Nonlinear State and Parameter Estimation using Discrete-Time Double Kalman Filter

Mohammad Abdollahpouri [*] **Mark Haring** [**]
Tor Arne Johansen [**] **Gergely Takács** [*] **Boris Roháľ-Ilkiv** [*]

[*] *Institute of Automation, Measurement and Applied Informatics,*
Faculty of Mechanical Engineering, Slovak University of Technology in
Bratislava, 812 31 Bratislava, Slovakia
(e-mail: mohammad.abdollahpouri@stuba.sk),
[**] *Center for Autonomous Marine Operations and Systems*
(NTNU-AMOS), Department of Engineering Cybernetics, Norwegian
University of Science and Technology, 7491 Trondheim, Norway.

Abstract: Dealing with nonlinear dynamics in conventional estimation methods like the extended Kalman filter (EKF) is challenging, since they are not guaranteed to have global convergence, and their instability can arise by selecting a poor initial guess. Recently, a double Kalman filter (DKF) has been proposed, where two stages of estimation are considered using cascade stability theory in the continuous time domain. The first stage guarantees global convergence through the use of a globally valid linear time-varying model transformation, but leads to sub-optimal accuracy in the presence of noise. The global model transformation is applicable to a class of nonlinear systems, where its state can be explicitly derived through a mapping of previous measurements and disturbances. Furthermore, the second stage compensates the lost performance using the estimate from the first stage via local linearization. Here, we derive the stability analysis of this globally convergent method in discrete time using a Lyapunov approach. Different Kalman filters are compared via simulation to validate the benefit of using DKF for nonlinear state and parameter estimation.

1. INTRODUCTION

The Kalman filter (KF) as an optimal (minimum variance) state estimation method is proven to be globally exponentially stable for uniformly completely observable (UCO) linear time varying (LTV) systems with white input and output disturbances (Jazwinski, 2007). The extended Kalman filter (EKF) is a modified KF for nonlinear dynamics which has enormously influenced the state estimation of real life applications (Gelb, 1974; Simon, 2010). However, this widely used estimator lacks a global stability guarantee; state and parameter estimates may diverge because of a poor initial guess or its dependency on system trajectory. This drawback is argued to initiate from the inherent feedback in its linearization, where the estimate from a poor initial guess is utilized as the linearization point. Divergence of the EKF has been reported by Perea

et al. (2007); Haseltine and Rawlings (2005); Lee and West (2010); Ficocelli and Janabi (2001).

A two stage state estimation approach has been recently developed by Johansen and Fossen (2016a). One of its variants, the double Kalman filter (DKF), has been analyzed in the continuous-time domain (Johansen and Fossen, 2016b) and further has been applied to a position estimation using pseudo-range measurements (Johansen et al., 2016). The first stage of DKF employs a technique which eliminates the nonlinearities of the system using a transformation that results in a LTV system. This transformation uses the measurements and possibly its derivatives without optimally considering the input and output disturbances. Suboptimal modeling of disturbances degrades the performance of estimation methods using this LTV model. The second stage of the DKF, a linearized Kalman filter (LKF), utilizes the estimates from the first stage as its linearization point.

The use of differential flatness theory is one approach to transform a continuous-time nonlinear system into a LTV system in the first stage of the DKF. This theory was first introduced by Fliess et al. (1995) and further developed by others in the field of control and state estimation (Rigatos, 2015). Based on Fliess et al. (1995), *flat systems* are among the class of nonlinear systems that are equivalent to a static system; using a *flat output* it is possible to explicitly express the state in terms of the flat output and a number of its derivatives. *Difference flatness* is the counterpart concept of differentially flatness in the discrete-time do-

main (Sira-Ramirez and Agrawal, 2004). Under the same topic, an immersion method with or without feedback is proposed by Levine and Marino (1986) to transform a class of nonlinear systems to linear systems. Necessary and sufficient conditions for linearizing transformations up to output injection are discussed by Besançon (1999); Hammouri and Gauthier (1992). Furthermore, conditions for a coordinate transformation to observer canonical form are presented by Califano et al. (2009).

This paper first discusses the transformation of a nonlinear system into a LTV system in Section 2. Next, Section 3 introduces the equations of the DKF in discrete time. A stability analysis of the DKF in discrete time is presented in Section 4. The stability analysis is derived utilizing a Lyapunov approach, where it is proven that the DKF is uniformly globally asymptotically stable using stability theory for discrete-time cascaded systems. Furthermore, in Section 5 different numerical examples validate the advantage of the DKF over the EKF that uses its last estimate as linearization point. Finally, this paper is concluded in Section 6.

Finally we define some useful definitions. A function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is in class of $\mathcal{K}$, if it is continuous, strictly increasing and $\alpha(0) = 0$. Furthermore, $\alpha \in \mathcal{K}_\infty$ if it is unbounded. A function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is of class $\mathcal{KL}$ if $\beta(\cdot, k) \in \mathcal{K}$ for all $k \geq 0$, and if $\beta(x, \cdot)$ is continuous, strictly decreasing and $\lim_{k \to 0} \beta(x, k) = 0$ for all $x > 0$. Diagonal matrix is denoted by *diag*.

2. MODEL TRANSFORMATION

This section provides the definitions and assumptions needed for applying a linearizing model transformation based on difference flatness. Let the nonlinear dynamics be described as follows

$$x_{k+1} = f(x_k, \acute{w}_k), \tag{1a}$$
$$y_k = h(x_k, \acute{v}_k), \tag{1b}$$

where the nonlinear dynamics is denoted by $f(\cdot) : \mathbb{R}^n \times \mathbb{R}^q \to \mathbb{R}^n$ and the nonlinear output function is expressed as $h(\cdot) : \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^p$. Furthermore, $x \in \mathbb{R}^n$ is the state vector, $y \in \mathbb{R}^p$ is the output, $\acute{w} \in \mathbb{R}^q$ is the input disturbance and $\acute{v} \in \mathbb{R}^r$ denotes the output disturbance. The time index is denoted by k. Although $f(\cdot)$ and $h(\cdot)$ are known functions, the state x_k and the disturbances $\acute{w}_k$ and $\acute{v}_k$ are unknown. The output y_k is obtained via measurement and is therefore known.

Definition 2.1. (Loria and Nešić, 2002) A nonlinear system of the form (1) is uniformly globally asymptotically stable (UGAS) if there exists a function $\beta \in \mathcal{KL}$ such that
$$\|x_k\| \leq \beta(\|x_{k_0}\|, k - k_0)$$
for all $x_{k_0} \in \mathbb{R}^n$ and all $k \geq k_0 \geq 0$.

Definition 2.2. (Loria and Nešić, 2002) A nonlinear system of the form (1) is uniformly globally bounded (UGB) if there exist a function $\kappa \in \mathcal{K}_\infty$ and a constant $c \in \mathbb{R}_{>0}$ such that
$$\|x_k\| \leq \kappa(\|x_{k_0}\|) + c$$
for all $x_{k_0} \in \mathbb{R}^n$ and all $k \geq k_0 \geq 0$.

We make the following assumptions regarding system (1).
Assumption 1. The nonlinear functions $f(\cdot)$ and $h(\cdot)$ are twice continuously differentiable.

Assumption 2. The state x_k is uniformly bounded.
Assumption 3. There exists a known map $\psi(\cdot) : \mathbb{R}^{pd} \times \mathbb{R}^{q(d-1)} \times \mathbb{R}^{rd} \to \mathbb{R}^n$ and a positive integer d such that the state of (1) can be written as
$$x_k = \psi(\mathcal{Y}_k^l, \mathcal{W}_{k-1}^l, \mathcal{V}_k^l),$$
where $\mathcal{Y}_k^l = \{y_l, \ldots, y_k\}$, $\mathcal{W}_{k-1}^l = \{\acute{w}_l, \ldots, \acute{w}_{k-1}\}$ and $\mathcal{V}_k^l = \{\acute{v}_l, \ldots, \acute{v}_k\}$, with $l = k - d + 1$.
Remark 2.3. Assumption 3 implies that system (1) is *difference flat* (Sira-Ramirez and Castro-Linares, 2000), with flat outputs $\mathcal{Y}_k^l$, $\mathcal{W}_{k-1}^l$ and $\mathcal{V}_k^l$. It is worth mentioning the fact that ψ is not necessarily unique. Moreover, it follows from Assumption 3 that the current state x_k can be uniquely determined from a number of current and past output measurements $\mathcal{Y}_k^l$ and disturbances $\mathcal{W}_{k-1}^l$ and $\mathcal{V}_k^l$, which implies observability of system (1) in some sense.

Using Assumption 3, the nonlinear system (1) can be globally transformed into a LTV system of the form
$$x_{k+1} = f(0,0) + F_k x_k + w_k, \tag{2a}$$
$$y_k = h(0,0) + H_k x_k + v_k, \tag{2b}$$
where the matrices F_k and H_k are known, and where the new input and output disturbances, denoted by w_k and v_k respectively, are uniformly zero (i.e., zero for all k) if $\acute{w}_k$ and $\acute{v}_k$ are uniformly zero. The transformation from (1) to (2) is not unique. One possible choice for F_k and H_k is

$$F_k = \int_0^1 \frac{\partial f}{\partial x}(s\psi(\mathcal{Y}_k^l, 0, 0), 0) \mathrm{d}s, \tag{3}$$

$$H_k = \int_0^1 \frac{\partial h}{\partial x}(s\psi(\mathcal{Y}_k^l, 0, 0), 0) \mathrm{d}s. \tag{4}$$

The matrices F_k in (3) and H_k in (4) are known because f, h, ψ and $\mathcal{Y}_k^l$ are known. The corresponding disturbances are given by

$$w_k = \int_0^1 \left(\frac{\partial f}{\partial x}(sx_k, 0) - \frac{\partial f}{\partial x}(s\psi(\mathcal{Y}_k^l, 0, 0), 0) \right) \mathrm{d}s \, x_k + f(x_k, \acute{w}_k) - f(x_k, 0), \tag{5}$$

$$v_k = \int_0^1 \left(\frac{\partial h}{\partial x}(sx_k, 0) - \frac{\partial h}{\partial x}(s\psi(\mathcal{Y}_k^l, 0, 0), 0) \right) \mathrm{d}s \, x_k + h(x_k, \acute{v}_k) - h(x_k, 0), \tag{6}$$

where we note that $\int_0^1 \frac{\partial f}{\partial x}(sx_k, 0)\mathrm{d}s \, x_k = f(x_k, 0) - f(0, 0)$ and $\int_0^1 \frac{\partial h}{\partial x}(sx_k, 0)\mathrm{d}s \, x_k = h(x_k, 0) - h(0, 0)$, which follows from Abraham et al. (2012, Prop. 2.4.7). From Assumption 3, we have that $x_k = \psi(\mathcal{Y}_k^l, 0, 0)$ if $\acute{w}_k$ and $\acute{v}_k$ are uniformly zero. By substituting $x_k = \psi(\mathcal{Y}_k^l, 0, 0)$ in (5) and (6), it is easy to see that w_k and v_k are uniformly zero if $\acute{w}_k$ and $\acute{v}_k$ are uniformly zero.

Example. Consider the nonlinear system
$$x_{k+1,1} = x_{k,1}(1 + \acute{w}_k) + x_{k,2},$$
$$x_{k+1,2} = -x_{k,1}x_{k,2}^2,$$
$$y_k = x_{k,1} + \acute{v}_k.$$
Let $\delta_k := y_k - \acute{v}_k$ such that $x_{k,1} = \delta_k$. From the first system equation, we obtain $x_{k-1,2} = \delta_k - \delta_{k-1}(1 + \acute{w}_{k-1})$. Substituting this in the second system equation yields $x_{k,2} = -\delta_{k-1}(\delta_k - \delta_{k-1}(1 + \acute{w}_{k-1}))^2$. Hence, we can define ψ in Assumption 3 as

$$\psi(\mathcal{Y}_k^{k-1}, \mathcal{W}_{k-1}^{k-1}, \mathcal{V}_k^{k-1}) = \begin{bmatrix} \delta_k \\ -\delta_{k-1}(\delta_k - \delta_{k-1}(1 + \acute{w}_{k-1}))^2 \end{bmatrix}.$$

Using the map ψ, the nonlinear system can be written as a LTV system of the form (2), with $f(0,0) = 0$, $h(0,0) = 0$ and

$$F_k = \begin{bmatrix} 1 & 1 \\ -\frac{1}{3}y_{k-1}^2(y_k - y_{k-1})^4 & \frac{2}{3}y_k y_{k-1}(y_k - y_{k-1})^2 \end{bmatrix},$$
$$H_k = [1 \ 0],$$

as in (3) and (4). We note that this linearizing transformation is not unique because the use of any matrix F_k that satisfies

$$F_k = \begin{bmatrix} 1 & 1 \\ -\zeta y_{k-1}^2(y_k - y_{k-1})^4 & (1-\zeta)y_k y_{k-1}(y_k - y_{k-1})^2 \end{bmatrix}$$

for some $\zeta \in \mathbb{R}$ results in a LTV system for which w_k and v_k are uniformly zero if $\acute{w}_k$ and $\acute{v}_k$ are uniformly zero.

Remark 2.4. The DKF in the next section can be applied to any nonlinear system (1) as long as it can be globally transformed into a LTV system. It is not strictly necessary that this transformation is based on difference flatness. Alternative methods to transform a nonlinear system into a LTV system are mentioned in Section 1.

3. DOUBLE KALMAN FILTERING

There are two stages of Kalman filtering in the DKF; see Fig. 1. The first stage of DKF utilizes the transformed LTV system for sub-optimal estimation and is called an auxiliary Kalman filter (AKF). This Kalman filter is acting as an auxiliary estimation to provide an operating point for the second stage. In the second stage, a linearized Kalman filter (LKF) improves the estimation quality using the results of AKF. This approach has been analyzed for continuous time domain by Johansen and Fossen (2016b). Here, we present the DKF in discrete time.
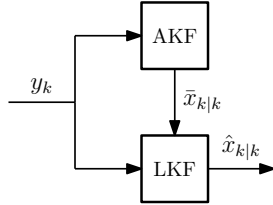

Fig. 1. Schematic of DKF in discrete time

Before we introduce the DKF formulation, let us define the notion of uniform complete observability for LTV systems.

Definition 3.1. A LTV system of the form (2) is uniformly completely observable (UCO) if there exist constants $c_1, c_2 \in \mathbb{R}_{>0}$ and a positive integer N such that

$$c_1 I \preceq \sum_{i=k}^{k+N-1} \Phi^{\mathrm{T}}(i,k)H_i^{\mathrm{T}}H_i\Phi(i,k) \preceq c_2 I$$

for all $k \geq 0$, where the transition matrix is defined as
$$\Phi(i,k) = F_{i-1}F_{i-2}\cdots F_k, \quad \forall i > k, \quad \Phi(k,k) = I.$$

The AKF estimates the state of the LTV system in (2). We assume the following.

Assumption 4. The LTV system in (2) is UCO.

The correction step of the AKF is formulated as

$$\bar{K}_k = \bar{P}_{k|k-1}H_k^{\mathrm{T}}(H_k\bar{P}_{k|k-1}H_k^{\mathrm{T}} + \bar{R}_k)^{-1}, \tag{7a}$$
$$\bar{x}_{k|k} = \bar{x}_{k|k-1} + \bar{K}_k(y_k - h(0,0) - H_k\bar{x}_{k|k-1}), \tag{7b}$$
$$\bar{P}_{k|k} = (I - \bar{K}_kH_k)\bar{P}_{k|k-1}, \tag{7c}$$

with initial conditions $\bar{x}_{0|-1}$ and $\bar{P}_{0|-1} = \bar{P}_{0|-1}^{\mathrm{T}} \succ 0$, where $\bar{R}_k = \bar{R}_k^{\mathrm{T}} \succ 0$ is a tuning matrix and $\bar{K}_k$ is the Kalman gain. The state estimate is propagated to the next time step using the equations

$$\bar{x}_{k+1|k} = f(0,0) + F_k\bar{x}_{k|k}, \tag{8a}$$
$$\bar{P}_{k+1|k} = F_k\bar{P}_{k|k}F_k^{\mathrm{T}} + \bar{Q}_k, \tag{8b}$$

where $\bar{Q}_k = \bar{Q}_k^{\mathrm{T}} \succ 0$ is a tuning matrix. The tuning matrices $\bar{R}_k$ and $\bar{Q}_k$ are chosen such that

$$c_{\bar{R}1}I \preceq \bar{R}_k \preceq c_{\bar{R}2}I,$$
$$c_{\bar{Q}1}I \preceq \bar{Q}_k \preceq c_{\bar{Q}2}I$$

for all $k \geq 0$ and some constants $c_{\bar{R}1}, c_{\bar{R}2}, c_{\bar{Q}1}, c_{\bar{Q}2} \in \mathbb{R}_{>0}$.

Remark 3.2. If the disturbances $\acute{w}_k$ and $\acute{v}_k$ in (1) are stochastic processes, the gain matrices $\bar{Q}_k$ and $\bar{R}_k$ can be chosen to be equal to the covariance matrices of the disturbances w_k and v_k in (2), respectively. However, the transformation from the nonlinear system in (1) to the LTV system in (2) makes it difficult to relate the stochastic properties of $\acute{w}_k$ and $\acute{v}_k$ to the stochastic properties of w_k and v_k. Moreover, the transformation can lead to estimation biases and nonlinear sensitivities due to correlations between the state and the disturbances. Hence, although setting the values of $\bar{Q}_k$ and $\bar{R}_k$ to the covariances of $\acute{w}_k$ and $\acute{v}_k$ often serves as a good starting point for the tuning of $\bar{Q}_k$ and $\bar{R}_k$, calculating the covariances of w_k and v_k may be difficult and additional tuning may be required. This motivates the need for the second stage KF, where the disturbances can usually be accommodated better.

For the second stage of the DKF, we linearize the nonlinear system in (1) about the state estimate $\bar{x}_{k|k}$ of the first stage of the DKF, which gives

$$x_{k+1} = f(\bar{x}_{k|k},0) + A_k(x_k - \bar{x}_{k|k}) + \mathcal{Q}_k + \check{w}_k, \tag{9a}$$
$$y_k = h(\bar{x}_{k|k},0) + C_k(x_k - \bar{x}_{k|k}) + \mathcal{R}_k + \check{v}_k, \tag{9b}$$

with

$$A_k = \frac{\partial f}{\partial x}(\bar{x}_{k|k},0), \qquad C_k = \frac{\partial h}{\partial x}(\bar{x}_{k|k},0),$$

where $\mathcal{Q}_k = [\mathcal{Q}_{1,k},\ldots,\mathcal{Q}_{n,k}]^{\mathrm{T}}$ and $\mathcal{R}_k = [\mathcal{R}_{1,k},\ldots,\mathcal{R}_{p,k}]^{\mathrm{T}}$ are the higher order remainder terms due to linearization and from (Folland, 1990)

$$\mathcal{Q}_{i,k} = (x_k - \bar{x}_{k|k})^{\mathrm{T}}\int_0^1 (1-s) \tag{10}$$
$$\times \frac{\partial^2 f_i}{\partial x \partial x^{\mathrm{T}}}(sx_k + (1-s)\bar{x}_{k|k},0)\mathrm{d}s \ (x_k - \bar{x}_{k|k}),$$

$$\mathcal{R}_{j,k} = (x_k - \bar{x}_{k|k})^{\mathrm{T}}\int_0^1 (1-s) \tag{11}$$
$$\times \frac{\partial^2 h_j}{\partial x \partial x^{\mathrm{T}}}(sx_k + (1-s)\bar{x}_{k|k},0)\mathrm{d}s \ (x_k - \bar{x}_{k|k})$$

for all $i \in \{1,\ldots,n\}$ and $j \in \{1,\ldots,p\}$, where f_i is the ith element of f and h_j is the jth element of h, and where the disturbances $\check{w}_k$ and $\check{v}_k$ are defined as

$$\check{w}_k = f(x_k,\acute{w}_k) - f(x_k,0), \tag{12}$$
$$\check{v}_k = h(x_k,\acute{v}_k) - h(x_k,0). \tag{13}$$

Similar to w_k and v_k in (2), we note that $\check{w}_k$ in (12) and $\check{v}_k$ in (13) are uniformly zero if $\acute{w}_k$ and $\acute{v}_k$ are uniformly zero.

Assumption 5. The LTV system in (9) is UCO.

The LKF formulation about $\bar{x}_{k|k}$ as an operating point is

$$\hat{K}_k = \hat{P}_{k|k-1}C_k^{\mathrm{T}}(C_k\hat{P}_{k|k-1}C_k^{\mathrm{T}} + \hat{R}_k)^{-1}, \tag{14a}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + \hat{K}_k(y_k - h(\bar{x}_{k|k}, 0) - C_k(\hat{x}_{k|k-1} - \bar{x}_{k|k})), \tag{14b}$$

$$\hat{P}_{k|k} = (I - \hat{K}_k C_k)\hat{P}_{k|k-1}, \tag{14c}$$

with initial conditions $\hat{x}_{0|-1}$ and $\hat{P}_{0|-1} = \hat{P}_{0|-1}^{\mathrm{T}} \succ 0$, where $\hat{R}_k = \hat{R}_k^{\mathrm{T}} \succ 0$ is a tuning matrix and $\hat{K}_k$ is the Kalman gain. The propagation for the state estimate is as follows

$$\hat{x}_{k+1|k} = f(\bar{x}_{k|k}, 0) + A_k(\hat{x}_{k|k} - \bar{x}_{k|k}), \tag{15a}$$

$$\hat{P}_{k+1|k} = A_k\hat{P}_{k|k}A_k^{\mathrm{T}} + \hat{Q}_k, \tag{15b}$$

where $\hat{Q}_k = \hat{Q}_k^{\mathrm{T}} \succ 0$ is a tuning matrix. The tuning matrices $\hat{R}_k$ and $\hat{Q}_k$ are selected such that they satisfy

$$c_{\hat{R}1}I \preceq \hat{R}_k \preceq c_{\hat{R}2}I, \tag{16}$$

$$c_{\hat{Q}1}I \preceq \hat{Q}_k \preceq c_{\hat{Q}2}I \tag{17}$$

for all $k \geq 0$ and some constants $c_{\hat{R}1}, c_{\hat{R}2}, c_{\hat{Q}1}, c_{\hat{Q}2} \in \mathbb{R}_{>0}$.

Remark 3.3. The expressions of the disturbances $\breve{w}_k$ in (12) and $\breve{v}_k$ in (13) of system (9) are much simpler than the expressions of w_k and v_k of the LTV system in (2); see for example (5) and (6). Therefore, it is relatively easy to relate the properties of the disturbances $\acute{w}_k$ and $\acute{v}_k$ of the nonlinear system in (1) to the properties of the disturbances $\breve{w}_k$ and $\breve{v}_k$ of the linearized system in (9). For example, if $\acute{w}_k$ and $\acute{v}_k$ are additive, then $\breve{w}_k = \acute{w}_k$ and $\breve{v}_k = \acute{v}_k$. For additive disturbances, the tuning matrices $\hat{Q}_k = \acute{Q}_k$ and $\hat{R}_k = \acute{R}_k$ are suitable choices, assuming that $\acute{w}_k$ and $\acute{v}_k$ are zero-mean white noise processes with covariances $\acute{Q}_k$ and $\acute{R}_k$, respectively. Moreover, for nonadditive disturbances, the tuning matrices $\hat{Q}_k$ and $\hat{R}_k$ may be chosen as $\hat{Q}_k = \left(\frac{\partial f}{\partial \acute{w}}(\bar{x}_{k|k}, 0)\right)^{\mathrm{T}} \acute{Q}_k \frac{\partial f}{\partial \acute{w}}(\bar{x}_{k|k}, 0)$ and $\hat{R}_k = \left(\frac{\partial h}{\partial \acute{v}}(\bar{x}_{k|k}, 0)\right)^{\mathrm{T}} \acute{R}_k \frac{\partial h}{\partial \acute{v}}(\bar{x}_{k|k}, 0)$, similar to Simon (2006), under the additional assumption that $\bar{x}_{k|k}$ is a reasonably good estimate of x_k.

4. NOMINAL STABILITY ANLAYSIS

In this section, we study the stability of the DKF in Fig. 1 under nominal conditions, i.e., $\acute{w}_k = 0$ and $\acute{v}_k = 0$ for all k. We define the following estimation errors of the AKF:

$$\tilde{x}_{k|k-1} = \bar{x}_{k|k-1} - x_k, \qquad \tilde{x}_{k|k} = \bar{x}_{k|k} - x_k. \tag{18}$$

From (2), (7) and (8), it follows that the corresponding difference equations under nominal conditions are

$$\tilde{x}_{k|k} = (I - \bar{K}_k H_k)\tilde{x}_{k|k-1}, \tag{19a}$$

$$\tilde{x}_{k+1|k} = F_k\tilde{x}_{k|k}, \tag{19b}$$

where we note that w_k and v_k are uniformly zero because $\acute{w}_k$ and $\acute{v}_k$ are uniformly zero; see Section 2. The stability of (19) follows directly from the well-known stability analysis for Kalman filters in Jazwinski (2007).

Lemma 4.1. Under Assumptions 1-4, the nominal error system of the AKF in (19) is UGAS.

Proof. See Jazwinski (2007, Theorem 7.4). □

We denote the estimation errors of the LKF by

$$\breve{x}_{k|k-1} = \hat{x}_{k|k-1} - x_k, \qquad \breve{x}_{k|k} = \hat{x}_{k|k} - x_k. \tag{20}$$

The difference equations of the corresponding error system of the LFK under nominal conditions follow from (9), (14) and (15):

$$\breve{x}_{k|k} = (I - \hat{K}_k C_k)\breve{x}_{k|k-1} + \hat{K}_k\mathcal{R}_k, \tag{21a}$$

$$\breve{x}_{k+1|k} = A_k\breve{x}_{k|k} - \mathcal{Q}_k, \tag{21b}$$

where we note that the disturbances $\breve{w}_k$ and $\breve{v}_k$ are uniformly zero because $\acute{w}_k$ and $\acute{v}_k$ are uniformly zero, as shown in Section 3. Remainder terms $\mathcal{Q}_k$ and $\mathcal{R}_k$ depends on the estimation error of the AKF.

Proposition 4.2. Under Assumptions 1-4, there exist constants $\epsilon_q, \epsilon_r \in \mathbb{R}_{>0}$ such that

$$\|\mathcal{Q}_k\| \leq \epsilon_q\|\tilde{x}_{k|k}\|^2, \quad \|\mathcal{R}_k\| \leq \epsilon_r\|\tilde{x}_{k|k}\|^2, \quad \forall k \geq 0.$$

Proof. From (10), (11) and (18), it follows that

$$\mathcal{Q}_{i,k} = \tilde{x}_{k|k}^{\mathrm{T}} \int_0^1 (1-s)\frac{\partial^2 f_i}{\partial x \partial x^{\mathrm{T}}}(x_k + (1-s)\tilde{x}_{k|k}, 0)\mathrm{d}s\, \tilde{x}_{k|k}, \tag{22}$$

$$\mathcal{R}_{j,k} = \tilde{x}_{k|k}^{\mathrm{T}} \int_0^1 (1-s)\frac{\partial^2 h_j}{\partial x \partial x^{\mathrm{T}}}(x_k + (1-s)\tilde{x}_{k|k}, 0)\mathrm{d}s\, \tilde{x}_{k|k} \tag{23}$$

for $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, p\}$. From Lemma 4.1 and Assumption 2, it follows that the state x_k and the error $\tilde{x}_k$ are uniformly bounded. Moreover, Assumption 1 implies that the Hessians of the functions f and h are continuous. Therefore, there exist constants $\epsilon_q, \epsilon_r \in \mathbb{R}_{>0}$ such that

$$\left\|\frac{\partial^2 f_i}{\partial x \partial x^{\mathrm{T}}}(x_k + (1-s)\tilde{x}_{k|k}, 0)\right\| \leq \frac{2\epsilon_q}{\sqrt{n}}, \tag{24}$$

$$\left\|\frac{\partial^2 h_j}{\partial x \partial x^{\mathrm{T}}}(x_k + (1-s)\tilde{x}_{k|k}, 0)\right\| \leq \frac{2\epsilon_r}{\sqrt{p}} \tag{25}$$

for all $k \geq 0$, $i \in \{1, \ldots, n\}$, $j \in \{1, \ldots, p\}$ and $s \in [0,1]$. The proposition follows from (22)-(25). □

The combined error dynamics of the AKF in (19) and the LKF in (21) can be regarded as a cascaded system, where the estimation error of the AKF enters the error system of the LKF via the remainder terms $\mathcal{Q}_k$ and $\mathcal{R}_k$. If the estimation error of the AKF is zero, then the remainder terms are zero (see Proposition 4.2), the error system of the LKF in (21) reduces to

$$\breve{x}_{k|k} = (I - \hat{K}_k C_k)\breve{x}_k, \tag{26a}$$

$$\breve{x}_{k+1|k} = A_k\breve{x}_{k|k} \tag{26b}$$

and we obtain the following result.

Lemma 4.3. Under Assumptions 1-5, the (reduced) nominal error system of the LKF in (26) is UGAS.

Proof. See Jazwinski (2007, Theorem 7.4). □

Since the estimation error of the AKF has converged to zero, the estimation error of the LKF converges to zero. To prove that the cascaded nominal error system in (19) and (21) is UGAS, it is sufficient to show in addition that the state of cascaded nominal error system is UGB; see Loria and Nešić (2002). The following lemma proves that the cascaded nominal error system of the DKF in (19) and (21) is UGB under the assumptions in this work.

Lemma 4.4. Under Assumptions 1-5, the cascaded nominal error system of the DKF in (19) and (21) is UGB.

Proof. Lemma 4.1 directly implies that the error system of the AKF in (19) is UGB because it is UGAS. Therefore,

the estimation error $\tilde{x}_k$ of the AKF is uniformly bounded for any initial condition. To prove that the error system of the LKF in (21) is UGB if $\tilde{x}_k$ is uniformly bounded, we introduce the Lyapunov function candidate

$$V_k = \breve{x}_{k|k-1}^{\mathrm{T}} \hat{P}_{k|k-1}^{-1} \breve{x}_{k|k-1}. \qquad (27)$$

Under Assumption 5, it can be shown with a similar line of reasoning as in the proofs of Jazwinski (2007, Lemmas 7.1-7.2), that there exist constants $c_{\hat{P}1}, c_{\hat{P}2} \in \mathbb{R}_{>0}$ and a positive integer N such that

$$c_{\hat{P}1} I \preceq \hat{P}_{k|k-1} \preceq c_{\hat{P}2} I, \qquad (28)$$

for all $k \geq N$. This implies that V_k can be bounded by

$$\frac{1}{c_{\hat{P}2}} \|\breve{x}_{k|k-1}\|^2 \leq V_k \leq \frac{1}{c_{\hat{P}1}} \|\breve{x}_{k|k-1}\|^2, \qquad (29)$$

for all $k \geq N$. From (14), (15), (21), (27) and the matrix inversion lemma, we obtain

$$V_{k+1} \leq V_k + \mathcal{Q}_k^{\mathrm{T}} \hat{Q}_k^{-1} \mathcal{Q}_k + \mathcal{R}_k^{\mathrm{T}} \hat{R}_k^{-1} \mathcal{R}_k - (C_k \breve{x}_{k|k-1} - \mathcal{R}_k)^{\mathrm{T}}$$
$$\times (C_k \hat{P}_{k|k-1} C_k + \hat{R}_k)^{-1} (C_k \breve{x}_{k|k-1} - \mathcal{R}_k), \qquad (30)$$

similar to the proof of Jazwinski (2007, Theorem 7.4). From (30) and Young's inequality, it follows that

$$V_{k+1} \leq V_k - \frac{1}{2} \breve{x}_{k|k-1}^{\mathrm{T}} C_k^{\mathrm{T}} (C_k \hat{P}_{k|k-1} C_k + \hat{R}_k)^{-1} C_k \breve{x}_{k|k-1}$$
$$+ \mathcal{Q}_k^{\mathrm{T}} \hat{Q}_k^{-1} \mathcal{Q}_k + 2 \mathcal{R}_k^{\mathrm{T}} \hat{R}_k^{-1} \mathcal{R}_k. \qquad (31)$$

From Proposition 4.2 and the bounds on $\hat{R}_k$ in (16) and $\hat{Q}_k$ in (17), we obtain that for any uniformly bounded solution of (19), there exist constants $\rho_Q, \rho_R \in \mathbb{R}_{>0}$ such that $2 \mathcal{R}_k^{\mathrm{T}} \hat{R}_k^{-1} \mathcal{R}_k \leq \rho_R$ and $\mathcal{Q}_k^{\mathrm{T}} \hat{Q}_k^{-1} \mathcal{Q}_k \leq \rho_Q$ for all $k \geq 0$. Moreover, from the boundedness of C_k and the bounds in (16) and (28), it follows that there exists a constant $\rho_C \in \mathbb{R}_{>0}$ such that

$$V_{k+1} \leq V_k - \rho_C \breve{x}_{k|k-1}^{\mathrm{T}} C_k^{\mathrm{T}} C_k \breve{x}_{k|k-1} + \rho_Q + \rho_R \qquad (32)$$

for all $k \geq 0$. By recursively using (32), we get

$$V_{k+N} \leq V_k - \rho_C \sum_{i=k}^{k+N-1} \breve{x}_{i|i-1}^{\mathrm{T}} C_i^{\mathrm{T}} C_i \breve{x}_{i|i-1} + N\rho_Q + N\rho_R \qquad (33)$$

for any positive integer N. From (21), we have

$$\breve{x}_{i|i-1} = \Psi(i,k) \breve{x}_{k|k-1} + \sum_{j=k}^{i-1} \Psi(i,j+1) \left(A_j \hat{K}_j \mathcal{R}_j - \mathcal{Q}_j \right) \qquad (34)$$

for all $i \geq k$, where $\Psi(i,k) = G_{i-1} G_{i-2} \cdots G_k$ for all $i \geq k$, with $G_j = A_j(I - \hat{K}_j C_j)$ for all $j \geq k$, and $\Psi(k,k) = I$. From (34), Young's inequality and the boundedness of A_k, C_k, $\hat{K}_k$, $\mathcal{R}_k$ and $\mathcal{Q}_k$, we obtain that there exists a constant $\rho_\Psi \in \mathbb{R}_{>0}$ such that

$$-\sum_{i=k}^{k+N-1} \breve{x}_{i|i-1}^{\mathrm{T}} C_i^{\mathrm{T}} C_i \breve{x}_{i|i-1} \leq \frac{\rho_\Psi}{\rho_C}$$
$$-\frac{1}{2} \breve{x}_{k|k-1}^{\mathrm{T}} \left(\sum_{i=k}^{k+N-1} \Psi(i,k)^{\mathrm{T}} C_i^{\mathrm{T}} C_i \Psi(i,k) \right) \breve{x}_{k|k-1}. \qquad (35)$$

Because it follows from Assumption 5 and Anderson and Moore (1981, Lemma 3.1) that

$$\sum_{i=k}^{k+N-1} \Psi(i,k)^{\mathrm{T}} C_i^{\mathrm{T}} C_i \Psi(i,k) \succeq c_o I \qquad (36)$$

for some constant $c_o \in \mathbb{R}_{>0}$ and a sufficiently large integer N, we obtain from (29), (33), (35) and (36) that

$$V_{k+N} \leq \left(1 - \frac{c_o c_{\hat{P}1} \rho_C}{2} \right) V_k + N\rho_Q + N\rho_R + \rho_\Psi \qquad (37)$$

for all $k \geq N$. By recursively applying (37), we obtain that

$$V_{k_0 + iN} \leq V_{k_0} + \frac{2(N\rho_Q + N\rho_R + \rho_\Psi)}{c_o c_{\hat{P}1} \rho_C} \qquad (38)$$

for all $k_0 \geq N$ and all nonnegative integers i. From (29) and (38), it follows that

$$\|\breve{x}_{k_0+iN|k_0+iN-1}\| \leq \sqrt{\frac{c_{\hat{P}2}}{c_{\hat{P}1}}} \|\breve{x}_{k_0|k_0-1}\|$$
$$+ \sqrt{\frac{2c_{\hat{P}2}(N\rho_Q + N\rho_R + \rho_\Psi)}{c_o c_{\hat{P}1} \rho_C}} \qquad (39)$$

for all $k_0 \geq N$ and all nonnegative integers i. Because the boundedness of A_k, C_k, $\hat{K}_k$, $\mathcal{R}_k$ and $\mathcal{Q}_k$ implies that $\breve{x}_{k|k-1}$ remains bounded for all $0 \leq k \leq 2N$ under arbitrary initial conditions, we conclude from (39) that the error system of the LKF in (21) is UGB. Hence, the cascaded system in (19) and (21) is UGB. □

By combining Lemmas 4.1, 4.3 and 4.4, we obtain the following result.

Theorem 4.5. Under Assumptions 1-5, the cascaded nominal error system of the DKF in (19) and (21) is UGAS.

Proof. The proof follows directly from Lemmas 4.1, 4.3 and 4.4 and Loria and Nešić (2002, Theorem 3). □

Theorem 4.5 implies that the state estimate $\hat{x}_{k|k}$ provided by the DKF globally converges to the state x_k of the nonlinear system in (1) under nominal conditions if the given assumptions hold.

5. NUMERICAL EXAMPLES

In this section, two different nonlinear systems are considered. The first example involves an inherent nonlinearity in the dynamics. The second example has linear dynamics, although, parameter estimation casts as a nonlinear system. The estimation results using the DKF and the EKF are compared.

5.1 Nonlinearity in dynamics

Consider the nominal nonlinear dynamics of the Van der Pol oscillator

$$\dot{x}_1 = x_2 + w_1, \quad \dot{x}_2 = (1 - x_1^2)x_2 - x_1 + w_2, \quad y = x_1 + v$$

that can be discretized using Euler's method with sampling time $T_s = 0.01$ as

$$x_{k+1,1} = x_{k,1} + T_s x_{k,2} + T_s w_{k,1}, \qquad (40a)$$
$$x_{k+1,2} = x_{k,2} + T_s((1 - x_{k,1}^2)x_{k,2} - x_{k,1} + w_{k,2}), \quad (40b)$$
$$y_k = x_{k,1} + v_k, \qquad (40c)$$

where the states are $x_{k,1}$ and $x_{k,2}$. The disturbances w_k and v_k are modeled as Gaussian white noise with variances $diag[0.001, 0.001]$ and 0.1, respectively.

DKF design: The model (40) corresponds to Assumption 3 that allows us to use the measurement y_k to transform the nominal nonlinear dynamics to a LTV model with
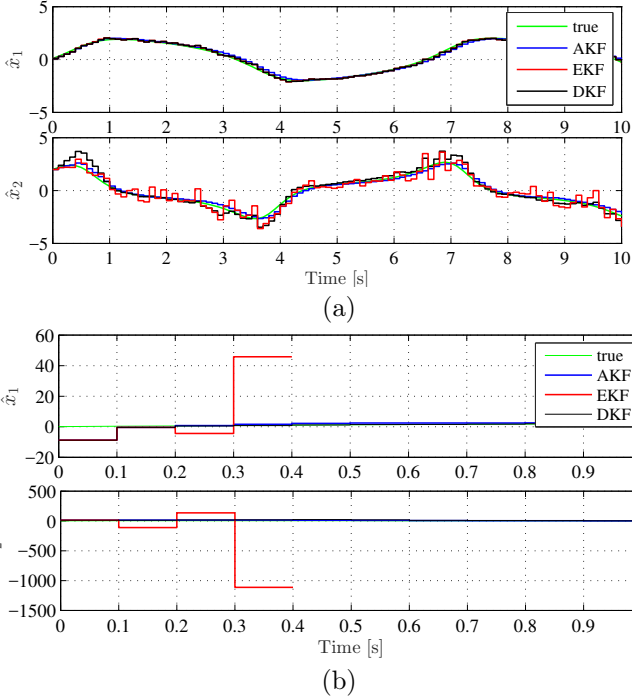
Fig. 2. Different estimation methods for the numerical example, (a) with correct initial estimate $\hat{x}_{0|0} = [0.1, 2]^{\mathrm{T}}$, for which the results with the EKF and the DKF are almost indistinguishable, and (b) with perturbed initial estimate $\hat{x}_{0|0} = [-8.8, 12]^{\mathrm{T}}$

$$F_k = \begin{bmatrix} 1 & T_s \\ -T_s & 1 + T_s(1 - y_k^2) \end{bmatrix}, \quad H_k = [1 \quad 0],$$

where the system matrix F_k is time varying. The implementation of the AKF is summarized by (7) and (8). The matrix $C_k = H_k$ and A_k of the LKF in (14) and (15) of the DKF are given as follows

$$A_k = \begin{bmatrix} 1 & T_s \\ -T_s(1 + 2\bar{x}_{k|k,1}\bar{x}_{k|k,2}) & 1 + T_s(1 - \bar{x}_{k|k,1}^2) \end{bmatrix}.$$

EKF design: We make use of the discrete-time EKF that linearizes the system about the state estimate to update the state covariance matrix as described in (Gelb, 1974). The tuning matrices of the EKF and the LKF

Table 1. accumulative estimation error

	AKF	DKF	EKF
Acc. error (a)	0.79	0.75	0.78
Acc. error (b)	271	263	N/A

are chosen as the covariances of w_k and v_k, i.e. $\hat{Q}_k = diag[0.001, 0.001]$ and $\hat{R}_k = 0.1$. The tuning matrices of the AKF are chosen differently to take into account the effects described in Remark 3.2, i.e. $\bar{Q}_k = diag[0.01, 0.01]$ and $\bar{R}_k = 0.1$. Since the state correction of the EKF is based on linearized dynamics, the wrong choice of initial condition can result in a suboptimal solution. The results are illustrated in Fig. 2, where the erroneous initial guess made the EKF to diverge and fail, even though the initial guess for covariance matrix has been updated to compensate the wrong choice of initial state. On the other hand, a correct initial guess for estimates, makes the algorithms have similar performance. Table 1 also provides a comparison between different estimation methods in the sense of accumulated error (Acc. error) for all scenarios. It

should be noted that the main benefit of the DKF over the EKF is its global convergence and any optimality analysis is outside the scope of this study.

5.2 Nonlinear augmented parameter estimation problem

In this example, a linear model for a vibration system is discretized using Euler's method with the sampling period T_s as follows

$$x_{k+1,1} = x_{k,1} + T_s(x_{k,2} + w_{k,1}), \quad y_k = x_{k,1} + v_k,$$
$$x_{k+1,2} = T_s(ax_{k,1} + bx_{k,2} + u_k + w_{k,2}),$$

where a and b are unknown parameters, u_k is the excitation signal, and the states are denoted by $x_{k,1}$ and $x_{k,2}$. By assuming the parameters constant over a sampling period, one can augment them into the system ($x_{k,3} := a$ and $x_{k,4} := b$) for estimation purposes. So, the nonlinear model for parameter estimation is

$$x_{k+1,1} = x_{k,1} + T_s(x_{k,2} + w_{k,1}), \tag{41a}$$
$$x_{k+1,2} = T_s(x_{k,3}x_{k,1} + x_{k,4}x_{k,2} + u_k + w_{k,2}), \tag{41b}$$
$$x_{k+1,3} = x_{k,3}, \quad x_{k+1,4} = x_{k,4}, \quad y_k = x_{k,1} + v_k, \tag{41c}$$

For the AKF design, the multiplicative nonlinearity can be removed under nominal conditions (i.e. $w_{k,1} = w_{k,2} = v_k = 0$). For $w_{k,1} = w_{k,2} = v_k = 0$, we obtain from (41) that $T_s x_{k-j,2} = y_{k-j+1} - y_{k-j}$ and $x_{j,1} = y_j$, for all $j = \{0, 1, \dots, k\}$. We need to calculate $x_{k,2}$ based on previous measurements to formulate a causal model for the AKF. So, rewriting the dynamics backward in time gives

$$T_s x_{k-2,2} = y_{k-1} - y_{k-2} = aT_s^2 y_{k-3} + bT_s(y_{k-2} - y_{k-3})$$
$$T_s x_{k-1,2} = y_k - y_{k-1} = aT_s^2 y_{k-2} + bT_s(y_{k-1} - y_{k-2})$$

where $a = x_{k,3}$ and $b = x_{k,4}$ are functions of y_{k-j+1}:

$$a = \frac{y_k y_{k-3} - y_{k-2}(y_k + y_{k-1}) + y_{k-1}^2}{T_s^2(y_{k-1}y_{k-3} - y_{k-2}^2)} - \frac{1}{T_s^2},$$
$$b = \frac{y_k y_{k-3} - y_{k-1}y_{k-2}}{T_s(y_{k-1}y_{k-3} - y_{k-2}^2)} - \frac{1}{T_s}.$$

Hence, $x_{k,2} = aT_s y_{k-1} + b(y_k - y_{k-1}) =: \phi(\mathcal{Y}_k^{k-3})$. Therefore, the nonlinear model in (41) can be transformed to the LTV model

$$x_{k+1} = \begin{bmatrix} 1 & T_s & 0 & 0 \\ 0 & 0 & T_s y_k & T_s \phi(\mathcal{Y}_k^{k-3}) \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ T_s \\ 0 \\ 0 \end{bmatrix} u_k, \tag{42a}$$
$$y_k = [1\ 0\ 0\ 0] x_k \tag{42b}$$

under nominal conditions. The DKF and the EKF are applied to estimate the parameters a and b, where (41) is linearized to obtain the LKF and the EKF, and where (42) is used for the AKF. The disturbance w_k and v_k are assumed to be Gaussian white noise with variances $diag[10^{-4}, 10^{-4}]$ and 0.01, respectively. The tuning matrices of the EKF and the LKF are chosen accordingly as $\hat{Q}_k = diag[10^{-4}, 10^{-4}, 10, 10]$ and $\hat{R}_k = 0.01$ to obtain the best performance, while the ones for AKF are slightly different: $\bar{Q}_k = diag[10^{-3}, 10^{-3}, 10^4, 10^2]$ and $\bar{R}_k = 0.01$.

Simulation results are presented in Fig. 3, where two scenarios are compared like in the previous example. In Fig. 3 (a), the second parameter estimate (denoted by $\hat{x}_4$) of the AKF shows a bias that is due to the neglected disturbances, which is compensated for by the LKF. Finally, in Fig. 3 (b) it is illustrated that a wrong
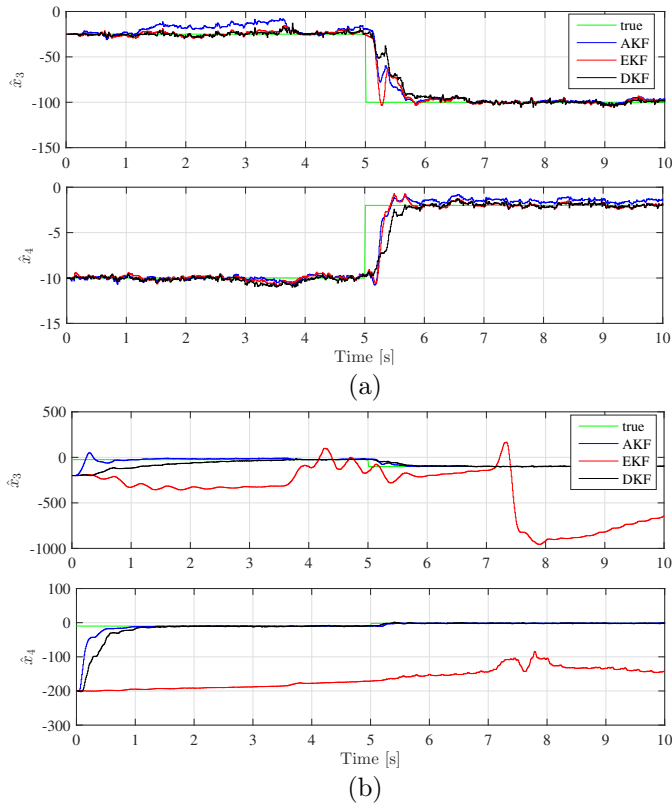
Fig. 3. Comparison of estimation methods, (a) with correct initial estimate $\hat{x}_{0|0} = [0, 0, -25, -10]^{\mathrm{T}}$ and (b) with perturbed initial estimate $\hat{x}_{0|0} = [0, 0, -200, -200]^{\mathrm{T}}$.

choice of initial guess for the parameters results in an erroneous linearization point for the EKF. Consequently, the estimates diverge to wrong values, although the initial covariance matrix has been updated accordingly. On the other hand, the DKF converges to the right value even though it uses the same initialization.

6. CONCLUSION

This paper summarizes the nominal error stability analysis of a two stage estimation method for nonlinear systems in the discrete-time domain. Since the so-called DKF uses an equivalent LTV model in its first stage, global convergence is achieved. Sufficient conditions for uniformly globally asymptotic stability under nominal conditions are derived. Furthermore, simulation results validate the stability of the method.

REFERENCES

Abraham, R., Marsden, J.E., and Ratiu, T. (2012). *Manifolds, tensor analysis, and applications*. Springer.

Anderson, B.D.O. and Moore, J.B. (1981). Detectability and stabilizability of time-varying discrete-time linear systems. *SIAM Journal on Control and Optimization*, 19(1), 20–32.

Besançon, G. (1999). On output transformations for state linearization up to output injection. *IEEE Transactions on Automatic Control*, 44(10), 1975–1981.

Califano, C., Monaco, S., and Normand-Cyrot, D. (2009). Canonical observer forms for multi-output systems up to

coordinate and output transformations in discrete time. *Automatica*, 45(11), 2483–2490.

Ficocelli, M. and Janabi, F. (2001). Adaptive filtering for pose estimation in visual servoing. In *Intelligent Robots and Systems, 2001. Proceedings. 2001 IEEE/RSJ International Conference on*, volume 1, 19–24.

Fliess, M., Lévine, J., Martin, P., and Rouchon, P. (1995). Flatness and defect of non-linear systems: introductory theory and examples. *International journal of control*, 61(6), 1327–1361.

Folland, G. (1990). Remainder estimates in Taylor's theorem. *The American Mathematical Monthly*, 97(3), 233–235.

Gelb, A. (1974). *Applied optimal estimation*. MIT press.

Hammouri, H. and Gauthier, J. (1992). Global time-varying linearization up to output injection. *SIAM journal on control and optimization*, 30(6), 1295–1310.

Haseltine, E.L. and Rawlings, J.B. (2005). Critical evaluation of extended Kalman filtering and moving-horizon estimation. *Industrial & engineering chemistry research*, 44(8), 2451–2460.

Jazwinski, A.H. (2007). *Stochastic processes and filtering theory*. Courier Corporation.

Johansen, T.A. and Fossen, T.I. (2016a). The eXogenous Kalman filter (XKF). *International Journal of Control*.

Johansen, T.A., Fossen, T.I., and Goodwin, G.C. (2016). Three-stage filter for position estimation using pseudo-range measurements. *IEEE Trans. Aerospace and Electronic Systems*, 52.

Johansen, T.A. and Fossen, T.I. (2016b). Nonlinear filtering with eXogenous Kalman filter and double Kalman filter. In *European Control Conference*.

Lee, S.H. and West, M. (2010). Performance comparison of the distributed extended Kalman filter and Markov chain distributed particle filter (MCDPF). *IFAC Proceedings Volumes*, 43(19), 151–156.

Levine, J. and Marino, R. (1986). Nonlinear system immersion, observers and finite-dimensional filters. *Systems & Control Letters*, 7(2), 133–142.

Loria, A. and Nešić, D. (2002). Stability of time-varying discrete-time cascades. *IFAC Proceedings Volumes*, 35(1), 263–268.

Perea, L., How, J., Breger, L., and Elosegui, P. (2007). Nonlinearity in sensor fusion: Divergence issues in EKF, modified truncated SOF, and UKF. In *AIAA Guidance, Navigation and Control Conference and Exhibit*.

Rigatos, G.G. (2015). Differential flatness theory and flatness-based control. In *Nonlinear Control and Filtering Using Differential Flatness Approaches*, 47–101. Springer.

Simon, D. (2006). *Optimal State Estimation: Kalman, H_∞, and Nonlinear Approaches*. Wiley–Interscience.

Simon, D. (2010). Kalman filtering with state constraints: a survey of linear and nonlinear algorithms. *IET Control Theory & Applications*, 4(8), 1303–1318.

Sira-Ramirez, H. and Castro-Linares, R. (2000). Sliding mode rest-to-rest stabilization and trajectory tracking for a discretized flexible joint manipulator. *Dynamics and Control*, 10(1), 87–105.

Sira-Ramirez, H. and Agrawal, S.K. (2004). *Differentially flat systems*. CRC Press.