a discrete-time stochastic system the form

$$x^{l+1}_{k-1} = x^l_{k-1} + \tau F\big(t^l_{k-1}, x^l_{k-1}\big) + G\tilde{w}^l_{k-1} \quad (11)$$

with the discretized noise $\tilde{w}^l_{k-1} \sim \mathcal{N}(0, \tau Q)$ at each node of the equidistant mesh introduced in a sampling interval $[t_{k-1}, t_k]$. We emphasize that equation (11) gives the m-step discretization scheme and, hence, its step size $\tau = \delta/m$.

The main idea of the cubature Kalman filtering invented in [10] is to approximate the first two moments of the random variable x^{l+1}_{k-1}, which are n-dimensional Gaussian-weighted integrals, by means of the third-degree spherical-radial cubature rule in every subdivision step. For that, one defines first the cubature nodes (vectors)

$$\xi_i = \sqrt{n}e_i, \quad \xi_{n+i} = -\sqrt{n}e_i, \quad i = 1, 2, \ldots, n, \quad (12)$$

where e_i denotes the i-th unit coordinate vector in $\mathbb{R}^n$ and n refers to the size of SDE (1). Arasaratnam et al. [11, Appendix B] design their square-root CKF, which is particularly advantageous in practical implementations. So, we accommodate the cited method to the discrete-time stochastic system (11) and arrive at the following state estimation algorithm.

Algorithm 2. EM-0.5 CKF (*Square-root implementation*)

INITIALIZATION: $(k = 0)$
 Apply Cholesky: $\Pi_0 = \Pi_0^{1/2}\Pi_0^{\top/2}$, $Q = Q^{1/2}Q^{\top/2}$
 where $\Pi_0^{1/2}$ and $Q^{1/2}$ are lower triangular factors;
 Set $\hat{x}_{0|0} := \bar{x}_0$ and $P^{1/2}_{0|0} := \Pi_0^{1/2}$.
TIME UPDATE: $(k = \overline{1, K})$ $\triangleright$ PRIORI ESTIMATION
Set $\hat{x}^0_{k-1|k-1} := \hat{x}_{k-1|k-1}$ and $S^0_{k-1|k-1} := P^{1/2}_{k-1|k-1}$;
For $l = 0, 1, \ldots, m-1$ **do** (with $\tau := (t_k - t_{k-1})/m$)
 Generate vectors ξ_i $(i = 1, \ldots, 2n)$ by rule (12);
 Calculate $\mathcal{X}^l_{i,k-1|k-1} := S^l_{k-1|k-1}\xi_i + \hat{x}^l_{k-1|k-1}$;
 Matrix $\mathcal{X}^l_{k-1|k-1} := \big[\mathcal{X}^l_{1,k-1|k-1}, \ldots, \mathcal{X}^l_{2n,k-1|k-1}\big]$;
 Find $\mathcal{Y}^{l+1}_{i,k-1|k-1} := \mathcal{X}^l_{i,k-1|k-1} + \tau F\big(t^l_{k-1}, \mathcal{X}^l_{i,k-1|k-1}\big)$;
 Matrix $\mathcal{Y}^{l+1}_{k-1|k-1} := \big[\mathcal{Y}^{l+1}_{1,k-1|k-1}, \ldots, \mathcal{Y}^{l+1}_{2n,k-1|k-1}\big]$;
 Compute mean $\hat{x}^{l+1}_{k-1|k-1} := \mathcal{Y}^{l+1}_{k-1|k-1}\mathbf{1}_{2n}/(2n)$
 where $\mathbf{1}_{2n}$ is the unitary column-vector of size $2n$;
 $\mathbb{Y}^{l+1}_{k-1|k-1} := \big(\mathcal{Y}^{l+1}_{k-1|k-1} - \mathbf{1}^\top_{2n} \otimes \hat{x}^{l+1}_{k-1|k-1}\big)/\sqrt{2n}$
 where $\otimes$ is the Kronecker tensor product (`kron`);
 Collect the pre-array and transform
 $\big[\mathbb{Y}^{l+1}_{k-1|k-1}, \sqrt{\tau}GQ^{1/2}\big]\Theta_l = \big[S^{l+1}_{k-1|k-1}\big]$
 where Θ_l lower-triangularizes the left-hand matrix;
 Read-off from the post-array: $S^{l+1}_{k-1|k-1}$;
Set $\hat{x}_{k|k-1} := \hat{x}^m_{k-1|k-1}$ and $P^{1/2}_{k|k-1} := S^m_{k-1|k-1}$.
MEASUREMENT UPDATE: $\triangleright$ POSTERIORI ESTIMATION
 Calculate $\mathcal{X}_{i,k|k-1} := P^{1/2}_{k|k-1}\xi_i + \hat{x}_{k|k-1}$ $(i = \overline{1, 2n})$;
 Form matrix $\mathcal{X}_{k|k-1} := \big[\mathcal{X}_{1,k|k-1}, \ldots, \mathcal{X}_{2n,k|k-1}\big]$;
 Calculate $\mathcal{Z}_{i,k|k-1} := h\big(\mathcal{X}_{i,k|k-1}\big)$, $i = 1, 2, \ldots, 2n$;
 Form matrix $\mathcal{Z}_{k|k-1} := \big[\mathcal{Z}_{1,k|k-1}, \ldots, \mathcal{Z}_{2n,k|k-1}\big]$;
 Compute mean $\hat{z}_{k|k-1} := \mathcal{Z}_{k|k-1}\mathbf{1}_{2n}/(2n)$;
 Matrix $\mathbb{X}_{k|k-1} := \big(\mathcal{X}_{k|k-1} - \mathbf{1}^\top_{2n} \otimes \hat{x}_{k|k-1}\big)/\sqrt{2n}$;
 Matrix $\mathbb{Z}_{k|k-1} := \big(\mathcal{Z}_{k|k-1} - \mathbf{1}^\top_{2n} \otimes \hat{z}_{k|k-1}\big)/\sqrt{2n}$;
 Apply Cholesky decomposition: $R_k = R_k^{1/2}R_k^{\top/2}$;
 Collect the pre-array and transform

$$\begin{bmatrix} \mathbb{Z}_{k|k-1} & R_k^{1/2} \\ \mathbb{X}_{k|k-1} & \mathbf{0} \end{bmatrix} \Theta_k = \begin{bmatrix} R^{1/2}_{e,k} & \mathbf{0} \\ \bar{P}_{xz,k} & P^{1/2}_{k|k} \end{bmatrix}$$

 where Θ_k lower-triangularizes the left-hand matrix;
 Read-off from the post-array: $\bar{P}_{xz,k}$, $R^{1/2}_{e,k}$ and $P^{1/2}_{k|k}$;
 Calculate $\mathbb{W}_k := \bar{P}_{xz,k}R^{-1/2}_{e,k}$;
 Derive estimate $\hat{x}_{k|k} := \hat{x}_{k|k-1} + \mathbb{W}_k(z_k - \hat{z}_{k|k-1})$.

C. The Unscented Kalman Filter Grounded in EM-0.5

The *Unscented Kalman Filtering* (UKF) is designed as an advanced state estimation EKF alternative for treating discrete-time nonlinear stochastic systems [7]–[9]. It is grounded in the concept of *Unscented Transform* (UT) presented in the cited papers in detail. Briefly, the UT considers that the set of $2n+1$ deterministically selected sigma points (vectors) $\mathcal{X}_i$ is determined in line with the following rule:

$$\mathcal{X}_0 = \hat{x}, \; \mathcal{X}_i = \hat{x} + \sqrt{3}P^{1/2}e_i, \; \mathcal{X}_{i+n} = \hat{x} - \sqrt{3}P^{1/2}e_i, \quad (13)$$

$i = 1, 2, \dots, n$, where e_i denotes the i-th unit coordinate vector in $\mathbb{R}^n$, and $P^{1/2}$ refers to the covariance square root, which is the lower triangular factor in the Cholesky-type factorization of the covariance in the n-dimensional normally distributed variable $x \sim \mathcal{N}(\hat{x}, P)$, i.e. $P = P^{1/2}P^{\top/2}$. Below, we restrict ourselves to the classical UT parametrization, which stems from the coefficients

$$\begin{aligned} w_0^{(m)} &= \lambda/(n+\lambda), \; w_0^{(c)} = \lambda/(n+\lambda) + 1 - \alpha^2 + \beta, \\ w_i^{(m)} &= w_i^{(c)} = 1/(2n+2\lambda), \quad i = 1, \dots, 2n, \end{aligned} \quad (14)$$

with the constants fixed as follows: $\alpha = 1$, $\beta = 0$ and $\lambda = 3 - n$ (see the above-cited papers for further details).

Formulas (13) and (14) are utilized for presentation of the mean and covariance of the given random variable in the form

$$\hat{x} = \sum_{i=0}^{2n} w_i^{(m)}\mathcal{X}_i, \quad P = \sum_{i=0}^{2n} w_i^{(c)}(\mathcal{X}_i - \hat{x})(\mathcal{X}_i - \hat{x})^\top. \quad (15)$$

The main property of this UT is that coefficients (14) are preserved under any sufficiently smooth mapping $F(x)$ of the random variable x, but its sigma vectors (13) and the mean $\hat{x}$ must change to $F(\mathcal{X}_i)$, $i = 1, 2, \dots, n$, and $\widehat{F}(x)$, respectively. Then, the same formulas (15) are applied for approximation of the mean and covariance in the transformed random variable $F(x)$ as well (see more explanation in [7]–[9]).

For treating the above SDE models, the continuous-discrete UKF is designed in [13]. Here, in contrast to the cited paper, we implement our novel UKF method by means of the EM-0.5 discretization of SDE (1). In other words, we replace the latter continuous-time equation with the discrete-time one (11) and, then, apply the additive (zero-mean) noise case UKF algorithm in [8, Table 7.3] for treating the resulting discretized model. So, we arrive at the following EM-0.5-based UKF technique.

Algorithm 3. EM-0.5 UKF (*conventional implementation*)

INITIALIZATION: ($k = 0$) Set $\hat{x}_{0|0} := \bar{x}_0$, $P_{0|0} := \Pi_0$;

Define vector $W_m := \left[w_0^{(m)}, \dots, w_{2n}^{(m)}\right]^\top$ and matrix $\mathcal{W} := \left(I_{2n+1} - \mathbf{1}_{2n+1}^\top \otimes W_m\right)$ $\times \operatorname{diag}\left\{w_0^{(c)}, \dots, w_{2n}^{(c)}\right\}\left(I_{2n+1} - \mathbf{1}_{2n+1}^\top \otimes W_m\right)^\top$ where $\mathbf{1}_{2n+1}$ is the unitary column-vector of size $2n+1$, I_{2n+1} is the identity matrix of size $2n+1$, $\operatorname{diag}\left\{w_0^{(c)}, \dots, w_{2n}^{(c)}\right\}$ stands for the diagonal matrix and $\otimes$ is the Kronecker tensor product (`kron`).

TIME UPDATE: ($k = \overline{1, K}$) ▷ PRIORI ESTIMATION

Set $\hat{x}_{k-1|k-1}^0 := \hat{x}_{k-1|k-1}$ and $P_{k-1|k-1}^0 := P_{k-1|k-1}$;

For $l = 0, 1, \dots, m-1$ **do** (with $\tau := (t_k - t_{k-1})/m$)

 Factorize $P_{k-1|k-1}^l = \left[S_{k-1|k-1}^l\right]\left[S_{k-1|k-1}^l\right]^\top$;

 Generate sigma vectors $\mathcal{X}_{i,k-1|k-1}^l$ $(i = 0, \dots, 2n)$ by (13) with $P^{1/2} := S_{k-1|k-1}^l$ and $\hat{x} := \hat{x}_{k-1|k-1}^l$;

 Matrix $\mathcal{X}_{k-1|k-1}^l := \left[\mathcal{X}_{0,k-1|k-1}^l, \dots, \mathcal{X}_{2n,k-1|k-1}^l\right]$;

 Find $\mathcal{Y}_{i,k-1|k-1}^{l+1} := \mathcal{X}_{i,k-1|k-1}^l + \tau F\left(t_{k-1}^l, \mathcal{X}_{i,k-1|k-1}^l\right)$;

 Matrix $\mathcal{Y}_{k-1|k-1}^{l+1} := \left[\mathcal{Y}_{0,k-1|k-1}^{l+1}, \dots, \mathcal{Y}_{2n,k-1|k-1}^{l+1}\right]$;

 Compute mean $\hat{x}_{k-1|k-1}^{l+1} := \mathcal{Y}_{k-1|k-1}^{l+1} W_m$;

 $P_{k-1|k-1}^{l+1} := \left[\mathcal{Y}_{k-1|k-1}^{l+1}\right]\mathcal{W}\left[\mathcal{Y}_{k-1|k-1}^{l+1}\right]^\top + \tau GQG^\top$;

Set $\hat{x}_{k|k-1} := \hat{x}_{k-1|k-1}^m$ and $P_{k|k-1} := P_{k-1|k-1}^m$.

MEASUREMENT UPDATE: ▷ POSTERIORI ESTIMATION

 Cholesky factorization: $P_{k|k-1} = P_{k|k-1}^{1/2} P_{k|k-1}^{\top/2}$;

 Generate sigma vectors $\mathcal{X}_{i,k|k-1}$ $(i = 0, \dots, 2n)$ by (13) with $P^{1/2} := P_{k|k-1}^{1/2}$ and $\hat{x} := \hat{x}_{k|k-1}$;

 Form matrix $\mathcal{X}_{k|k-1} := \left[\mathcal{X}_{0,k|k-1}, \dots, \mathcal{X}_{2n,k|k-1}\right]$;

 Calculate $\mathcal{Z}_{i,k|k-1} := h\left(\mathcal{X}_{i,k|k-1}\right)$, $i = 0, 1, \dots, 2n$;

 Form matrix $\mathcal{Z}_{k|k-1} := \left[\mathcal{Z}_{0,k|k-1}, \dots, \mathcal{Z}_{2n,k|k-1}\right]$;

 Compute mean $\hat{z}_{k|k-1} := \mathcal{Z}_{k|k-1} W_m$;

 Find $P_{zz,k|k-1} := \mathcal{Z}_{k|k-1}\mathcal{W}\,\mathcal{Z}_{k|k-1}^\top + R_k$;

 Calculate $P_{xz,k|k-1} := \mathcal{X}_{k|k-1}\mathcal{W}\,\mathcal{Z}_{k|k-1}^\top$;

 Find $\mathbb{W}_k := P_{xz,k|k-1} P_{zz,k|k-1}^{-1}$;

 Compute estimate $\hat{x}_{k|k} := \hat{x}_{k|k-1} + \mathbb{W}_k(z_k - \hat{z}_{k|k-1})$;

 Calculate $P_{k|k} := P_{k|k-1} + \mathbb{W}_k P_{zz,k|k-1} \mathbb{W}_k^T$.

In the next section, we intend for examination of the filtering methods presented in Sec. II and their comparison to the earlier-published EKF, CKF and UKF techniques on stiff variants of the stochastic Van der Pol oscillator studied in [1], [2] and the stochastic Oregonator reaction considered in [3].

III. FILTERS' ACCURACY EXAMINATION

Here, our intention is to see the link of the filters' accuracies to the type of Kalman filtering implemented (i.e. EKF, CKF or UKF) and the SDE dicretization scheme utilized (i.e. EM-0.5 or IT-1.5). The filters' performances are assessed exactly in the way that has been already presented in [2], [3] at large. Briefly, we simulate true states of each stochastic system via solving the corresponding SDE (1) by the Euler-Maruyama method with the tiny step size equal to 10^{-5} in the simulation interval $[t_0, t_{end}]$ in MATLAB. These are used in the true measurement generation procedure of our numerical study and for evaluation of errors committed by the state estimators under consideration. Such errors are assessed on average in the form of the *Accumulated Root Mean Square Error* (ARMSE) derived via 100 independent Monte Carlo runs. For a fair comparison, all the filters utilize the same measurement histories and process and measurement noise realizations. Eventually, we come to the standard ARMSE estimation formula

$$\text{ARMSE} := \left(\frac{1}{100K}\sum_{l=1}^{100}\sum_{k=1}^{K}\sum_{i=1}^{n}\left(xi_{ref,l}(t_k) - \widehat{xi}_{k|k,l}\right)^2\right)^{1/2}$$

where the subscript ref distinguishes our reference stochastic solutions, i indicates the true and estimated state entries (i.e.

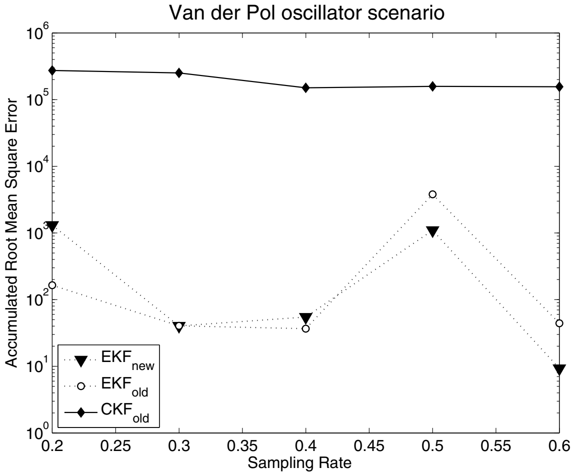

Fig. 1. The accuracies of the old and new EKF, CKF and UKF methods observed in our stiff stochastic Van der Pol oscillator scenario.

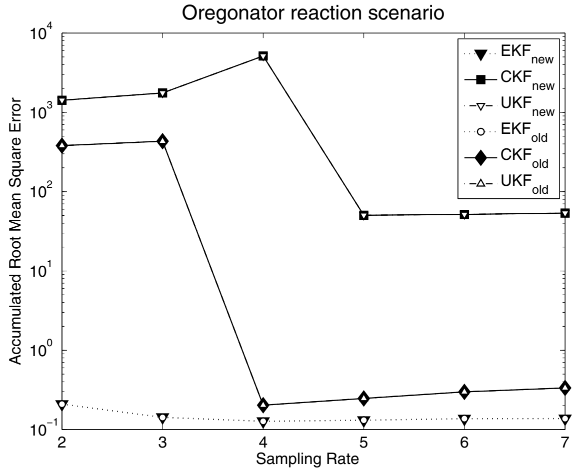

Fig. 2. The accuracies of the old and new EKF, CKF and UKF methods observed in our stiff stochastic Oregonator reaction scenario.

n is the size of the given SDE model), l refers to a particular Monte Carlo simulation, k marks the corresponding sampling instant t_k and $K := [(t_{end} - t_0)/\delta]$, where $[\cdot]$ stands for the integer part of the number and δ is a sampling rate accepted in our experiment.

A. The Stiff Stochastic Van Der Pol Oscillator Scenario

Following [1], [2], our first SDE model is taken as follows:

$$d\begin{bmatrix} x1(t) \\ x2(t) \end{bmatrix} = \begin{bmatrix} x2(t) \\ 10^4\big[\big(1 - x1^2(t)\big)x2(t) - x1(t)\big] \end{bmatrix} dt + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} dw(t)$$

with the noise $w(t) \sim \mathcal{N}(0, I_2)$ and the simulation interval $[0, 2]$. Similarly to the cited papers, the filters' initial values are set to be: $\bar{x}_0 = [\bar{x}1(0), \bar{x}2(0)]^\top = [2, 0]^\top$ and $\Pi_0 = \mathrm{diag}\{10^{-1}, 10^{-1}\}$. This SDE model is observed partially, i.e. with the noise $v_k \sim \mathcal{N}(0, 0.04)$, it entails the scheme

$$z_k = x1(t_k) + v_k.$$

The accuracies of all the EKF, CKF and UKF discussed in Sec. II and observed within our stiff Van der Pol oscillator scenario are exposed in Fig. 1. Following [2], these filters are implemented with 2×10^5 equidistant subdivisions of each sampling interval $[t_k, t_{k-1}]$, i.e. with $m = 2 \times 10^5$ in the above m-step filtering algorithms. The ARMSEs committed by our new filtering techniques have the subscript `new` in contrast to the errors of the remaining methods whose accuracies are copied from [2] and distinguished with the subscript `old` in Fig. 1. Here and below, all plots are scaled logarithmically.

First, in contrast to the IT-1.5-based CKF in [2], our novel EM-0.5-based CKF and UKF fail and return `NaN` at all sampling rates $\delta = 0.2, 0.3, \ldots, 0.6$ utilized in our Van der Pol oscillator scenario. That is why these data are absent in Fig. 1. Second, the situation becomes more complicated in the EKF-type Kalman filtering methods, i.e. the discretization EM-0.5 allows for the smaller ARMSEs observed at the shorter sampling times, whereas the other one IT-1.5 yields higher accuracies at the lower sampling rates considered here. All this can be explained by the larger discretization errors of EM-0.5. Third, both EM-0.5- and IT-1.5-based EKFs outperform the CKF and UKF with either discretization. The latter fact has been addressed for the continuous-discrete filtering in [5].

B. The Stiff Stochastic Oreginator Reaction Scenario

Following [3], our second SDE model is taken as follows:

$$d\begin{bmatrix} x1(t) \\ x2(t) \\ x3(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0.1 \end{bmatrix} dw(t) + \begin{bmatrix} 77.27\big[x2(t) + x1(t)\big(1 - 8.375\times 10^{-6} x1(t) - x2(t)\big)\big] \\ \big[x3(t) - x2(t)\big(1 + x1(t)\big)\big]/77.27 \\ 0.161\big(x1(t) - x3(t)\big) \end{bmatrix} dt$$

with the noise $w(t) \sim \mathcal{N}(0, I_3)$ and the simulation interval $[0, 60]$. The initial values are: $\bar{x}_0 = [\bar{x}1(0), \bar{x}2(0), \bar{x}3(0)]^\top = [4, 1.1, 4]^\top$, $\Pi_0 = \mathrm{diag}\{10^{-2}, 10^{-2}, 10^{-2}\}$ in all the filters. The above SDE model is supposed to be fully observed, i.e.

$$z_k = x1(t_k) + x2(t_k) + x3(t_k) + v_k$$

with the measurement noise $v_k \sim \mathcal{N}(0, 0.04)$. The ARMSEs of our filtering methods computed at the sampling rates $\delta = 2, 3, \ldots, 10$ in MATLAB are displayed in Fig. 2. Following [3], these filters are applied with $m = 5 \times 10^5$ equidistant subdivisions of each sampling interval $[t_k, t_{k-1}]$, i.e. with $m = 5 \times 10^5$ in the above m-step filtering algorithms.

In our stiff Oregonator scenario, the IT-1.5-based CKF and UKF methods are more accurate in comparison to the EM-0.5-based ones, again. Besides, the EKF-type estimators expose the same ARMSEs irrespective of the discretization used. Nevertheless, similarly to the outcome of Sec. III-A, the EKF-type methods outperform always the CKF and UKF ones with either discretization in our stiff stochastic Oregonator reaction scenario.

IV. CONCLUSION

This paper has studied state estimation accuracies of some discrete-discrete filters depending on the type of Kalman filtering technique implemented (i.e. EKF, CKF or UKF) and on the SDE dicretization scheme utilized (i.e. EM-0.5 or IT-1.5) when applied to treating stiff continuous-discrete stochastic systems. In particular, we have developed the EM-0.5-based CKF and UKF methods and the EKF grounded in IT-1.5, as well. These novel state estimators have been used for solving two stiff practical problems in electrical and chemical engineering, namely, for estimating the stochastic Van der Pol oscillator and Oregonator reaction models considered also in [2] and [3], respectively. In addition, the committed state estimation errors have been compared to those derived in the cited papers. Eventually, we have concluded that the EKF-type state estimation technique and the SDE discretization IT-1.5 are preferable and produce more accurate state estimates in both stiff stochastic scenarios considered in our case study.

REFERENCES

[1] G. Yu. Kulikov and M. V. Kulikova, "Estimating the state in stiff continuous-time stochastic systems within extended Kalman filtering," *SIAM J. Sci. Comput.*, vol. 38, no. 6, pp. A3565–A3588, 2016.

[2] ——, "Do the cubature and unscented Kalman filtering methods outperform always the extended Kalman filter?" *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 3762–3767, 2017.

[3] ——, "Accurate state estimation of stiff continuous-time stochastic models in chemical and other engineering," *Math. Comput. Simulation*, vol. 142, pp. 62–81, 2017.

[4] ——, "Square-root Kalman-like filters for estimation of stiff continuous-time stochastic systems with ill-conditioned measurements," *IET Control Theory Appl.*, vol. 11, no. 9, pp. 1420–1425, 2017.

[5] ——, "Stability analysis of Extended, Cubature and Unscented Kalman Filters for estimating stiff continuous-discrete stochastic systems," *Automatica*, vol. 90, pp. 91–97, 2018.

[6] S. Särkkä, "On unscented Kalman filter for state estimation of continuous-time nonlinear systems," *IEEE Trans. Automat. Contr.*, vol. 52, no. 9, pp. 1631–1641, Sep. 2007.

[7] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Trans. Automat. Contr.*, vol. 45, no. 3, pp. 477–482, Mar. 2000.

[8] E. A. Wan and R. Van der Merwe, "The unscented Kalman filter," in S. Haykin ed. *Kalman Filtering and Neural Networks*. New York: John Wiley & Sons, Inc., 2001, pp. 221–280.

[9] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, Mar. 2004.

[10] I. Arasaratnam and S. Haykin, "Cubature Kalman filters," *IEEE Trans. Automat. Contr.*, vol. 54, no. 6, pp. 1254–1269, Jun. 2009.

[11] I. Arasaratnam, S. Haykin, and T. R. Hurd, "Cubature Kalman filtering for continuous-discrete systems: Theory and simulations," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 4977–4993, Oct. 2010.

[12] G. Yu. Kulikov and M. V. Kulikova, "Accurate numerical implementation of the continuous-discrete extended Kalman filter," *IEEE Trans. Automat. Contr.*, vol. 59, no. 1, pp. 273–279, Jan. 2014.

[13] ——, "The accurate continuous-discrete extended Kalman filter for radar tracking," *IEEE Trans. Signal Process.*, vol. 64, no. 4, pp. 948–958, Feb. 2016.