www.elsevier.com/locate/ejcon

High-order accurate continuous-discrete extended Kalman filter for chemical engineering

G.Yu. Kulikov, M.V. Kulikova

PII: S0947-3580(14)00095-8
DOI: http://dx.doi.org/10.1016/j.ejcon.2014.11.003
Reference: EJCON105

To appear in: *European Journal of Control*

Received date: 8 February 2014
Revised date: 21 November 2014
Accepted date: 30 November 2014

Cite this article as: G.Yu. Kulikov, M.V. Kulikova, High-order accurate continuous-discrete extended Kalman filter for chemical engineering, *European Journal of Control*, http://dx.doi.org/10.1016/j.ejcon.2014.11.003

High-order accurate continuous-discrete extended Kalman filter for chemical engineering

G. Yu. Kulikov[a,*], M. V. Kulikova[a]

[a] *CEMAT, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal.*

Abstract

This paper elaborates a new version of extended Kalman filtering (EKF) for state estimation in chemical nonlinear continuous-discrete stochastic systems. Such a state estimation always compounds real measurements of some system's variables (depending on the utilized technology) with computation of remaining (not measurable) parameters by means of appropriate filtering algorithms. Here, we consider the continuous-discrete EKF and show that its quality is raised by using the adaptive sixth-order nested implicit Runge-Kutta (NIRK) method of Gauss type with automatic local and global error controls. Through case studies the new filtering technology is compared to another EKF implementation based on an adaptive ODE solver but with the sole local error control. Our numerical results exhibit that the designed state estimation algorithm not only outperforms the earlier published adaptive EKF method, but also resolves a so-called "EKF failure" case reported recently.

Keywords: Extended Kalman filter, continuous-discrete model, moment differential equations, adaptive ODE solver, automatic local and global error control

1. Introduction

Mathematical models in chemical research and industrial applications are often presented in the form of stochastic differential equation (SDE)

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

where $x(t) \in \mathbb{R}^{n_1}$ is the n_1-dimensional vector of system's state at time t, $u(t) \in \mathbb{R}^{n_2}$ is the measurable input at time t, $F : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}^{n_1}$ is a nonlinear function representing the chemical reaction kinetics, $G\big(x(t), u(t)\big)$ is a matrix of dimension $n_1 \times q$ and $\{w(t),\, t > 0\}$ is a Brownian process with square diffusion matrix $Q(t) \geq 0$ of the size q. We point out that chemical process models are based on conservation laws and described conventionally by ordinary differential equations (ODEs), whereas the stochastic term in SDE (1) simulates possible random disturbances and uncertainties in the reaction and also a plant-model mismatch always existing in reality.

Here, we deal with mathematical models of the form (1) where the measurable input $u(t)$ is assumed to be a known function of time. In other words, we consider that its value is known at any

*Corresponding author.

Email addresses: gkulikov@math.ist.utl.pt (G. Yu. Kulikov), maria.kulikova@ist.utl.pt (M. V. Kulikova)

time instant. If the discussed chemical process model does not correspond to a particular situation, i.e. when the input of chemical system is unknown and should be measured in real experimentation, one can treat such a chemical system by augmenting the system's state with the unknown input entries. In this case, the input $u(t)$ is evaluated as a part of the augmented state vector $x(t)$. Then, removing the term $u(t)$ from all the below formulas allows the augmented chemical model to be estimated by the designed method as well.

The initial state x_0 of chemical process (1) is supposed to be a random variable. More precisely, $x_0 \sim \mathcal{N}(\bar{x}_0, \Pi_0)$ with $\Pi_0 \geq 0$, where the notation $\mathcal{N}(\bar{x}_0, \Pi_0)$ stands for the normal distribution with mean $\bar{x}_0$ and covariance Π_0.

The task of state estimation in chemical system (1) always compounds real measurements of measurable system's variables (depending on the utilized technology) with computation of remaining (not measurable) parameters by means of an appropriate nonlinear filtering algorithm. It is usually assumed that some observation information arrives discretely and in equidistant intervals of size $\delta = t_k - t_{k-1}$. This time interval δ is called the sampling period (or waiting time) in filtering theory. The relation of observations y_k to the state vector x_k in chemical system (1) is fixed by the formula

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

where k stands for a discrete time index (i.e. x_k means $x(t_k)$), $y_k \in \mathbb{R}^m$ is the information available at time t_k, $h : \mathbb{R}^{n_1} \to \mathbb{R}^m$ is a linear or nonlinear function and the measurement noise v_k is a zero-mean Gaussian white-noise process with covariance matrix $R_k > 0$. We emphasize that formula (2) covers both linear and nonlinear observation models. Also, all realizations of $w(t)$, v_k and x_0 are assumed to be taken from mutually independent Gaussian distributions. Thus, the continuous-discrete stochastic state-space model (1), (2) is best suited for state estimation in chemical systems and widely used in chemistry research and industrial applications (see, for instance, Wilson et al. (1998); Soroush (1998); Dochain (2003); Haseltine and Rawlings (2002, 2005); Jørgensen (2007); Rawlings and Bakshi (2006); Romanenko and Castro (2004); Romanenko et al. (2004)).

Concerning state estimation algorithms, we have to remark that, at present, there exist a great variety of different methods starting from a rigorous probabilistic approach solving Kolmogorov's (Fokker-Planck's) forward equation (as discussed, for instance, in Jazwinski (1970); Maybeck (1982)) till approximate approaches including various nonlinear modifications and implementations of the well-known Kalman filter (see Lewis (1986); Singer (2002, 2006); Julier et al. (2000); Julier and Uhlmann (2004); Ito and Xiong (2000); Nørgaard et al. (2000); Arasaratnam and Haykin (2008, 2009); Arasaratnam et al. (2010); Frogerais et al. (2012); Jørgensen et al. (2007); Kulikov and Kulikova (2014); Rawlings and Bakshi (2006); Romanenko and Castro (2004); Romanenko et al. (2004); Schneider and Georgakis (2013)) as well as optimization based approaches usually referred to as the moving horizon estimation (studied by Jang et al. (1986); Rao et al. (2001); Haseltine and Rawlings (2002, 2005); Rawlings and Bakshi (2006) and so on). Undoubtedly, the extended Kalman filter (EKF) still remains among the most popular and widely used numerical techniques for practical state estimation in nonlinear stochastic systems because of its implementation simplicity and good performance. It originates from the optimum state estimation theory developed by Kalman (1960) in linear discrete-time stochastic state-space systems. The progress made in the Kalman filtering by now has resulted in a number of fast, numerically stable and many other algorithms (see, for instance, Lewis (1986); Kailath et al. (2000); Grewal and Andrews (2001); Goodwin and Sin (1984); Maybeck (1982); Simon (2006)).

Despite EKF's popularity, this method has been criticized on its performance for offline models

and industrial applications in chemical research by Wilson et al. (1998); Soroush (1998); Dochain (2003); Haseltine and Rawlings (2002, 2005); Jørgensen (2007); Rawlings and Bakshi (2006); Romanenko and Castro (2004); Romanenko et al. (2004). For example, Haseltine and Rawlings (2002, 2005) report that their EKF fails for two types of chemical reactors meaning that wrong steady-states are calculated and negative concentrations are observed after convergence, which are of no physical sense. Jørgensen (2007) claims that his EKF is not able to reconstruct offset free concentrations in the Van der Vusse reaction scenario on the basis of temperature measurements, only. One more difficulty mentioned in relation to the EKF is that it may fail for nonlinear systems with infrequent observations. So, Soroush (1998) writes: "In the chemical/petrochemical and biochemical industries, there are many processes wherein the choice of sampling rate is limited by the availability of the output measurements. For example, composition analyzers such as gas chromatographs have a cycle time say 5-10 min compared to a desired control interval of say 0.1-1 min. If the control interval is increased to match the availability of measurements then control performance deteriorates significantly."

Recently, Kulikov and Kulikova (2014) presented a way to resolve some cases of the "EKF failure" by means of adaptive ODE solvers with automatic error control. Therefore, the task of searching for the most appropriate ODE solver in the frame of EKF technology has arisen. We emphasize that the formulated task corresponds well to practitioners' expectations in control theory, as, for instance, stated in Arasaratnam et al. (2010). Below, we contribute to the research topic announced in the cited paper and conduct a detailed theoretical and numerical study of two efficient ODE solvers. The first one is grounded in the embedded pair of explicit-first-stage singly diagonal implicit Runge-Kutta (ESDIRK) methods of orders 3 and 4 and published in Kristensen et al. (2004). It is further denoted as `ESDIRK3(4)`. The second ODE solver is hybrid and designed here on the basis of two different schemes for computing the predicted state expectation and the predicted error covariance matrix. The predicted state mean is calculated by an embedded Runge-Kutta pair of orders 4 and 6. The higher-order method in this pair belongs to the family of nested implicit Runge-Kutta (NIRK) formulas of Gauss type. These NIRK schemes are introduced and studied by Kulikov and Shindin (2009); Kulikov (2009, 2013) at large. The predicted error covariance matrix is determined by the corresponding part of the numerical scheme designed and explored in Mazzoni (2007), but it is modified for a square-root implementation of the EKF in this paper. So, the latter hybrid ODE solver is further referred to as `NIRK6(4)M2`. In the next section, we present the `ESDIRK3(4)` and `NIRK6(4)M2` based EKF variants with all necessary implementation particulars for a detailed theoretical and numerical comparison and study.

It is worthwhile to remark that our `NIRK6(4)M2` method distinguishes from what was used and published earlier in Kulikov and Kulikova (2014). First, the embedded NIRK pair of orders 2 and 4 with the global error control from Kulikov (2013) was applied for the simultaneous solution of the state mean and error covariance equations in the earlier published research. This resulted in an accurate but time-consuming state estimator where the authors solved linear systems of size $n_1(n_1+3)/2$ with n_1 standing for the dimension of the state vector. The global error was controlled in all entries of the moment differential equations. Here, we design another method that treats the state mean and error covariance equations, separately. The global error control is also implemented in numerical integration of the first moment differential equation, only. All this reduces the cost of `NIRK6(4)M2` in comparison to the earlier published state estimator. Second, Kulikov and Kulikova (2014) studied the standard continuous-discrete EKF technology whereas the present paper deals with a more stable square-root implementation of the EKF, as explained in Section 2.3.

At the end of this introduction, we point out that the purpose of the present paper is not to address the above-mentioned "EKF failure" phenomena, but to look for a more efficient version of the EKF, which may be successful in practice. In addition, we expose that the state estimator designed here works well for offline chemical models and, hence, may be potentially useful in industrial environment. Any comparison of our EKF technique to other effective nonlinear state estimation algorithms, as, for example, particle, unscented and ensemble filters, and its practical testing in real experimentation are beyond the scope of this paper and expected in future.

2. Theory and implementation

2.1. Continuous-discrete extended Kalman filter

Frogerais et al. (2012) identify two main approaches for implementing the EKF method (namely, the continuous-discrete and discrete-discrete EKF implementations), and Kulikov and Kulikova (2014) explain that the continuous-discrete EKF is more accurate and reliable in practice. Therefore, it is most suited for treating the continuous-discrete stochastic state-space model (1), (2) arising often in chemical research and industrial applications. We restrict ourselves to the latter state estimation technology and present its implementation particulars, below.

It is well-known from the cited literature that the continuous-discrete EKF is based on replacement of the predicted values of state mean and error covariance matrix determined in the time-update step of the Kalman filtering with the values satisfying the moment differential equations (MDEs)

$$\frac{d\hat{x}(t)}{dt} = F\big(\hat{x}(t),u(t)\big), \tag{3a}$$

$$\frac{dP(t)}{dt} = J\big(\hat{x}(t),u(t)\big)P(t) + P(t)J^T\big(\hat{x}(t),u(t)\big) + G\big(\hat{x}(t),u(t)\big)Q(t)G^T\big(\hat{x}(t),u(t)\big) \tag{3b}$$

where $J\big(\hat{x}(t),u(t)\big)$ denotes the Jacobian of the drift function $F\big(\hat{x}(t),u(t)\big)$ form the SDE (1) (i.e. $J\big(\hat{x}(t),u(t)\big) = \partial F\big(\hat{x}(t),u(t)\big)/\partial\hat{x}(t)\big)$, $G\big(\hat{x}(t),u(t)\big)$ is the matrix from the stochastic noise term of this equation, $Q(t)$ is the covariance matrix of the zero-mean Gaussian white-noise process $w(t)$, $\hat{x}(t)$ means the state expectation of the random system's state vector $x(t)$ at time t (i.e. $x(t)$ is a solution to the SDE (1)), and $u(t)$ is the measurable input (i.e. the known function of time), as explained above. The matrix $P(t)$ in formula (3b) has the physical meaning of being the variance of the state prediction error, i.e. $x(t)-\hat{x}(t)$, and, then, has to be positive semi-definite.

One uses values of the state mean and error covariance matrix estimated at the previous sampling point t_{k-1} as initial values for the MDEs (3) formulated in the time interval $[t_{k-1},t_k]$, i.e. $\hat{x}(t_{k-1}) = \hat{x}_{k-1|k-1}$, $P(t_{k-1}) = P_{k-1|k-1}$. Then, having solved this initial value problem in the mentioned interval, one determines predicted values of the state mean and error covariance matrix at the sampling instant t_k as follows: $\hat{x}_{k|k-1} = \hat{x}(t_k)$, $P_{k|k-1} = P(t_k)$. After arrival a new measurement y_k, the measurement-update step of the Kalman filtering is performed in the form of the following algorithm:

$$R_{e,k} = R + H_kP_{k|k-1}H_k^T, \quad K_k = P_{k|k-1}H_k^TR_{e,k}^{-1}, \tag{4a}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_ke_k, \quad e_k = y_k - h(\hat{x}_{k|k-1}), \tag{4b}$$

$$P_{k|k} = P_{k|k-1} - K_kH_kP_{k|k-1} \tag{4c}$$

where the Jacobian $H_k = dh(\hat{x}_{k|k-1})/dx_k$, and $e_k \sim \mathcal{N}(0, R_{e,k})$ are innovations of the Kalman filter. Eventually, the linear least-square estimate $\hat{x}_{k|k}$ of the system's state $x(t_k)$ based on measurements $Y^k = \{y_1, \ldots, y_k\}$ has been computed.

The main difficulty associated with state estimator (3), (4) is that the arisen MDEs are often nonlinear in chemical research and industrial applications and, hence, should be treated numerically. Strictly speaking, the computed predicted state mean $\hat{x}_{k|k-1}$ and the error covariance matrix $P_{k|k-1}$ do not further satisfy the MDEs (3) because they are always determined with some error, which is referred to as the discretization error. The magnitude of the discretization error depends on the quality of the implemented MDE solver. Unfortunately, large discretization errors of inaccurate numerical MDE solutions might reduce accuracy of the corresponding EKF based state estimation, significantly. For example, Frogerais et al. (2012); Kulikov and Kulikova (2014) show divergence of the EKF on a number of practical models when the MDEs are solved for one step of the size δ and the sampling period δ is not too small. The theory presented by Hairer et al. (1993); Hairer and Wanner (2010); Butcher (2008) says that small discretization errors in numerical methods applied to the MDEs will be ensured only for sufficiently short waiting times δ if the above-mentioned fixed-stepsize implementation is utilized. However, it is absolutely impossible to identify *a priori* a proper size of the sampling period δ ensuring small discretization errors in practice because it depends on the MDEs (3) themselves, i.e. it is a problem-dependent value.

Kulikov and Kulikova (2014) suggest using more advanced and accurate ODE solvers with automatic global error control to overcome the so-called "curse of discretization" and to improve the accuracy and robustness of the continuous-discrete EKF method. In particular, the cited EKF implementation is able to regulate automatically the above-mentioned discretization error, i.e. the user needs only to set the required accuracy of numerical integration (one number) and the code does all remaining work for computing numerical solutions to the MDEs (3) with the error corresponding to the set accuracy level. This boosts the accuracy and reliability of the entire state estimation algorithm (3), (4), considerably. The only drawback associated with the MDE solver utilized by Kulikov and Kulikova (2014) is that the positive semi-definiteness of the predicted error covariance matrix $P(t)$ in formula (3b) is not guaranteed in the course of numerical integration. However, the latter property is important for the proper work of the EKF method. We further resolve this inconsistency by means of our new adaptive `NIRK6(4)M2` triple with automatic global error control presented in Section 2.3 in detail. To exhibit advantages of the implemented global error control mechanism, we compare it to the other adaptive `ESDIRK3(4)` method. The latter MDE solver is implemented with only local error control. Jørgensen et al. (2007) have shown its superiority to the built-in Matlab ODE solver `ODE15s` in terms of efficiency of computation. However, our numerical experiments in Section 3.2 say that the accuracy of `ESDIRK3(4)` may compromise the reliability of state estimation in chemical and other engineering. We further discuss implementation particulars of both MDE solvers at large.

2.2. Adaptive ESDIRK3(4) MDE solver with local error control

The first EKF method is built by Jørgensen et al. (2007) and presents the computationally efficient and robust implementation of the continuous-discrete EKF grounded in the `ESDIRK3(4)` embedded pair with local error control. The latter adaptive MDE solver is designed by Kristensen et al. (2004). Moreover, in order to boost the robustness of their `ESDIRK3(4)` based continuous-discrete EKF method against round-off, Jørgensen et al. (2007) suggest propagating not the error covariance matrix itself but its square root, only. Unfortunately, the square-root implementation presented in the concluding part of Section IIB of the cited paper is not obvious because the

coefficient b_2 of the ESDIRK method of order 3 is negative (see Kristensen et al. (2004)) and the authors do not explain what the square root of this negative real number means. Therefore, we further restrict ourselves to the non-square-root `ESDIRK3(4)` based continuous-discrete EKF method, which is denoted as `EKF-ESDIRK3(4)` in the discussion and numerical experiments of Section 3. The latter state estimator is implemented and run with the accuracy parameter `Tol` $= 10^{-4}$. For more particulars and explanation of this non-square-root `EKF-ESDIRK3(4)` technique and for its precise implementation algorithm, the readers are referred to Section 2.5 in Kulikov and Kulikova (2015).

Jørgensen et al. (2007) exhibit that their `ESDIRK3(4)` based EKF method outperforms the `ODE15s` based continuous-discrete EKF in terms of efficiency of computation. On the other hand, the implemented local error control does not allow the true error of numerical solutions to MDE (3a) to be effectively regulated. For instance, Kulikov and Weiner (2011) present three numerical examples for which the well-known and efficient code `RODAS` with local error control (see details in Hairer and Wanner (2010)) exhibits ratios "True Error"/"Tolerance" in the range from 10 to 10^{+8} even for sufficiently stringent accuracy conditions (i.e. for small values of "Tolerance"). We recall that "Tolerance" means the size of tolerated local errors in the course of numerical integration (see more explanation in Kulikov and Kulikova (2015)). This disadvantage of the local error control implemented in the `ESDIRK3(4)` MDE solver may yield its divergence, i.e. huge total errors produced by the entire `EKF-ESDIRK3(4)` technique for difficult SDE models. That is why we further present an MDE solver that controls additionally the global (true) error of numerical solution. The latter facility enables an efficient regulation of the discretization error in the course of numerical integration of MDEs and leads to the new concept of *accurate continuous-discrete extended Kalman filtering*. The superiority of this nonlinear filtering algorithm is discussed in Section 3.1 and illustrated on numerical examples in Section 3.2 of the paper.

2.3. Adaptive NIRK6(4)M2 MDE solver with local and global error controls

To introduce the adaptive hybrid `NIRK6(4)M2` solver with the combined local and global error control, we remark firstly that MDEs (3) are rather independent and can easily be decoupled, as explained in Mazzoni (2007). More precisely, Eq. (3a) does not depend on the other MDE in system (3) and, hence, can be solved separately. This is important since formula (3a) is the unique nonlinear equation in the discussed MDEs. Formula (3b) represents the linear differential matrix equation with respect to the error covariance. Therefore, a numerical treatment of MDE (3a) introduces the main part of discretization error and one has to pay a special attention to it. Here, we implement the advanced ODE solver `NIRK6(4)` with the automatic global error control presented and examined extensively in Kulikov (2013). As we have already mentioned in the concluding part of Section 1, the higher-order numerical method underlying `NIRK6(4)` belongs to the efficient family of NIRK formulas of Gauss type. The latter numerical schemes possess many useful properties such as A-stability, stiff accuracy and symmetry, which allow for successful numerical treatments of various MDEs, including stiff and very stiff equations. The NIRK formulas are also applicable to Hamiltonian and reversible problems and to differential-algebraic systems as well. We stress that the NIRK method utilized for computing numerical solutions to MDE (3a) is of stage order 3, classical order 6 and it does not extend the dimension of MDEs. All these properties are proved theoretically and examined numerically in Kulikov and Shindin (2009); Kulikov (2009).

To describe the solution method at large, we suppose further that the state evaluation has been completed at time instant t_{k-1}. Hence, filtered values of the state mean $\hat{x}_{k-1|k-1}$ and the error covariance matrix $P_{k-1|k-1}$ are available. So, our intention is to calculate a numerical solution to

the MDEs (3) with the initial values $\hat{x}(t_{k-1}) = \hat{x}_{k-1|k-1}$, $P(t_{k-1}) = P_{k-1|k-1}$ at the next sampling point t_k. At first, we assume that a mesh

$$\{t_l\}_{l=0}^{L} = \{t_0 = t_{k-1},\ t_{l+1} = t_l + \tau_l,\ l = 0, 1, \ldots, L-1,\ t_L = t_k\} \tag{5}$$

is introduced in the integration interval $[t_{k-1}, t_k]$. We point out that the introduced mesh is supposed to be variable. The latter means that the step size τ_l may vary as much as we want. In the concluding part of this section, we explain how to generate such a mesh in automatic mode.

Having discretized the MDE (3a) by means of the Gauss-type NIRK method of order 6 on mesh (5), we yield the nonlinear problem

$$\hat{x}_{l1}^2 = a_{11}^2\hat{x}_l + a_{12}^2\hat{x}_{l+1} + \tau_l\big[d_{11}^2F(\hat{x}_l, u_l) + d_{12}^2F(\hat{x}_{l+1}, u_{l+1})\big], \tag{6a}$$

$$\hat{x}_{l2}^2 = a_{21}^2\hat{x}_l + a_{22}^2\hat{x}_{l+1} + \tau_l\big[d_{21}^2F(\hat{x}_l, u_l) + d_{22}^2F(\hat{x}_{l+1}, u_{l+1})\big], \tag{6b}$$

$$\begin{aligned}\hat{x}_{l1}^3 &= a_{11}^3\hat{x}_l + a_{12}^3\hat{x}_{l+1} + \tau_l\big[d_{11}^3F(\hat{x}_l, u_l) + d_{12}^3F(\hat{x}_{l+1}, u_{l+1})\\ &+ d_{13}^3F(\hat{x}_{l1}^2, u_{l1}^2) + d_{14}^3F(\hat{x}_{l2}^2, u_{l2}^2)\big],\end{aligned} \tag{6c}$$

$$\begin{aligned}\hat{x}_{l2}^3 &= a_{21}^3\hat{x}_l + a_{22}^3\hat{x}_{l+1} + \tau_l\big[d_{21}^3F(\hat{x}_l, u_l) + d_{22}^3F(\hat{x}_{l+1}, u_{l+1})\\ &+ d_{23}^3F(\hat{x}_{l1}^2, u_{l1}^2) + d_{24}^3F(\hat{x}_{l2}^2, u_{l2}^2)\big],\end{aligned} \tag{6d}$$

$$\begin{aligned}\hat{x}_{l3}^3 &= a_{31}^3\hat{x}_l + a_{32}^3\hat{x}_{l+1} + \tau_l\big[d_{31}^3F(\hat{x}_l, u_l) + d_{32}^3F(\hat{x}_{l+1}, u_{l+1})\\ &+ d_{33}^3F(\hat{x}_{l1}^2, u_{l1}^2) + d_{34}^3F(\hat{x}_{l2}^2, u_{l2}^2)\big],\end{aligned} \tag{6e}$$

$$\hat{x}_{l+1} = \hat{x}_l + \tau_l\left[b_1F(\hat{x}_{l1}^3, u_{l1}^3) + b_2F(\hat{x}_{l2}^3, u_{l2}^3) + b_3F(\hat{x}_{l3}^3, u_{l3}^3)\right] \tag{6f}$$

for each subscript $l = 0, 1, 2, \ldots, L-1$ with respect to an unknown vector $\hat{x}_{l+1}$ approximating the state mean at time t_{l+1}. The stage value $\hat{x}_{lj}^i$ of method (6) implies an approximation to the state expectation $\hat{x}(t_{lj}^i)$ evaluated at the time point $t_{lj}^i = t_l + c_j^i\tau_l$, and the measurable input $u_{lj}^i = u(t_{lj}^i)$, $j = 1, 2$, $i = 2, 3$. We recall that $u(t)$ is the known function of time (see Section 1). As customary, the function F represents the right-hand side of the MDE (3a) or, this is the same, the drift function in the SDE (1). All constant coefficients of method (6) are defined as follows: $b_1 = b_3 = 5/18$, $b_2 = 4/9$, $c_1^2 = (3-\sqrt{3})/6$, $c_2^2 = (3+\sqrt{3})/6$, $a_{11}^2 = a_{22}^2 = 1/2 + 2\sqrt{3}/9$, $a_{12}^2 = a_{21}^2 = 1/2 - 2\sqrt{3}/9$, $d_{11}^2 = -d_{22}^2 = (3+\sqrt{3})/36$, $d_{12}^2 = -d_{21}^2 = (-3+\sqrt{3})/36$, $c_1^3 = (5-\sqrt{15})/10$, $c_2^3 = 1/2$, $c_3^3 = (5+\sqrt{15})/10$, $a_{11}^3 = a_{32}^3 = (125+39\sqrt{15})/250$, $a_{12}^3 = a_{31}^3 = (125-39\sqrt{15})/250$, $a_{21}^3 = a_{22}^3 = 1/2$, $d_{11}^3 = -d_{32}^3 = (7+2\sqrt{15})/200$, $d_{12}^3 = -d_{31}^3 = (-7+2\sqrt{15})/200$, $d_{13}^3 = -d_{34}^3 = (18\sqrt{15}+15\sqrt{3})/1000$, $d_{14}^3 = -d_{33}^3 = (18\sqrt{15}-15\sqrt{3})/1000$, $d_{21}^3 = -d_{22}^3 = 1/32$, $d_{23}^3 = -d_{24}^3 = 3\sqrt{3}/32$. We stress that formulas (6) present the nested 3-level form of the Gauss-type NIRK method of order 6, which is preferable for practical utilization (see more explanation in Kulikov and Shindin (2009)).

Method (6) seems to be complicated for implementation. However, Kulikov (2009) has shown that it admits the following simplified Newton iteration, which is efficient in practice:

$$\hat{x}_{l1}^{2,\ell} = a_{11}^2\hat{x}_l^4 + a_{12}^2\hat{x}_{l+1}^{\ell} + \tau_l\big[d_{11}^2F(\hat{x}_l^4, u_l) + d_{12}^2F(\hat{x}_{l+1}^{\ell}, u_{l+1})\big], \tag{7a}$$

$$\hat{x}_{l2}^{2,\ell} = a_{21}^2\hat{x}_l^4 + a_{22}^2\hat{x}_{l+1}^{\ell} + \tau_l\big[d_{21}^2F(\hat{x}_l^4, u_l) + d_{22}^2F(\hat{x}_{l+1}^{\ell}, u_{l+1})\big], \tag{7b}$$

$$\hat{x}_{l1}^{3,\ell} = a_{11}^3\hat{x}_l^4 + a_{12}^3\hat{x}_{l+1}^\ell + \tau_l\big[d_{11}^3F(\hat{x}_l^4,u_l) + d_{12}^3F(\hat{x}_{l+1}^\ell,u_{l+1}) + d_{13}^3F(\hat{x}_{l1}^{2,\ell},u_{l1}^2) + d_{14}^3F(\hat{x}_{l2}^{2,\ell},u_{l2}^2)\big], \tag{7c}$$

$$\hat{x}_{l2}^{3,\ell} = a_{21}^3\hat{x}_l^4 + a_{22}^3\hat{x}_{l+1}^\ell + \tau_l\big[d_{21}^3F(\hat{x}_l^4,u_l) + d_{22}^3F(\hat{x}_{l+1}^\ell,u_{l+1}) + d_{23}^3F(\hat{x}_{l1}^{2,\ell},u_{l1}^2) + d_{24}^3F(\hat{x}_{l2}^{2,\ell},u_{l2}^2)\big], \tag{7d}$$

$$\hat{x}_{l3}^{3,\ell} = a_{31}^3\hat{x}_l^4 + a_{32}^3\hat{x}_{l+1}^\ell + \tau_l\big[d_{31}^3F(\hat{x}_l^4,u_l) + d_{32}^3F(\hat{x}_{l+1}^\ell,u_{l+1}) + d_{33}^3F(\hat{x}_{l1}^{2,\ell},u_{l1}^2) + d_{34}^3F(\hat{x}_{l2}^{2,\ell},u_{l2}^2)\big], \tag{7e}$$

$$\Big[I_{n_1} - \frac{\tau_l}{6}J\big(\hat{x}_l^4,u_l\big)\Big]^3\big(\hat{x}_{l+1}^{\ell+1} - \hat{x}_{l+1}^\ell\big) = -\hat{x}_{l+1}^\ell + \hat{x}_l^4 + \tau_l\Big[b_1F(\hat{x}_{l1}^{3,\ell},u_{l1}^3) + b_2F(\hat{x}_{l2}^{3,\ell},u_{l2}^3) + b_3F(\hat{x}_{l3}^{3,\ell},u_{l3}^3)\Big], \quad \ell = 0,1,2,3. \tag{7f}$$

The matrix $J(\hat{x}_l^4,u_l)$ in linear problem (7f) means the Jacobian defined after formulas (3) and evaluated at the point $(\hat{x}_l^4,u_l)$, and I_{n_1} stands for the identity matrix of size n_1. Formulas (7) imply that 4 iteration steps are performed per mesh node and $\hat{x}_{l+1}^4$ is taken as the outcome approximation to the state mean at time point t_{l+1}. Hence, the superscript ℓ marks the current iterate. Notice that iteration (7) is implemented with trivial predictor, i.e. the solution $\hat{x}_l^4$ from the previous numerical integration step is considered as its initial guess or, more formally, $\hat{x}_{l+1}^0 = \hat{x}_l^4$.

It is worthwhile to mention that iteration (7) is as cheap as the iteration recommended for the ESDIRK method by Jørgensen et al. (2007) (see also implementation details of the ESDIRK solver in Kulikov and Kulikova (2015)). This is because the cubic matrix $\big[I_{n_1} - \tau_l J\big(\hat{x}_l^4,u_l\big)/6\big]^3$ must not be evaluated for solving system (7f). This linear problem is treated effectively as follows:

$$\Big[I_{n_1} - \frac{\tau_l}{6}J\big(\hat{x}_l^4,u_l\big)\Big]\Big[I_{n_1} - \frac{\tau_l}{6}J\big(\hat{x}_l^4,u_l\big)\Big]\Big[I_{n_1} - \frac{\tau_l}{6}J\big(\hat{x}_l^4,u_l\big)\Big]\big(\hat{x}_{l+1}^{\ell+1} - \hat{x}_{l+1}^\ell\big) = RHS$$

where RHS stands for the right-hand side of formula (7f). Thus, the latter representation of linear system (7f) implies that one iterate of the simplified Newton method (7) is calculated by means of successive solution of three linear systems with the same coefficient matrix $I_{n_1} - \tau_l J\big(\hat{x}_l^4,u_l\big)/6$. If we now take into account that 4 iteration steps of the iterative scheme (7) (with the initial guess $\hat{x}_{l+1}^0 = \hat{x}_l^4$) are sufficient for providing the sixth-order convergence of method (6) (this follows from Theorem 2.3 in Kulikov et al. (2007)) we will conclude that advancing a step of our NIRK method demands one evaluation and LU-factorization of the matrix $I_{n_1} - \tau_l J\big(\hat{x}_l^4,u_l\big)/6$ and 12 solutions of the linear systems with the same factorized coefficient matrix. This is exactly how the ESDIRK method of order 3 has been implemented (see Section 2.5 in Kulikov and Kulikova (2015)). In other words, both methods studied in this paper compute the predicted state expectation at time t_{l+1} for a similar amount of computational work, but with different accuracy. The ESDIRK method gives the third-order approximation to the mentioned state mean whereas the NIRK method ensures the sixth-order convergence. This is the evident superiority of the hybrid `NIRK6(4)M2` MDE solver designed here. More advantages and disadvantages of the `ESDIRK3(4)` and `NIRK6(4)M2` based EKF variants are discussed in Section 3.1 al large.

Further, we explain how to treat numerically the matrix MDE (3b), which is solved after each step of the above method (6). We recall that this differential problem is linear. So even simple

numerical schemes can be utilized successfully. Here, we implement the corresponding part of the hybrid method presented by Mazzoni (2007), which, for Eq. (3b), reads

$$\begin{aligned} P_{l+1} &= M_{l+1/2}P_lM_{l+1/2}^T+\tau_lK_{l+1/2}G\big(\hat{x}_{l+1/2},u(t_{l+1/2})\big) \\ &\times\ Q(t_{l+1/2})G^T\big(\hat{x}_{l+1/2},u(t_{l+1/2})\big)K_{l+1/2}^T \end{aligned} \tag{8}$$

where $t_{l+1/2}=t_l+\tau_l/2$ is the mid-point of the current step of size τ_l. The variable matrices defined at $t_{l+1/2}$ are calculated by the formulas:

$$K_{l+1/2} = \Big[I_{n_1}-\frac{\tau_l}{2}J\big(\hat{x}_{l+1/2},u(t_{l+1/2})\big)\Big]^{-1}, \tag{9a}$$

$$M_{l+1/2} = K_{l+1/2}\Big[I_{n_1}+\frac{\tau_l}{2}J\big(\hat{x}_{l+1/2},u(t_{l+1/2})\big)\Big] \tag{9b}$$

where $J\big(\hat{x}_{l+1/2},u(t_{l+1/2})\big)$ denotes the conventional Jacobian evaluated at the point $\big(\hat{x}_{l+1/2},u(t_{l+1/2})\big)$. We emphasize that, in contrast to the hybrid numerical scheme designed by Mazzoni (2007) where the mid-point state expectation $\hat{x}_{l+1/2}$ is not available and a special (time-consuming) polynomial interpolation formula is utilized to compute it in addition, our method (6) allows one stage value to be reused for approximation of the state mean at the mid-point of each step of the method. More formally, we utilize $\hat{x}_{l+1/2}=\hat{x}_{l2}^{3,4}$ in formulas (9). This follows from the definition of the mentioned stage value done after method (6) and the value of the coefficient c_2^3 given there. Notice that the accuracy of this stage value suffices for the second-order convergence of the method (8), (9) proved by Mazzoni (2007). The latter fact follows from Theorem 3.1 in Kulikov (2013). We also remark that formulas (7a)–(7d) are applied one extra time to increase the accuracy of calculation of the state mean at the mid-point $t_{l+1/2}$.

Mazzoni (2007) proves that his numerical approximation (8), (9) of the MDE (3b) is A-stable, and also preserves the symmetry and positive semi-definiteness of the predicted error covariance matrix P_{l+1} for any step size τ_l in exact arithmetic provided that the matrix P_l is symmetric and positive semi-definite. This is the other advantage of our hybrid `NIRK6(4)M2` solver because the positive semi-definiteness of the error covariance matrix P_{l+1} from the ESDIRK method is not guaranteed (see the detailed explanation in Kulikov and Kulikova (2015)). This positive semi-definiteness is not only the necessary condition for a proper performance of the EKF method (3), (4) but it is also important for a square-root implementation of the numerical scheme (8). It means that not the error covariance matrix itself but only its square root denoted by $P_l^{1/2}$ is propagate in time. Many authors present a clear evidence that the square-root implementation allows the round-off error to be effectively reduced and, hence, the accuracy and reliability of the Kalman filter and its sensitivity evaluation to be raised, considerably (see, for instance, Bierman (1977); Bierman et al. (1990); Kailath et al. (2000); Grewal and Andrews (2001); Kulikova (2009); Kulikova and Pacheco (2013); Tsyganova and Kulikova (2013)).

Again, the idea of square-root implementation of the Kalman filter elaborated in the cited literature is grounded in finding the square root of the initial covariance matrix and only this square root is propagated in time. The latter requires some change in the measurement-update step of the EKF method discussed in Section 2.1, as shown in the detailed filtering algorithm presented in Appendix A. Initially, the Cholesky decomposition of the covariance matrix Π_0 is implemented for converting it to the form $\Pi_0=\Pi_0^{T/2}\Pi_0^{1/2}$ with $\Pi_0^{1/2}$ standing for the upper triangular matrix. The square roots $Q^{1/2}(t)$ and $R_k^{1/2}$ of the covariance matrices of the process and measurement noises are

determined, similarly. Moreover, the matrices Π_0, $Q(t)$ and R_k are diagonal and with nonnegative entries in all numerical examples of Section 3.2. So, their square roots are obviously identified. Then, the predicted square root $P_{l+1}^{1/2}$ is computed at all nodes of mesh (5) as follows:

$$\Theta_l \begin{pmatrix} P_l^{1/2} M_{l+1/2}^T \\ Q^{1/2}(t_{l+1/2}) G^T\big(\hat{x}_{l+1/2}, u(t_{l+1/2})\big) K_{l+1/2}^T \sqrt{\tau_l} \end{pmatrix} = \begin{pmatrix} P_{l+1}^{1/2} \\ 0 \end{pmatrix} \tag{10}$$

where the matrices $K_{l+1/2}$ and $M_{l+1/2}$ are evaluated by formulas (9) and Θ_l means an orthogonal rotation that upper-triangularizes the left-hand matrix of this formula. The requested square root $P_{l+1}^{1/2}$ appears as the result of this triangularization.

The most complicated part of the hybrid `NIRK6(4)M2` MDE solver is the algorithm of automatic generation of mesh (5) in each sampling subinterval $[t_{k-1}, t_k]$. We emphasize that the accuracy of the numerical solution to the MDEs (3) is entirely determined by the quality of the mesh generation procedure. Here, we accommodate the effective error estimation and control mechanism that equips the embedded pair `NIRK6(4)` in Kulikov (2013) to produce a good mesh in automatic mode.

As we said above, only the local and global errors of the numerical solution to the MDE (3a) are to be evaluated and regulated in the course of numerical integration, because of its nonlinearity. Thus, having completed the iteration (7) we derive the predicted state expectation vector $\hat{x}_{l+1}^4$ at mesh node t_{l+1} and also the third-level stage values $\hat{x}_{l1}^{3,4}$, $\hat{x}_{l2}^{3,4}$, $\hat{x}_{l3}^{3,4}$ of this method. We emphasize that formulas (7a)–(7e) have been applied one more time to determine these more accurate stage values for the output approximate expectation $\hat{x}_{l+1}^4$. Then, the local error le_{l+1} associated with $\hat{x}_{l+1}^4$ is evaluated as follows:

$$le_{l+1} = -\frac{\tau_l}{3}\Big[\frac{1}{2}F(\hat{x}_l^4, u_l) - \frac{5}{6}F(\hat{x}_{l1}^{3,4}, u_{l1}^3) + \frac{2}{3}F(\hat{x}_{l2}^{3,4}, u_{l2}^3) - \frac{5}{6}F(\hat{x}_{l3}^{3,4}, u_{l3}^3) + \frac{1}{2}F(\hat{x}_{l+1}^4, u_{l+1})\Big]. \tag{11}$$

Kulikov and Weiner (2011); Kulikov (2013) explain that controlling the local error of numerical schemes does not ensure that the actual error of numerical integration is as small as one wants *a priori*. Therefore, the best solution is to evaluate and regulate additionally the so-called global (or true) error of the numerical method. Here, we utilize the cheap technology designed for the embedded pair `NIRK6(4)` in Kulikov (2013). It is based on summation of the local error estimates (11). More precisely, we denote the global error at mesh node t_{l+1} by $\Delta\hat{x}_{l+1}$ and, then, evaluate it by the formula

$$\Delta\hat{x}_{l+1} = \Delta\hat{x}_l + le_{l+1} \tag{12}$$

where the initial integration error $\Delta\hat{x}_0$ is set to be zero. We recall that the numerical solution of the initial value problem (3a), whose initial value is known from the measurement-update step of the EKF algorithm (3), (4) in the previous waiting period, i.e. $\hat{x}_0 = \hat{x}_{k-1|k-1}$, is discussed here. Hence, we do not use the numerical integration scheme `NIRK6(4)` to calculate it. That is why $\Delta\hat{x}_0 = 0$ is set in the beginning of numerical integration on each sampling interval $[t_{k-1}, t_k]$.

Again, formula (12) does not calculate the total error of the `EKF-NIRK6(4)M2` technique, but only its part that comes from the variable-stepsize ODE solver `NIRK6(4)` applied to the state expectation equation (3a). The local and global errors estimated by formulas (11) and (12), respectively, are utilized for regulating the accuracy of the numerical integration procedure, as shown in Algorithm 3.2 in Kulikov (2013). Having used the cited stepsize selection algorithm we arrive at the `NIRK6(4)M2` based *array square-root accurate continuous-discrete extended Kalman filter*, which is presented with all necessary implementation details in Appendix A.

Table 1: Theoretical comparison of the ESDIRK3(4) and NIRK6(4)M2 MDE solvers

	Property	ESDIRK3(4)	NIRK6(4)M2
1.	Convergence order	3	6
2.	Stage order	2	3
3.	A-stability	+	+
4.	L-stability	+	–
5.	Stiff accuracy	+	+
6.	Symmetry	–	+
7.	Applicability to Hamiltonian systems	–	+
8.	Applicability to reversible systems	–	+
9.	Local error control	+	+
10.	Global error control	–	+
11.	Symmetry of the error covariance matrix	+	+
12.	Positive semi-definiteness of the error covariance matrix	?	+
13.	Square-root implementation	?	+

Finally, we stress that the EKF-NIRK6(4)M2 method does not require any manual tuning, maybe, except for altering the value of one parameter ϵ_g, which restricts the magnitude of the committed global errors in the course of numerical integration of MDEs and should be set by the user. In this paper, $\epsilon_g = 10^{-4}$. It is also worthwhile to mention that the implemented local and global error control mechanism has already been presented within Runge-Kutta formulas, Nordsieck schemes and two-step peer W-methods and examined extensively on difficult test problems (see Kulikov (2013, 2012); Kulikov and Weiner (2011)). That is why it is expected to work successfully for short and sufficiently long sampling intervals δ, including variable waiting periods. Further, we discuss advantages and disadvantages of the EKF-ESDIRK3(4) and EKF-NIRK6(4)M2 algorithms in detail.

3. Discussion and comparison

3.1. Theoretical comparison

Main theoretical properties of the EKF-ESDIRK3(4) and EKF-NIRK6(4)M2 filtering technologies are completely determined by the underlying MDE solvers. For comparison, they are summarized in Table 1, where "+" means that the corresponding property is available, "–" implies that the corresponding property is absent, and the question sign "?" says that the corresponding property is compromised.

Having analyzed Table 1 we conclude that the MDE solver NIRK6(4)M2 is superior to the ESDIRK3(4) in all properties except for L-stability. However, this property is important only for solving very stiff ODEs. Actually, the absence of L-stability does not influence performance of our hybrid solver for stiff MDEs because Kulikov and Shindin (2009) prove that their NIRK formula of order 6 is A-stable and stiffly accurate. The latter properties admit its efficient application to very stiff ODEs and to differential-algebraic systems as well (see more explanation, for instance, in Hairer and Wanner (2010); Butcher (2008)). This is also confirmed numerically by Kulikov (2009).

In contrast, the MDE solver built here works well for reversible and Hamiltonian problems for which the `ESDIRK3(4)` is not effective. Also, the `NIRK6(4)M2` is more accurate than the other solver, and, what is more important, it includes the automatic global error control mechanism, which is not available in the method constructed by Kristensen et al. (2004). The latter means that the accuracy of the `EKF-ESDIRK3(4)` may be insufficient to solve successfully difficult SDE models of the form (1), whereas the `EKF-NIRK6(4)M2` method is effective and accurate. Moreover, the `EKF-ESDIRK3(4)` does not ensure the positive semi-definiteness of the predicted error covariance, that may also affect the accuracy of state estimation in practice. Nevertheless, both utilized MDE solvers are adaptive and implement the local error control for regulating the size of current step in automatic mode. This feature allows some "EKF failures" reported in the literature to be overcome. In Section 3.2, we illustrate all these conclusions on a mathematical model from chemical engineering and also on one difficult artificial SDE.

Concerning efficiency of the `EKF-ESDIRK3(4)` and `EKF-NIRK6(4)M2` methods in terms of time expenditures, we recall first that the cost of one integration step in solving Eq. (3a) by the MDE solvers `ESDIRK3(4)` and `NIRK6(4)M2` is almost the same, as explained in Section 2.3. However, our `EKF-NIRK6(4)M2` technique demands additional QR factorizations when solving the error covariance equation (3b) and in the measurement-update step of this algorithm (see Appendix A for further details). This means that every step of the new state estimator is expected to be more expensive in comparison to the `EKF-ESDIRK3(4)` designed earlier. However, the `EKF-NIRK6(4)M2` method is much more stable to round-off errors than the other algorithm. Furthermore, time expenditures of these filters depend mostly not on the complexity of one step of each state estimator but determined by the number of steps, i.e. by the mesh (5) generated in every sampling interval $[t_{k-1}, t_k]$ in automatic mode. On one hand, the NIRK method is of the higher convergence order (see Table 1) and, hence, allows for larger step sizes in comparison to the `ESDIRK3(4)` for the same accuracy conditions. The latter reduces the number of steps in the generated mesh and saves the execution time, dramatically. On the other hand, the implemented global error control may impose more stringent accuracy conditions and, hence, require more careful step sizes when solving MDEs (3) and even their repeated integrations. All this might increase the filtering time of the `EKF-NIRK6(4)M2`. This is a usual trade-off between the accuracy and the speed. Eventually, we arrive at the conclusion that the `EKF-NIRK6(4)M2` can be a complicated and time-consuming but accurate and reliable state estimation technique. Actually, its computational complexity creates no obstacle to powerful computational devices available at present. The difference in the execution time is not visible for all SDE models treated on a conventional home PC in Section 3.2, and the time of measurement simulation by the Euler-Maruyama method is longer than the state estimation time of any CD-EKF method discussed in this paper.

3.2. Numerical comparison

At first, we confirm numerically that the `EKF-ESDIRK3(4)` and `EKF-NIRK6(4)M2` filtering technologies resolve successfully the "EKF failure" reported by Jørgensen (2007). The cited paper claims: "However, even though it is in principle possible to estimate the unmeasured states from the subset of measured states, it is often so in practice that in the face of model-plant mismatch the result is quite disappointing in the sense that one cannot substitute concentration measurements with a soft sensor as the EKF [26]. There is simply no substitute for a good sensor (except for the Utopian wish for a perfect model.)" Then, Jørgensen (2007) concludes: "The continuous-discrete extended Kalman filter is applied to the Van der Vusse benchmark example. In this example, temperature measurements are *not* sufficient to provide steady-state offset free state estimation".

Fortunately, our own calculations made by means of the `EKF-ESDIRK3(4)` and `EKF-NIRK6(4)M2` do not confirm the cited claim. That is why we explore this test example in detail. It is also worthwhile to mention that the Van der Vusse reaction is a classical benchmark in a number of control studies (see, for instance, Chen et al. (1995); Engell and Klatt (1993); Jørgensen (2007); Ogunnaike and Ray (1994)). The exploited state estimators are coded and run in Matlab.

Following Jørgensen (2007), the Van der Vusse reaction consists of four species denoted by A, B, C and D. The desired product is B, while C and D are unwanted by-products. This reaction is conducted in a continuously stirred tank reactor (CSTR) with a cooling jacket and simulated by the following ODEs:

$$
\begin{aligned}
c_A'(t) &= \frac{F}{V_R}\big[c_{A0} - c_A(t)\big] - k_{10}\exp\left\{-\frac{E_1}{RT}\right\}c_A(t) - k_{30}\exp\left\{-\frac{E_3}{RT}\right\}c_A^2(t), && \text{(13a)}\\
c_B'(t) &= -\frac{F}{V_R}c_B(t) + k_{10}\exp\left\{-\frac{E_1}{RT}\right\}c_A(t) - k_{20}\exp\left\{-\frac{E_2}{RT}\right\}c_B(t), && \text{(13b)}\\
T'(t) &= \frac{F}{V_R}\big[T_0 - T'(t)\big] + \frac{k_w A_R}{\rho C_p V_R}\big[T_J(t) - T(t)\big] - \left[k_{10}\Delta H_{r_1}\exp\left\{-\frac{E_1}{RT}\right\}c_A(t)\right.\\
&+ \left. k_{20}\Delta H_{r_2}\exp\left\{-\frac{E_2}{RT}\right\}c_B(t) + k_{30}\Delta H_{r_3}\exp\left\{-\frac{E_3}{RT}\right\}c_A^2(t)\right]/\big[\rho C_p\big], && \text{(13c)}\\
T_J'(t) &= \frac{1}{m_J C_{PJ}}\Big(\dot{Q}_J + k_w A_R\big[T(t) - T_J(t)\big]\Big). && \text{(13d)}
\end{aligned}
$$

Values of all constant parameters in the mathematical model (13) and its steady-state (called the nominal operating point) are taken from Tables I and II in Jørgensen (2007). In other words, the initial values of ODEs (13) are given as follows:

$$
c_A = 2.1404\,\text{mol/L},\ c_B = 1.0903\,\text{mol/L},\ T = 387.34\,\text{K},\ T_J = 386.06\,\text{K}. \tag{14}
$$

So, each entry of the solution vector to problem (13) will be the straight line corresponding to its initial value in formulas (14) if the Van der Vusse model under consideration is not corrupted by random and/or preassigned disturbances.

The deterministic model (13) is augmented by the stochastic term $Gw(t)$ in which

$$
G = 0.03\begin{pmatrix} 2.1404 & 0 & 0 & 0\\ 0 & 1.0903 & 0 & 0\\ 0 & 0 & 387.34 & 0\\ 0 & 0 & 0 & 386.06 \end{pmatrix} \tag{15}
$$

and $w(t)$ stands for the zero-mean Gaussian white-noise process with the identity covariance matrix of proper size, i.e. with $Q \equiv I_4$. Following Jørgensen (2007), we assume that the constant parameter c_{A0} is increased by 20% at time $t = 4$ hr. Deterministic and stochastic solutions to this Van der Vusse example are computed in the time interval [0,10 hr].

First of all we determine a reference solution (i.e. true states) to the ODEs (13) augmented with the stochastic term $Gw(t)$, where the matrix G defined in (15), by means of the Euler-Maruyama method with the fixed step size equal to 0.0001 hr. Then, we apply the measurement equation

$$
y_k = \begin{pmatrix} T(t_k)\\ T_J(t_k) \end{pmatrix} + v_k, \tag{16}
$$

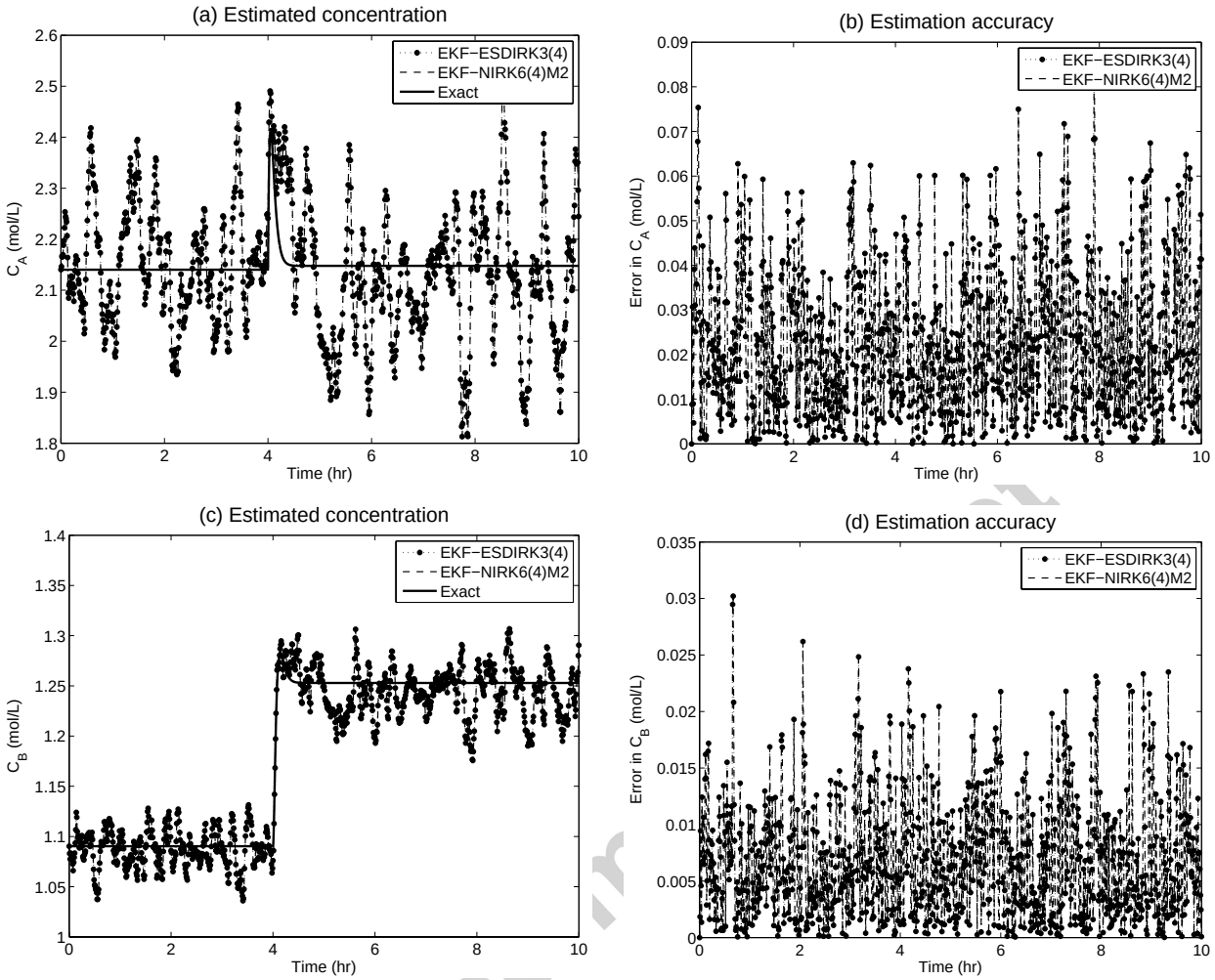

Figure 1: The deterministic (exact) and `EKF-ESDIRK3(4)` and `EKF-NIRK6(4)M2` estimated concentrations (left-hand plots) for the stochastic Van der Vusse model with the small disturbance in the feed concentration of A and the corresponding errors of these filtering techniques (right-hand plots).

where the sampling time points $t_k = 0.01k$ hr, $k = 1, 2, \ldots$, i.e. $\delta = 0.01$ hr in this numerical experiment, and the measurement noise is $v_k \sim \mathcal{N}(0, R_k)$ with the diagonal covariance matrix $R_k = 0.003\,\mathrm{diag}\{387.34, 386.06\}$, to generate the measurement history for the stochastic Van der Vusse reaction at the sampling instants. We point out that the most complicated case when only the temperatures $T(t_k)$ and $T_J(t_k)$ are measured, i.e. exactly as in the measurement model (16), is considered here. A more detailed description of this benchmark example and results of numerical simulation can be found in Jørgensen (2007). In other words, we repeat precisely the numerical experiment elaborated in the cited paper.

Having applied the `EKF-ESDIRK3(4)` and `EKF-NIRK6(4)M2` methods to the stochastic Van der Vusse reaction with the measurement equation (16) and the simulated observation history we arrive at Figure 1(a,c) where the deterministic (marked as Exact) and stochastic concentrations c_A and c_B are displayed. Additionally, Figure 1(b,d) presents the corresponding errors of the used state estimators. We stress that these errors are not the absolute values of the difference

Table 2: The average absolute and relative errors of the `EKF-ESDIRK3(4)` and `EKF-NIRK6(4)M2` estimated concentrations for the stochastic Van der Vusse model with the small disturbance in the feed concentration of A

State vector	Absolute error EKF-ESDIRK3(4)	Relative error EKF-ESDIRK3(4)	Absolute error EKF-NIRK6(4)M2	Relative error EKF-NIRK6(4)M2
c_A	0.0232	1.11%	0.0232	1.11%
c_B	0.0068	0.58%	0.0068	0.58%

between the deterministic and estimated concentrations, but the magnitudes of the difference of the reference solution (i.e. true states) computed to the stochastic Van der Vusse reaction by the Euler-Maruyama method and the concentrations estimated by the `EKF-ESDIRK3(4)` (or `EKF-NIRK6(4)M2`) method at the sampling time points.

Figure 1 displays that both state estimation algorithms reconstruct the state of the stochastic Van der Vusse reaction with no offset observed in the estimated concentrations and with small errors. The average absolute and relative errors of the `EKF-ESDIRK3(4)` and `EKF-NIRK6(4)M2` estimated concentrations c_A and c_B are shown in Table 2. All these data confirm that the temperature measurements are indeed sufficient for the accurate concentration estimation in the Van der Vusse example at least. Furthermore, we even did not augment this chemical model with additional integrators in the feed concentration of A (c_{A0}) and the feed temperature (T_0) as those recommended in Section IIIA in Jørgensen (2007) to calculate the offset free concentrations c_A and c_B in the case of full state feedback, there. So, the numerical results reported by Jørgensen (2007) seem to be an implication of mistakes in coding or, maybe, of an inadequate use of the mathematical theory. For instance, the cited paper utilizes the square-root version of the `ESDIRK3(4)` based EKF method described in Jørgensen et al. (2007) at large. On the other hand, we have already explained in Section 2.2 that the EKF implementation published in Jørgensen et al. (2007) does not ensure the positive semi-definiteness of the predicted error covariance matrix and, hence, its square root may nonexist. The latter might also affect the state estimation quality and lead to the wrong conclusion made for the Van der Vusse benchmark example in Jørgensen (2007).

Another possible reason of such a disappointing result could be an incorrect generation of the true states and measurements. For example, the "true" stochastic and deterministic concentrations of the product B are displayed to be around 1.5 after the disturbance in the cited paper (see Figure 2 in Jørgensen (2007)). However, our calculations say that they should be about 1.25 (see Figure 1, here).

Besides, Figure 1 and Table 2 state that the methods `EKF-ESDIRK3(4)` and `EKF-NIRK6(4)M2` work equally well for the chemical system under examination. So, in order to range them, we boost the difficulty of this test problem. First, we enlarge its time interval to [0,100 hr] and the sampling period to $\delta = 2$ hr, i.e. now the sampling instants are $t_k = 2k$ hr, $k = 1, 2, \ldots$. Second, we increase the disturbance, i.e. it is assumed that the constant parameter c_{A0} is increased by 100% at time $t = 50$ hr. A reference solution (i.e. true states) and the corresponding temperature measurements are calculated precisely as explained above. The outcome of this numerical experiment is exhibited in Figures 2, 3 and in Table 3.

Having analyzed the mentioned experimental data we arrive at the following conclusion. First, the state estimators under consideration are able to reconstruct the offset free state estimates to this more complicated version of the Van der Vusse reaction on the basis of the temperature feedback, only. Second, the `EKF-NIRK6(4)M2` produces the concentration and temperature estimates that are closer to the corresponding deterministic solution (marked as Exact) in Figures 2(a,c) and

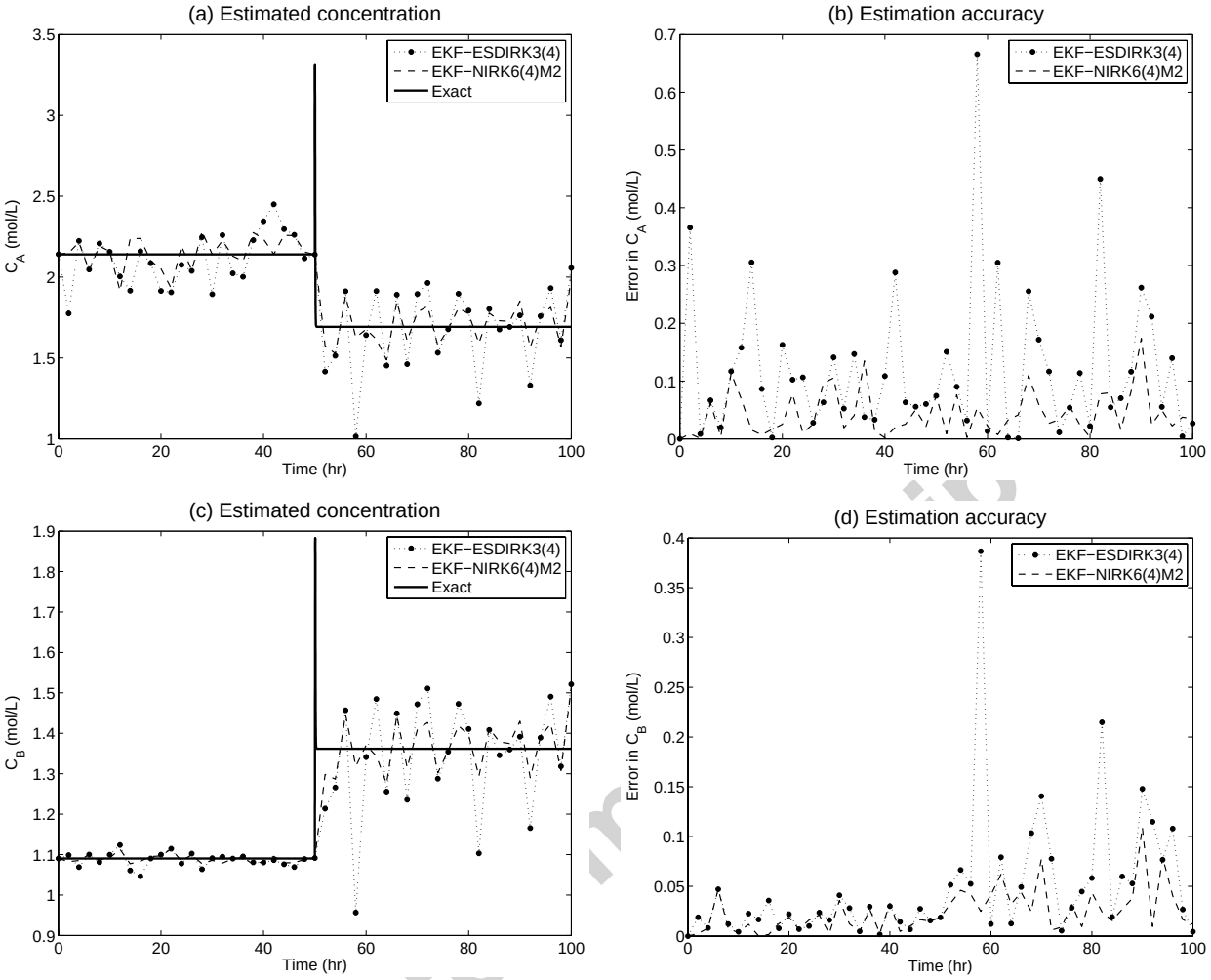

Figure 2: The deterministic (exact) and `EKF-ESDIRK3(4)` and `EKF-NIRK6(4)M2` estimated concentrations (left-hand plots) for the stochastic Van der Vusse model with the increased disturbance in the feed concentration of A and the corresponding errors of these filtering techniques (right-hand plots).

Table 3: The average absolute and relative errors of the `EKF-ESDIRK3(4)` and `EKF-NIRK6(4)M2` estimated concentrations and temperatures for the stochastic Van der Vusse model with the increased disturbance in the feed concentration of A

State vector	Absolute error EKF-ESDIRK3(4)	Relative error EKF-ESDIRK3(4)	Absolute error EKF-NIRK6(4)M2	Relative error EKF-NIRK6(4)M2
c_A	0.1186	6.34%	0.0430	2.27%
c_B	0.0483	3.73%	0.0243	1.91%
T	4.1854	1.06%	0.7297	0.18%
T_J	2.5636	0.65%	0.5653	0.14%

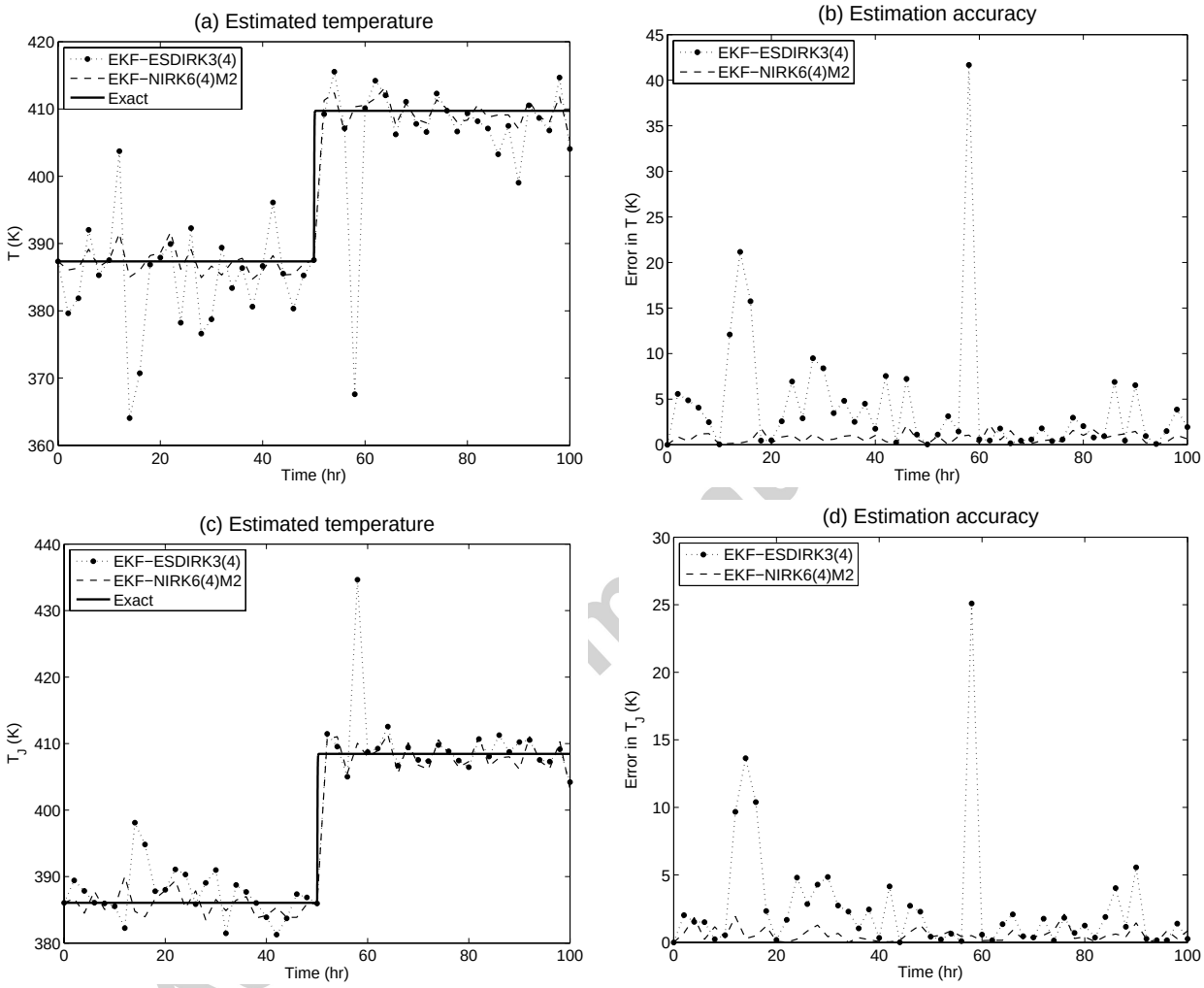

Figure 3: The deterministic (exact) and `EKF-ESDIRK3(4)` and `EKF-NIRK6(4)M2` estimated temperatures (left-hand plots) for the stochastic Van der Vusse model with the increased disturbance in the feed concentration of A and the corresponding errors of these filtering techniques (right-hand plots).

3(a,c). Moreover, the errors of the new state estimator designed in this paper are smaller than the corresponding errors of the EKF variant presented by Jørgensen et al. (2007) (see Figures 2(b,d), 3(b,d) and Table 3). All this confirms that the `EKF-NIRK6(4)M2` method outperforms the earlier published `EKF-ESDIRK3(4)` technique in the quality of state estimation in the stochastic Van der Vusse reaction scenario. Hence, it can be potentially effective in treating other chemical systems.

Our last numerical experiment serves for showing that the state estimator `EKF-NIRK6(4)M2` works successfully even in a situation when the `EKF-ESDIRK3(4)` may fail. For that, we construct an artificial example by taking the most difficult ODE test utilized in Kulikov and Weiner (2011) and adding a noise. The resulting SDE has the form

$$d\begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} \lambda\big(x_2^2(t) - x_1(t)\big) + 2x_1(t)/x_2(t) \\ x_1(t) - x_2^2(t) + 1 \\ -50\big(x_2(t) - 2\big)x_3(t) \end{pmatrix} dt + \begin{pmatrix} 0.01 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} dw(t) \tag{17}$$

where the constant parameter $\lambda = 100$. We remark that this λ represents the stiffness level of the underlying ODEs. The process noise is chosen to be $w(t) \sim \mathcal{N}(0, I_3)$ where, as customary, I_3 means the identity matrix of dimension 3. The SDE (17) is solved in the interval $[0, 2]$ and with the initial values

$$\bar{x}_0 = \begin{pmatrix} 1 \\ 1 \\ \exp(-25) \end{pmatrix} \text{ and } \Pi_0 = \begin{pmatrix} 0.01 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{18}$$

The measurements are discrete and assumed to satisfy the equation

$$y_k = x_2(t_k) + v_k \tag{19}$$

where the measurement noise $v_k \sim \mathcal{N}(0, 0.04)$. Again, we explore the partial measurement model where only the second entry of the state vector is measured. Here, we also examine two different values of the sampling interval δ, namely, $\delta = 0.1$ and $\delta = 0.25$. We intend to show that the accuracy of state estimation by the `EKF-ESDIRK3(4)` method may deteriorate essentially for the longer waiting time, whereas the new state estimation method `EKF-NIRK6(4)M2` is robust and works equally well for both δ's and all noise samples.

As above, we apply first the Euler-Maruyama method, but with the smaller fixed step size equal to 0.00002, to compute a reference solution (i.e. true states) to the SDE (17) with the initial values set by the first formula in (18). Then, the observation history is calculated at sampling points by formula (19) with use of the computed reference solution. After that, we solve the reverse problem, i.e. we treat numerically the continuous-discrete system (17), (19) by means of the `EKF-ESDIRK3(4)` and `EKF-NIRK6(4)M2` methods with the simulated measurements. The absolute values of the output errors committed by the state estimators under examination are exhibited in Figure 4. The top diagrams present numerical results for $\delta = 0.1$ and two typical behaviors of the `EKF-ESDIRK3(4)` (i.e. the left-hand plot corresponds to the accurately estimated states, and the right-hand plot shows the inaccurate state estimation). The bottom graphs exhibit the errors for the larger $\delta = 0.25$ and also two estimation regimes of the `EKF-ESDIRK3(4)`. We emphasize that the magnitude of the error in the most difficult entry $x_3(t)$ of the solution vector is taken as the output. Furthermore, all plots of Figure 4 are scaled logarithmically because of the size of the errors committed by the `EKF-ESDIRK3(4)`.

We observe that the tested methods work differently. For both values of the sampling interval δ and all noise samples, the new `EKF-NIRK6(4)M2` always outperforms the earlier published

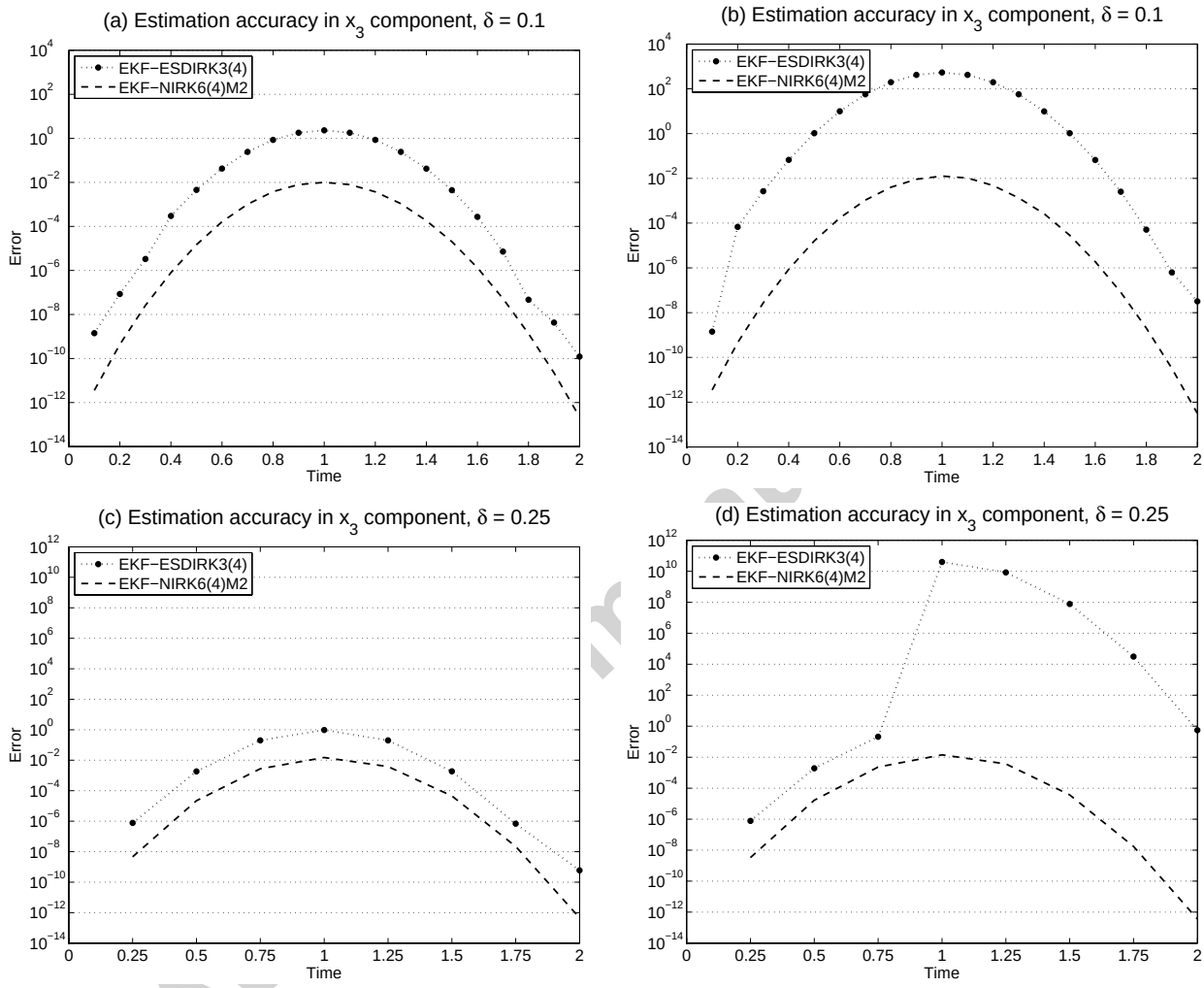

Figure 4: Errors in the entry $x_3(t)$ of the state vector to the SDE (17) committed by the `EKF-ESDIRK3(4)` and `EKF-NIRK6(4)M2` methods for $\delta = 0.1$ and different noise samples (the top plots) and for $\delta = 0.25$ and different noise samples (the bottom plots).

EKF-ESDIRK3(4) technique in the accuracy of state estimation. Moreover, our method is sensitive to neither the particular noise sample nor the length of the waiting time considered in this numerical experiment, whereas the EKF-ESDIRK3(4) can exhibit a reasonable accuracy for some noise samples (see the left-hand plots in Figure 4) or an unacceptable accuracy for others (see the right-hand plots in Figure 4). The latter state estimation algorithm is also sensitive to the duration of the sampling period and, hence, can be hardly recommended for practical use. All this confirms the superiority of the EKF-NIRK6(4)M2 technique and creates the necessary theoretical background for its further implementation in offline chemical models and industrial environment, including real experimentation.

4. Conclusion

In this paper, we have designed and examined the new state estimation technology for chemical engineering. Furthermore, it can easily be extended to other areas of study where continuous-discrete stochastic state-space systems (1), (2) may arise. We recall that the practical state evaluation is often based on available measurements of some parameters (depending of the utilized technology) coupled with computation of remaining (not measurable) variables by means of appropriate software implementing this or that filtering method. Here, we have presented the state estimator grounded in the well-known extended Kalman filtering implemented with use of the efficient NIRK6(4)M2 MDE solver with the local and global error controls. This solver possesses many important properties for applied calculations (see Section 3.1), which create a solid basis for its successful implementation in practice, including industrial applications. The designed EKF-NIRK6(4)M2 method has been compared theoretically and numerically with the EKF method published earlier by Jørgensen et al. (2007) and also grounded in an adaptive MDE solver but with only local error control. Undoubtedly, the EKF implementation constructed here is superior in the sense of theoretical properties and practical efficiency (see Section 3). Our numerical experiments in Section 3.2 exhibit that the EKF-NIRK6(4)M2 method outperforms the other EKF-ESDIRK3(4) method in the stochastic Van der Vusse reaction scenario (see the first test example). Furthermore, as the second test in Section 3.2 says, the new state estimation technique works successfully even in the situation when the EKF-ESDIRK3(4) fails.

Additionally, we have shown that some situations attributed as "EKF failure" can be resolved with use of the new state estimation method. Thus, in contrast to Jørgensen (2007), our numerical experiments confirm that the temperature feedback is sufficient for the accurate concentration estimation in the Van der Vusse reaction (see the first test example in Section 3.2). Eventually, the state estimator EKF-NIRK6(4)M2 presented in Appendix A in detail creates a solid background for developing efficient software sensors for offline models and industrial implementation in chemical and other engineering. Its comparison to other known effective nonlinear state estimation algorithms, as, for example, particle, unscented, cubature and ensemble filters, and also its practical testing in real experimentation are expected in future.

Appendix A: `NIRK6(4)M2` **based array square-root accurate continuous-discrete extended Kalman filter**

Start: Determine the square roots $\Pi_0^{1/2}$ and $R^{1/2}$. Set $P_{0|0}^{1/2} := \Pi_0^{1/2}$, $\hat{x}_{0|0} := \bar{x}_0$ and $\epsilon_g := 10^{-4}$.
Loop body: For $k := 1, 2, \ldots, K$ (where K is the number of sampling points in the integration interval of SDE (1)), do:

- *Time updates.* Predict $\hat{x}_{k|k-1}$ and $P_{k|k-1}^{1/2}$ in the k-th sampling interval (i.e. in $[t_{k-1}, t_k]$) as follows:

  1. Set $\hat{x}_0 := \hat{x}_{k-1|k-1}$ and $P_0^{1/2} := P_{k-1|k-1}^{1/2}$.
  2. Apply the adaptive `NIRK6(4)M2` method to solve numerically the MDEs (3) with the set initial values in the interval $[t_{k-1}, t_k]$ by the following algorithm:

0. Set $\epsilon_{loc} := \epsilon_g^{5/4}$, $\tau_0 := \min\{0.01, \delta\}$, $t_0 := t_{k-1}$, $\tau_{max} := 0.1$, $M := 1$;
1. While $M = 1$ do;
2. $l := 0$, $M := 0$, $\hat{x}_0 := \hat{x}_{k-1|k-1}$, $\Delta\hat{x}_0 := 0$, $P_0^{1/2} := P_{k-1|k-1}^{1/2}$;
3. While $(t_l < t_k)$ & $(\|\Delta\hat{x}_l\|_\infty \le 10\epsilon_g)$ do;
4. $t_{l+1} := t_l + \tau_l$, compute $\hat{x}_{l1}^{3,4}$, $\hat{x}_{l2}^{3,4}$, $\hat{x}_{l3}^{3,4}$, $\hat{x}_{l+1}^{4}$ by the iteration (7);
5. Evaluate the local error le_{l+1} by formula (11);
6. Calculate the optimum step size $\tau_l^* := \min\left\{1.5, 0.8\left(\epsilon_{loc}/\|le_{l+1}\|_\infty\right)^{1/5}\right\}\tau_l$;
7. If $\|le_{k+1}\|_\infty > \epsilon_{loc}$,
then $\tau_l := \tau_l^*$;
else do;
8. Compute the global error $\Delta\hat{x}_{l+1}$ by formula (12);
9. If $\|\Delta\hat{x}_{l+1}\|_\infty > \epsilon_g$,
then $M := 1$;
end{If};
10. If $M = 0$,
then do;
11.* Apply the Cholesky decomposition
$Q(t_{l+1/2}) = Q^{T/2}(t_{l+1/2})Q^{1/2}(t_{l+1/2})$;
11. Calculate the square root $P_{l+1}^{1/2}$ by formulas (9), (10);
end{then};
end{If};
12. $\tau_{l+1} := \min\{\tau_l^*, t_k - t_{l+1}, \tau_{max}\}$;
13. $l := l + 1$;
end{else};
end{If};
end{while};
14. If $M = 1$,
then $\epsilon_{loc} := \left(0.8\epsilon_g / \max_l \|\Delta\hat{x}_l\|_\infty\right)^{5/4}\epsilon_{loc}$;
end{If};
end{while};
15. Stop.

3. The numerical solutions $\hat{x}_L^4$ and $P_L^{1/2}$, where the subscript L marks the last node in the generated mesh (5) (i.e. $t_L \equiv t_k$), are taken as the output of the `NIRK6(4)M2` MDE solver, i.e. the predicted state expectation and the predicted error covariance square root are set to be: $\hat{x}_{k|k-1} := \hat{x}_L^4$, $P_{k|k-1}^{1/2} := P_L^{1/2}$.

- *Measurement updates.* Having computed $\hat{x}_{k|k-1}$ and $P_{k|k-1}^{1/2}$, determine the filtered state $\hat{x}_{k|k}$ and the filtered error covariance square root $P_{k|k}^{1/2}$ as follows:

  1. Evaluate the Jacobian $H_k = dh(\hat{x}_{k|k-1})/dx_k$ at the predicted state $\hat{x}_{k|k-1}$.
  2. Introduce the notation $\bar{K}_{f,k} := P_{k|k-1} H_k^T R_{e,k}^{-1/2}$ and apply the following transformation:

$$\Theta_k \begin{pmatrix} R^{1/2} & 0 \\ P_{k|k-1}^{1/2} H_k^T & P_{k|k-1}^{1/2} \end{pmatrix} = \begin{pmatrix} R_{e,k}^{1/2} & \bar{K}_{f,k}^T \\ 0 & P_{k|k}^{1/2} \end{pmatrix}$$

  where Θ_k is any orthogonal rotation that upper-triangularizes the left-hand matrix of the latter formula. The requested square root $P_{k|k}^{1/2}$ is derived as a diagonal block of the right-hand matrix.
  3. Then, the filtered state $\hat{x}_{k|k}$ is calculated by

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + \bar{K}_{f,k} R_{e,k}^{-T/2} e_k$$

  where the innovation e_k is determined by the second formula in (4b).

Here is some recommendation on practical use of the presented `EKF-NIRK6(4)M2` algorithm.

First, item 2 in the *Time updates* intends for computing the predicted state expectation $\hat{x}_{k|k-1}$ and the predicted square root $P_{k|k-1}^{1/2}$ at the next sampling time point t_k with the error not exceeding the value of the user-supplied accuracy parameter ϵ_g determined in the **Start**. This is the only parameter that has to be set by the user. Its default value is 10^{-4} in this paper. However, one can alter ϵ_g depending on the requested accuracy of state estimation, with understanding that smaller values of this parameter may demand more computational work and, hence, extend the execution time of the algorithm. It is also worthwhile to remark that all the errors are measured in the sup-norm, which is defined for an arbitrary n-dimensional vector x as follows: $\|x\|_\infty = \max_{i=1,2,\ldots,n} |x_i|$. The user may also change it to any other norm appropriate to this or that situation.

Second, step *11** in item 2 of the *Time updates* is optional and may be absent in the situation when the covariance matrix Q is constant or possesses such a trivial structure that its square root is known in advance. In all such cases the square root $Q^{1/2}$ is determined and set in the beginning of the above filtering procedure, i.e. in the **Start**. All numerical examples considered in the paper satisfy this condition.

Acknowledgements

The authors thank the support from Portuguese National Funds through the *Fundação para a Ciência e a Tecnologia* (FCT) within the scope of projects PEst-OE/MAT/UI0822/2011 and SFRH/BPD/64397/2009 and within the *Investigador FCT 2013* programme. The authors are also grateful to anonymous referees for their valuable remarks and comments on the paper.

Arasaratnam, I., Haykin, S., 2008. Square-root quadrature Kalman filtering. IEEE Trans. on Signal Processing 56 (6), 2589–2593.

Arasaratnam, I., Haykin, S., 2009. Cubature Kalman filters. IEEE Trans. Automat. Contr. 54 (6), 1254–1269.

Arasaratnam, I., Haykin, S., Hurd, T. R., 2010. Cubature Kalman filtering for continuous-discrete systems: Theory and simulations. IEEE Trans. on Signal Processing 58 (10), 4977–4993.

Bierman, G. J., 1977. Factorization Methods for Discrete Sequential Estimation. Academic Press, New York.

Bierman, G. J., Belzer, M. R., Vandergraft, J. S., Porter, D. W., Dec. 1990. Maximum likelihood estimation using square root information filters. IEEE Trans. Automat. Contr. 35, 1293–1298.

Butcher, J. C., 2008. Numerical Methods for Ordinary Differential Equations. John Wiley and Sons, Chichester.

Chen, H., Kremling, A., Allgöwer, F., 1995. Nonlinear predictive control of a benchmark CSTR. In: Proceedings of the 3rd European Control Conference ECC'95. pp. 3247–3252.

Dochain, D., 2003. State and parameter estimation in chemical and biochemical processes: a tutorial. J. Process Control 13, 801–818.

Engell, S., Klatt, K.-U., 1993. Nonlinear control of a nonminimum phase CSTR. In: Proceedings of the American Control Conference. Vol. 2941. pp. 2941–2945.

Frogerais, P., Bellanger, J.-J., Senhadji, L., 2012. Various ways to compute the continuous-discrete extended Kalman filter. IEEE Trans. Automat. Contr. 57 (4), 1000–1004.

Goodwin, G. C., Sin, K. S., 1984. Adaptive Filtering Prediction and Control. Prentice-Hall, Englewood Cliffs, New Jersey.

Grewal, M. S., Andrews, A. P., 2001. Kalman Filtering: Theory and Practice. Prentice Hall, New Jersey.

Hairer, E., Nørsett, S. P., Wanner, G., 1993. Solving Ordinary Differential Equations I: Nonstiff Problems. Springer-Verlag, Berlin.

Hairer, E., Wanner, G., 2010. Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems. Springer-Verlag, Berlin.

Haseltine, E. L., Rawlings, J. B., 2002. A critical evaluation of extended kalman filtering and moving-horizon estimation. Tech. rep., Madison, WI.

Haseltine, E. L., Rawlings, J. B., 2005. Critical evaluation of extended Kalman filtering and moving-horizon estimation. Ind. Eng. Chem. Res. 44, 2451–2460.

Ito, K., Xiong, K., 2000. Gaussian filters for nonlinear filtering problems. IEEE Trans. Automat. Contr. 45 (5), 910–927.

Jang, S.-S., Joseph, B., Mukai, H., 1986. Comparison of two approaches to on-line parameter and state estimation of nonlinear systems. Ind. Eng. Chem. Process Des. Dev. 25, 809–814.

Jazwinski, A. H., 1970. Stochastic Processes and Filtering Theory. Academic Press, New York.

Jørgensen, J. B., 2007. A critical discussion of the continuous-discrete extended Kalman filter. In: European Congress of Chemical Engineering - 6, Copenhagen, Denmark. (Available at http://www.nt.ntnu.no/users/skoge/prost/proceedings/ecce6_sep07/upload/3520.pdf).

Jørgensen, J. B., Thomsen, P. G., Madsen, H., Kristensen, M. R., 2007. A computationally efficient and robust implementation of the continuous-discrete extended Kalman filter. In: Proceedings of the American Control Conference. pp. 3706–3712.

Julier, S., Uhlmann, J., 2004. Unscented filtering and nonlinear estimation. Proceedings of the IEEE 92 (3), 401–422.

Julier, S., Uhlmann, J., Durrant-Whyte, H., 2000. A new method for the nonlinear transformation of means and covariances in filters and estimators. IEEE Trans. Automat. Contr. 45 (3), 477–482.

Kailath, T., Sayed, A. H., Hassibi, B., 2000. Linear Estimation. Prentice Hall, New Jersey.

Kalman, R. E., 1960. A new approach to linear filtering and prediction problem. ASME J. Basic Eng. 82 (1), 35–45.

Kristensen, M. R., Jørgensen, J. B., Thomsen, P. G., Jørgensen, S. B., 2004. An ESDIRK method with sensitivity analysis capabilities. Comput. Chem. Eng. 28, 2695–2707.

Kulikov, G. Yu., 2009. Automatic error control in the Gauss-type nested implicit Runge-Kutta formula of order 6. Russian J. Numer. Anal. Math. Modelling 24 (2), 123–144.

Kulikov, G. Yu., 2012. Global error control in adaptive Nordsieck methods. SIAM J. Sci. Comput. 34 (2), A839–A860.

Kulikov, G. Yu., 2013. Cheap global error estimation in some Runge-Kutta pairs. IMA J. Numer. Anal. 33 (1), 136–163.

Kulikov, G. Yu., Kulikova, M. V., 2014. Accurate numerical implementation of the continuous-discrete extended Kalman filter. IEEE Trans. Automat. Contr. 59 (1), 273–279.

Kulikov, G. Yu., Kulikova, M. V., 2015. The accurate continuous-discrete extended Kalman filter for continuous-time stochastic systems. Russian J. Numer. Anal. Math. Modelling 30 (6), (in press).

Kulikov, G. Yu., Merkulov, A. I., Shindin, S. K., 2007. Asymptotic error estimate for general Newton-type methods

and its application to differential equations. Russian J. Numer. Anal. Math. Modelling 22 (6), 567–590.
Kulikov, G. Yu., Shindin, S. K., 2009. Adaptive nested implicit Runge-Kutta formulas of Gauss type. Appl. Numer. Math. 59 (3-4), 707–722.
Kulikov, G. Yu., Weiner, R., 2011. Global error estimation and control in linearly-implicit parallel two-step peer W-methods. J. Comput. Appl. Math. 236 (6), 1226–1239.
Kulikova, M. V., 2009. Likelihood gradient evaluation using square-root covariance filters. IEEE Trans. Automat. Contr. 54 (3), 646–651.
Kulikova, M. V., Pacheco, A., 2013. Kalman filter sensitivity evaluation with orthogonal and J-orthogonal transformations. IEEE Trans. Automat. Contr. 58 (7), 1798–1804.
Lewis, F. L., 1986. Optimal Estimation: with an Introduction to Stochastic Control Theory. John Wiley & Sons, New York.
Maybeck, P. S., 1982. Stochastic Models, Estimation and Control. Academic Press, London.
Mazzoni, T., 2007. Computational aspects of continuous-discrete extended Kalman filtering. Comput. Statist. 23 (4), 519–539.
Nørgaard, M., Poulsen, N. K., Ravn, O., 2000. New developments in state estimation for nonlinear systems. Automatica 36, 1627–1638.
Ogunnaike, B. A., Ray, W. H., 1994. Process Dynamics, Modeling, and Control. Oxford University Press, New York.
Rao, C. V., Rawlings, J. B., Lee, J. H., 2001. Constrained linear state estimation — a moving-horizon approach. Automatica 37, 1619–1628.
Rawlings, J. B., Bakshi, B. R., 2006. Particle filtering and moving horizon estimation. Comput. Chem. Eng. 30, 1529–1541.
Romanenko, A., Castro, J. A. A. M., 2004. The unscented filter as an alternative to the EKF for nonlinear state estimation: a simulation case study. Comput. Chem. Eng. 28 (3), 347–355.
Romanenko, A., Santos, L. O., Afonso, P. A. F. N. A., 2004. Unscented Kalman filtering of a simulated pH system. Ind. Eng. Chem. Res. 43, 7531–7538.
Schneider, R., Georgakis, C., 2013. How to not make the extended Kalman filter fail. Ind. Eng. Chem. Res. 52, 3354–3362.
Simon, D., 2006. Optimal State Estimation: Kalman, H Infinity and Nonlinear Approaches. Wiley, Hoboken, New Jersey.
Singer, H., 2002. Parameter estimation of nonlinear stochastic differential equations: Simulated maximum likelihood versus extended Kalman filter and Ito-Taylor expansion. Journal of Computational and Graphical Statistics 11 (4), 972–995.
Singer, H., 2006. Continuous-Discrete Unscented Kalman Filtering. FernUniversitat, Tech. Rep. 384, Hagen, Germany.
Soroush, M., 1998. State and parameter estimation and their applications in process control. Comput. Chem. Eng. 23, 229–245.
Tsyganova, J. V., Kulikova, M. V., 2013. State sensitivity evaluation within UD based array covariance filter. IEEE Trans. Automat. Contr. 58 (11), 2944–2950.
Wilson, D. I., Agarwal, M., Rippin, D. W. T., 1998. Experiences inplementing the extended Kalman filter on an industrial batch reactor. Comput. Chem. Eng. 22, 1653–1672.