Journal of Mathematical Systems, Estimation, and Control
Vol. 5, No. 1, 1995, pp. 59–78
© 1995 Birkhäuser-Boston

The Extended Kalman Filter as a Local Asymptotic Observer for Discrete-Time Nonlinear Systems*

Yongkyu Song[†] Jessy W. Grizzle[†]

Abstract

The convergence aspects of the extended Kalman filter, when used as a deterministic observer for a nonlinear discrete-time system, are analyzed. Systems with nonlinear output maps are treated, and the conditions needed to ensure the uniform boundedness of the error covariances are related to the observability properties of the underlying nonlinear system. Furthermore, the uniform asymptotic convergence of the observation error is established whenever the nonlinear system satisfies an observability rank condition and the states stay within a convex compact domain. This last result provides a theoretical foundation for this classic, approximate nonlinear filter.

Key words: discrete-time nonlinear systems, observability rank condition, extended Kalman filter, uniform asymptotic stability

AMS Subject Classifications: 93C55, 93D20, 93E11

1 Introduction

Designing an observer for a nonlinear system is quite a challenge. Thus, as a first step, it is interesting to see how classical linearization techniques work with nonlinear systems and what their limitations are. In [4], Baras et al. describe a method for constructing observers for dynamic systems as asymptotic limits of filters. They discuss the method as applied to the linear case, and a class of nonlinear systems with linear observations[1], in the continuous-time domain. Essentially the extended Kalman filter(EKF)

*Received April 8, 1992; received in final form April 6, 1993.
[†]Work supported by the National Science Foundation under contract NSF ECS-88-96136 with matching funds provided by the FORD MOTOR CO.
[1]See also [13] for the case of nonlinear outputs.

is used as their observer ([4], [5]). The extended Kalman filter is a well-known standard linearization method for approximate nonlinear filtering. The available literature is vast and we refer the reader to [14], [21], [28], and the references therein. In particular, in the context of parameter estimation for linear stochastic systems, a fairly systematic and comprehensive convergence analysis of the EKF is given in [22].

Our work is most closely related to the work of Baras et al. [4]. We will consider the system:

$$\begin{aligned} x_{k+1} &= f(x_k, u_k), \quad x_0 \text{ unknown}, \\ y_k &= h(x_k, u_k), \end{aligned} \tag{1.1}$$

and the EKF for the associated "noisy" system:

$$\begin{aligned} z_{k+1} &= f(z_k, u_k) + N w_k, \\ \xi_k &= h(z_k, u_k) + R v_k. \end{aligned} \tag{1.2}$$

Throughout the paper, we assume that $x, w \in \mathbf{R}^n, u \in \mathbf{R}^m$ and $y, v \in \mathbf{R}^p$, and that f, h are at least twice differentiable. As usual, z_0, v_k, and w_k are assumed jointly Gaussian and mutually independent. Furthermore $z_0 \sim \mathcal{N}(\bar{x}_0, \bar{P}_0), w_k \sim \mathcal{N}(0, I_n)$, and $v_k \sim \mathcal{N}(0, I_p)$. We also assume that the design variables N, R, and $\bar{P}_0$ are always chosen such that N has rank n and R and $\bar{P}_0$ are positive definite.

We denote by $|\cdot|$, the Euclidean norm of a vector, and by $\|\cdot\|$ and $|||\cdot|||$, the induced norms on matrices and tensors. We also adopt the following notations. Given two symmetric matrices P and Q, of the same dimension, the inequality $P \geq Q$ means that the difference $P - Q$ is nonnegative definite. Similarly, $P > Q$ means that $P - Q$ is positive definite. A symmetric matrix Q is said to be bounded from above (below, respectively) if there is a number $q > 0$ such that $Q \leq qI (qI \leq Q)$. The symbol ":=" means that the RHS is defined to be equal to the LHS; the reverse holds for " =: ".

In Section 2, we give a new, simple proof which shows that the Kalman filter is a global observer for (discrete-time) linear time-varying systems. Based on this proof, we make an extension to the case of nonlinear systems with nonlinear output maps in Section 3. The conditions needed to ensure the uniform boundedness of the error covariances are then related to the observability properties of the underlying nonlinear system in Section 4. In Section 5, convergence of the error is proven under the observability rank condition as long as the states stay within a convex compact set, which is not necessarily small. These results show that the EKF is a quasi-local observer [11]. Conclusions are made in Section 6.

2 The Kalman Filter: A Global Asymptotic Observer for Linear Time-Varying Systems

It is well known that, under stochastic controllability and observability assumptions, the Kalman filter for a linear time-varying system with artificial noises can be used as a global asymptotic observer for the underlying deterministic system [9]. This fact can be also seen from the duality of a linear optimal regulator problem [18, p. 535]. In this Section, we give a new, simple proof, which is essential for setting up the analysis on nonlinear systems done in Section 3 through Section 5.

Consider a linear system:

$$\begin{aligned} x_{k+1} &= A_k x_k + B_k u_k, \quad x_0 \text{ unknown}, \\ y_k &= C_k x_k, \end{aligned} \tag{2.1}$$

where A_k is assumed invertible[2], and consider also the associated "*noisy*" system:

$$\begin{aligned} z_{k+1} &= A_k z_k + B_k u_k + N w_k, \\ \xi_k &= C_k z_k + R v_k, \end{aligned} \tag{2.2}$$

where the design parameters N and R are to be chosen as positive definite matrices. Then the Kalman filter equations for (2.2) are given as follows [1].

Measurement update:

$$\begin{aligned} \hat{x}_k &= \bar{x}_k + K_k(\xi_k - C_k \bar{x}_k), \\ P_k^{-1} &= \bar{P}_k^{-1} + C_k^T (RR^T)^{-1} C_k, \end{aligned} \tag{2.3}$$

Time update:

$$\begin{aligned} \bar{x}_{k+1} &= A_k \hat{x}_k + B_k u_k, \\ \bar{P}_{k+1} &= A_k P_k A_k^T + NN^T, \end{aligned} \tag{2.4}$$

$$K_k = P_k C_k^T (RR^T)^{-1} = \bar{P}_k C_k^T (C_k \bar{P}_k C_k^T + RR^T)^{-1}$$

where $\bar{P}_k$ and P_k are the *a priori* and *a posteriori* error covariances, and $\bar{x}_k$ and $\hat{x}_k$ the *a priori* and *a posteriori* estimates of the state at time k, respectively. The filter is initiated with $\bar{x}_0$ and $\bar{P}_0$; $\bar{P}_0$ is used as a design parameter, assumed also positive definite.

[2]This assumption can be relaxed to singular state transition matrices if a linear system is considered [25]. Toward nonlinear systems, however, we make this stronger assumption here.

To obtain an error dynamics, let's rewrite the Kalman filter in terms of the *a priori* variables. From (2.3) and (2.4) we have, noting that we use y_k instead of ξ_k,

$$\begin{aligned} \bar{x}_{k+1} &= A_k(I - K_kC_k)\bar{x}_k + B_ku_k + A_kK_ky_k, &(2.5)\\ \bar{P}_{k+1} &= A_k(I - K_kC_k)\bar{P}_kA_k^T + NN^T. &(2.6)\end{aligned}$$

If we define the error as $e_k := x_k - \bar{x}_k$, then the error dynamics is given as

$$e_{k+1} = A_k(I - K_kC_k)e_k. \tag{2.7}$$

The associated Riccati equations for the error covariances are

$$\begin{aligned} \bar{P}_{k+1} &= A_k[\bar{P}_k^{-1} + C_k^T(RR^T)^{-1}C_k]^{-1}A_k^T + NN^T, &(2.8)\\ P_{k+1}^{-1} &= [A_kP_kA_k^T + NN^T]^{-1} + C_k^T(RR^T)^{-1}C_k. &(2.9)\end{aligned}$$

Note that $\bar{P}_0 > 0$ and rank $N = n$ implies $\bar{P}_k > 0$ and $P_k > 0$ for all $0 \le k < \infty$.

Since we are interested in the asymptotic behavior of the error, e_k, it is necessary to obtain bounds for $\|\bar{P}_k\|$ and $\|P_k^{-1}\|$. Deyst and Price [9] obtained a sufficient condition which gives lower and upper bounds on P_k as follows.

Lemma 2.1 *Consider the following "noisy" system:*

$$\begin{aligned} x_{k+1} &= A_kx_k + Nw_k,\\ y_k &= C_kx_k + Rv_k. \end{aligned} \tag{2.10}$$

Suppose that there are positive real numbers $\alpha_1, \alpha_2, \beta_1, \beta_2$ such that the following conditions hold for some finite $M \ge 0$ and for all $k \ge M$:

$$\alpha_1 I \ge \sum_{i=k-M}^{k-1} \Phi(k, i+1)NN^T\Phi^T(k, i+1) \ge \alpha_2 I, \tag{2.11}$$

$$\beta_1 I \le \sum_{i=k-M}^{k} \Phi^T(i,k)C_i^T(RR^T)^{-1}C_i\Phi(i,k) \le \beta_2 I; \tag{2.12}$$

then

$$\frac{1}{\beta_2 + 1/\alpha_2} I \le P_k \le (\alpha_1 + 1/\beta_1)I,$$

where

$$\Phi(k,i) := A_{k-1}A_{k-2}\cdots A_i.$$

Remark 2.2 The conditions (2.11) and (2.12) imply that the "noisy" system (2.10) is stochastically controllable and observable [3]. It is easily seen that the positive definiteness of N implies the stochastic controllability through the condition (2.11). For linear systems, this requirement can be weakened to stabilizability [2] or even to nonstabilizability under a few more assumptions [7]. For the ease of presentation and the nonlinear systems to be considered later, however, we use this stronger assumption here. On the other hand, let's take $R = I$, R being a design parameter; then condition (2.12) is satisfied if the deterministic part of the system (2.10), i.e., the pair (A_k, C_k), is uniformly completely observable [18]. □

Remark 2.3 Under the above conditions, it can be also shown that $\bar{P}_k$ is bounded from above and below. Indeed, from (2.4),

$$||\bar{P}_k|| \leq (\alpha_1 + 1/\beta_1)||A||^2 + ||N||^2.$$

Also, from (2.3),

$$\bar{P}_k \geq P_k \geq \frac{1}{\beta_2 + 1/\alpha_2} I.$$

Therefore,

$$\frac{1}{\beta_2 + 1/\alpha_2} I \leq \bar{P}_k \leq \{(\alpha_1 + 1/\beta_1)||A||^2 + ||N||^2\} I.$$

It is obvious that P_k^{-1} and $\bar{P}_k^{-1}$ are both bounded from above and below. □

Remark 2.4 Deyst and Price [9] have also shown that, under stochastic controllability and observability assumptions, the homogeneous filter equations of the *a posteriori* estimates are uniformly asymptotically stable. Since, in [9], $A_k^T A_k$ is assumed bounded from above and below in norm, and $\bar{x}_{k+1} = A_k \hat{x}_k$ when the control variable is not considered, uniform asymptotic stability also holds for the homogeneous filter equations of the *a priori* estimates (2.5), which is exactly the same as the error dynamics (2.7).

Baras et al. [4] have also obtained bounds for the error covariances in continuous-time via dual optimal control problems under some "stronger" observability and controllability assumptions (see conditions (28) and (29) in [4]) and used the bounds to show the convergence of the error. Similar methods yield bounds for the error covariances in discrete-time. Bounds for the case of linear time invariant systems are explicitly shown in [26], but in this case they follow simply from the detectability of the pair (A, C) and the invertibility of A. In Section 4 we will discuss how the observability of a nonlinear discrete-time system is related to the boundedness of the error covariances in the associated extended Kalman filter later.

Now we state a theorem on the convergence of the error, on which an extention is made to the nonlinear systems later.

Theorem 2.5 *Consider the system (2.1) and the Kalman filter equations (2.3) and (2.4) for the associated system (2.2). Suppose that the system (2.1) is uniformly observable and* A_k *is invertible for all* k*, and that* $\|A\| := sup\{\|A_k\| : k \geq 0\}$ *and* $\|C\| := sup\{\|C_k\| : k \geq 0\}$ *are bounded. Then the Kalman filter for the noisy system (2.2) is a global, uniform asymptotic observer for the deterministic system (2.1).*

Remark 2.6 In the proof the Lyapunov function $V(k, e_k) = e_k^T \bar{P}_k^{-1} e_k$ is used, and the Riccati equations for the error covariances are exploited to obtain the uniform asymptotic stability of the error dynamics. For the detailed proof the reader may refer to [26].

3 General Nonlinear Systems

In this section the results for linear systems are extended to general nonlinear systems of the form (1.1). For simplicity of notation[3], we consider a system without controls:

$$\begin{aligned} x_{k+1} &= f(x_k), \quad x_0 \text{ unknown}, \\ y_k &= h(x_k), \end{aligned} \tag{3.1}$$

and its associated "*noisy*" system:

$$\begin{aligned} z_{k+1} &= f(z_k) + N w_k, \\ \xi_k &= h(z_k) + R v_k. \end{aligned} \tag{3.2}$$

The extended Kalman filter for the associated system is given as follows [1].

Measurement update:

$$\begin{aligned} \hat{x}_k &= \bar{x}_k + K_k(\xi_k - h(\bar{x}_k), \\ P_k^{-1} &= \bar{P}_k^{-1} + C_k^T (RR^T)^{-1} C_k, \end{aligned} \tag{3.3}$$

Time update:

$$\begin{aligned} \bar{x}_{k+1} &= f(\hat{x}_k), \\ \bar{P}_{k+1} &= A_k P_k A_k^T + NN^T, \end{aligned} \tag{3.4}$$

[3] The modifications necessary to handle inputs are indicated at the end of the section.

where

$$
\begin{aligned}
K_k &:= \bar{P}_k C_k^T (C_k \bar{P}_k C_k^T + RR^T)^{-1}, \\
A_k &:= \frac{\partial f}{\partial x}(\hat{x}_k), \\
C_k &:= \frac{\partial h}{\partial x}(\bar{x}_k).
\end{aligned}
$$

The Riccati equations for the error covariances are the same as in (2.8) and (2.9) with the above matrices.

To begin with, we make the following assumptions for setting up the analysis. Section 4 addresses how Assumption 3.1.1 is implied by an observability property of (3.1); the other conditions are addressed in Section 5.

Assumption 3.1

1. *The linearized system along the estimated trajectory of the extended Kalman filter is uniformly observable, that is, (A_k, C_k) of (3.3) and (3.4) satisfies the uniform observability condition.*

2. *$A(x) := \frac{\partial f}{\partial x}(x)$ is invertible at each $x \in R^n$.*

3. *The following norms are bounded;*

$$
\|A\| := \sup_{x \in R^n} \|A(x)\|, \quad \|A^{-1}\| := \sup_{x \in R^n} \|[A(x)]^{-1}\|,
$$

$$
\|H\| := \sup_{x \in R^n} \|R^{-1}\frac{\partial h}{\partial x}(x)\|, \quad |||D^2 f||| = \sup_{x \in R^n} |||D^2 f(x)|||,
$$

$$
|||D^2 h||| = \sup_{x \in R^n} |||D^2 h(x)|||.
$$

4. *Let $g(x,y) := h(x) - h(y) - \frac{\partial h}{\partial x}(x)(x-y)$, and suppose that there exists $g < \infty$ such that $|g(x,y)| \le g|||D^2 h|||\,|x-y|^2$ for all $x, y \in R^n$.*

Assumption 3.1.1 implies that the error covariances are uniformly bounded. Thus let $0 < q, p_1 < \infty$ be the corresponding bounds for error covariances, that is, $\|\bar{P}_k\| \le q$ and $\|P_k^{-1}\| \le p_1$ for all $k \ge 0$. For later use we derive a few more bounds. From (3.3)

$$
\bar{P}_M^{-1} = P_M^{-1} - H_M^T H_M,
$$

thus giving

$$
\|\bar{P}_M^{-1}\| \le \|P_M^{-1}\| + \|H\|^2 \le p_1 + \|H\|^2 := p.
$$

Also from (3.3),

$$\|P_M\| \leq \|\bar{P}_M\| \leq q.$$

Furthermore,

$$\|I - K_M C_M\| = \|P_M \bar{P}_M^{-1}\| \leq pq$$

and

$$\|K_M\| = \|P_M C_M^T (RR^T)^{-1}\| \leq q\|H\|\|R^{-1}\|^2 =: \delta.$$

Now to prove convergence, set

$$\begin{aligned} V(k, e_k) &= e_k^T \bar{P}_k^{-1} e_k, \\ \phi(|e_k|, X, Y) &:= \delta g Y \|A\| + \tfrac{1}{2} X (pq + \delta g Y |e_k|)^2, \\ \varphi(|e_k|, X, Y) &:= -\tfrac{1}{rq^2} + p|e_k|\phi(|e_k|, X, Y)\{2pq\|A\| + \phi(|e_k|, X, Y)|e_k|\}. \end{aligned}$$

Theorem 3.2 *Consider the system (3.1) and the extended Kalman filter equations (3.3) and (3.4) for the associated system (3.2). Suppose that Assumption 3.1 holds. Then, if* $|e_0|, |||D^2 f|||$, *and* $|||D^2 h|||$ *are such that for some* $\gamma > 0$,

$$\varphi(q^{\frac{1}{2}} V^{\frac{1}{2}}(0, e_0), |||D^2 f|||, |||D^2 h|||) \leq -\gamma$$

then the extended Kalman filter for the noisy system (3.2) is a local, uniform asymptotic observer for the deterministic system (3.1).

Proof: Let $e_k = x_k - \bar{x}_k$. Then

$$\begin{aligned} e_{k+1} &= f(x_k) - f(\hat{x}_k) \\ &= \int_0^1 Df(\hat{x}_k + s\tilde{e}_k) ds \tilde{e}_k \end{aligned}$$

where

$$\tilde{e}_k = x_k - \hat{x}_k = x_k - \bar{x}_k - K_k(h(x_k) - h(\bar{x}_k)).$$

Note also that

$$\begin{aligned} \tilde{e}_k &= e_k - K_k(C_k e_k + g_k) \\ &= (I - K_k C_k) e_k - K_k g_k. \end{aligned}$$

Thus, using the above equation,

$$\begin{aligned} e_{k+1} &= [A_k + \int_0^1 (Df(\hat{x}_k + s\tilde{e}_k) - Df(\hat{x}_k)) ds] \tilde{e}_k \\ &= [A_k + \int_0^1 \int_0^1 D^2 f(\hat{x}_k + rs\tilde{e}_k) s\tilde{e}_k dr ds] \tilde{e}_k \\ &= [A_k + B_k] \tilde{e}_k \\ &= A_k[(I - K_k C_k) e_k - K_k g_k] + B_k \tilde{e}_k \\ &= A_k(I - K_k C_k) e_k + l_k, \end{aligned}$$

where

$$\begin{aligned} B_k &= \int_0^1 \int_0^1 D^2 f(\hat{x}_k + rs\tilde{e}_k) s\tilde{e}_k drds \\ l_k &= -A_k K_k g_k + B_k \tilde{e}_k. \end{aligned}$$

Hence,

$$\begin{aligned} e_{k+1}^T \bar{P}_{k+1}^{-1} e_{k+1} &= (e_k^T (I - K_k C_k)^T A_k^T + l_k^T) \bar{P}_{k+1}^{-1} (A_k (I - K_k C_k) e_k + l_k) \\ &= e_k^T (I - K_k C_k)^T A_k^T \bar{P}_{k+1}^{-1} A_k (I - K_k C_k) e_k \\ &\quad + l_k^T \bar{P}_{k+1}^{-1} A_k \times (I - K_k C_k) e_k \\ &\quad + e_k^T (I - K_k C_k)^T A_k^T \bar{P}_{k+1}^{-1} l_k + l_k^T \bar{P}_{k+1}^{-1} l_k. \end{aligned}$$

Using the linear results,

$$\begin{aligned} \Delta V(k, e_k) &= e_{k+1}^T \bar{P}_{k+1}^{-1} e_{k+1} - e_k^T \bar{P}_k^{-1} e_k \\ &\leq -e_k^T \bar{P}_k^{-1} (P_k^{-1} + A^T (N N^T)^{-1} A)^{-1} \bar{P}_k^{-1} e_k + l_k^T \bar{P}_{k+1}^{-1} A_k (I \\ &\quad - K_k C_k) e_k + e_k^T (I - K_k C_k)^T A_k^T \bar{P}_{k+1}^{-1} l_k + l_k^T \bar{P}_{k+1}^{-1} l_k. \end{aligned}$$

With the definition of $g_k = g(x_k, \hat{x}_k)$, since

$$\begin{aligned} |\tilde{e}_k| &= |(I - K_k C_k) e_k - K_k g_k| \\ &\leq \|I - K_k C_k\| \, |e_k| + \|K_k\| \, |g_k| \\ &\leq (pq + \delta g \||D^2 h\|| \, |e_k|) |e_k|, \end{aligned}$$

and

$$\begin{aligned} \|B_k\| &= \| \int_0^1 \int_0^1 D^2 f(\hat{x}_k + rs\tilde{e}_k) s\tilde{e}_k drds \| \\ &\leq \int_0^1 \int_0^1 \||D^2 f\|| s drds |\tilde{e}_k| = \frac{1}{2} \||D^2 f\|| \, |\tilde{e}_k|, \end{aligned}$$

it follows that

$$\begin{aligned} |l_k| &= | - A_k K_k g_k + B_k \tilde{e}_k| \\ &\leq \phi(|e_k|, \||D^2 f\||, \||D^2 h\||) |e_k|^2 \end{aligned}$$

and

$$\begin{aligned} & l_k^T \bar{P}_{k+1}^{-1} A_k (I - K_k C_k) e_k + e_k^T (I - K_k C_k)^T A_k^T \bar{P}_{k+1}^{-1} l_k + l_k^T \bar{P}_{k+1}^{-1} l_k \\ &\quad \leq \|\bar{P}_{l+1}^{-1}\| \, |l_k| (2\|A\| \, \|I - K_k C_k\| \, |e_k| + |l_k|) \\ &\quad \leq p |e_k|^3 \phi(|e_k|, \||D^2 f\||, \||D^2 h\||) \{2pq \|A\| \\ &\quad\quad + \phi(|e_k|, \||D^2 f\||, \||D^2 h\||) |e_k|\}. \end{aligned}$$

Therefore,

$$\Delta V(k,e_k) \le \varphi(|e_k|, |||D^2f|||, |||D^2h|||)|e_k|^2.$$

A simple argument shows that if $\varphi(q^{\frac{1}{2}}V^{\frac{1}{2}}(0,e_0), |||D^2f|||, |||D^2h|||) \le -\gamma$ then
$\Delta V(k,e_k) \le -\gamma |e_k|^2$ for all $k \ge 0$. Thus e_k uniformly converges to zero. □

Remark 3.3

(a) If the observation map is linear, i.e., $h(x) = Cx$, then $D^2h \equiv 0$. It follows that $\varphi(|e_k|, |||D^2f|||, |||D^2h|||) = -\frac{1}{rq^2} + \frac{p^4q^3}{2}|e_k| \cdot |||D^2f|||$ $(2\|A\| + \frac{pq}{2}|e_k| \cdot |||D^2f|||)$. Let ζ^+ be the real, positive solution of the equation $-\frac{1}{rq^2} + \frac{p^4q^3}{2}\zeta(2\|A\| + \frac{pq}{2}\zeta) = -\gamma, 0 < \gamma < \frac{1}{rq^2}$; ζ^+ is a function of the design variables $N, R, \bar{P}_0, \gamma$. Under Assumption 3.1, if

$$|e_0| \cdot |||D^2f||| \le \max_{N,R,\bar{P}_0,\gamma} \frac{\zeta^+}{(pq)^{1/2}}, \tag{3.5}$$

the extended Kalman filter (3.3) and (3.4) with $\xi_k = y_k$ is a local asymptotic observer for the deterministic system (3.1) with linear observations. We note that the condition (3.5) can be satisfied if either $|e_0|$ or $|||D^2f|||$ is small enough, in other words, if either the estimate of the initial state is close enough to the true value or f is only weakly nonlinear. If the observation map is nonlinear but $|||D^2h|||$ is small, similar reasoning applies.

(b) Even if the output map is nonlinear it may be *locally* transformed via a coordinate change into a linear form provided the Jacobian of h has constant rank. In order to obtain this result rigorously the noisy system has to be constructed after the coordinate change since otherwise the noisy terms become state dependent.

(c) With known controls we can construct in the same way a local asymptotic observer for systems with inputs of the form (1.1), using the extended Kalman filter for the associated "*noisy*" system (1.2).

4 Observability Conditions of a Nonlinear System and its Linearization

In this section we discuss the observability condition in relation to the EKF. First, consider the system (2.10). If we use $R = I$, the observability condition (2.12) becomes

$$\beta_1 I \le \sum_{i=k-M}^{k} \Phi^T(i,k)C_i^T C_i \Phi(i,k) \le \beta_2 I, \quad 0 < \beta_1, \beta_2 < \infty. \tag{4.1}$$

If we assume further that $A_k^T A_k \geq \nu I > 0 \ \forall k$, then condition (4.1) is equivalent to the following condition: for some η_1, η_2, $0 < \eta_1 \leq \eta_2 < \infty$,

$$\eta_1 I \leq O^T(k-M,k)O(k-M,k) \leq \eta_2 I, \tag{4.2}$$

where

$$O(k-M,k) := \begin{bmatrix} C_{k-M} \\ C_{k-M+1}A_{k-M} \\ \vdots \\ C_k A_{k-1} \cdots A_{k-M} \end{bmatrix}.$$

In order to apply this linear observability condition to the EKF (3.3) and (3.4) and, ultimately, to relate this to observability properties of the underlying nonlinear system, let's represent $O(k-M,k)$ in terms of the EKF variables in (3.3) and (3.4), i.e.,

$$O_e(\bar{x}_{k-M}, \cdots, \bar{x}_k, \hat{x}_{k-M}, \cdots, \hat{x}_{k-1}) := \begin{bmatrix} \frac{\partial h}{\partial x}(\bar{x}_{k-M}) \\ \frac{\partial h}{\partial x}(\bar{x}_{k-M+1})\frac{\partial f}{\partial x}(\hat{x}_{k-M}) \\ \vdots \\ \frac{\partial h}{\partial x}(\bar{x}_k)\frac{\partial f}{\partial x}(\hat{x}_{k-1}) \cdots \frac{\partial f}{\partial x}(\hat{x}_{k-M}) \end{bmatrix} \tag{4.3}$$

Define the map $H : \mathbb{R}^n \to (\mathbb{R}^p)^n$ by

$$H(x) := (h(x), h(f(x)), \cdots, h(f^{n-1}(x)))^T. \tag{4.4}$$

A system is said to satisfy the *observability rank condition* at x_0 [24] if the rank[4] of the map H at x_0 equals n. A system satisfies the *observability rank condition on $\mathcal{O}$* if this is true for every $x \in \mathcal{O}$; if $\mathcal{O} = \mathbb{R}^n$, then one says that the system satisfies the observability rank condition. By the chain rule,

$$\begin{aligned} \frac{\partial H}{\partial x}(x_0) &= \begin{bmatrix} \frac{\partial h}{\partial x}(x_0) \\ \frac{\partial h}{\partial x}(x_1)\frac{\partial f}{\partial x}(x_0) \\ \vdots \\ \frac{\partial h}{\partial x}(x_{n-1})\frac{\partial f}{\partial x}(x_{n-2}) \cdots \frac{\partial f}{\partial x}(x_0) \end{bmatrix} \\ &=: \frac{\partial H}{\partial x}(x_0, x_1, \cdots, x_{n-1}) \end{aligned} \tag{4.5}$$

where $x_{k+1} = f(x_k)$, $k = 0, 1, \cdots, n-2$. It follows that $O_e = \frac{\partial H}{\partial x}$ if $\bar{x}_k$ and $\hat{x}_k$ are equal to the true state x_k, for $k = 0, 1, \cdots, n-1$. By continuity, we can argue that if the system (3.1) satisfies the observability rank condition, then its associated EKF satisfies the observability condition (4.2), for $M =$

[4] Recall that the rank of H at x_0 equals the rank of $\frac{\partial H}{\partial x}(x)$ evaluated at x_0.

$n-1$, whenever the estimates $\bar{x}_k$ and $\hat{x}_k$ are "sufficiently" close to the true state x_k. The boundedness of the error covariances would then follow from Deyst and Price [9]. This line of reasoning is made precise in Proposition 4.1 below and in Section 5. Let $\bar{X}_{n-1} := (\bar{x}_0, \cdots, \bar{x}_{n-1}) \in (\mathbb{R}^n)^n$ and $\hat{X}_{n-2} := (\hat{x}_0, \cdots, \hat{x}_{n-2}) \in (\mathbb{R}^n)^{n-1}$ and view O_e as a function of $\bar{X}_{n-1}$ and $\hat{X}_{n-2}$.

Proposition 4.1 *Suppose that the system (3.1) satisfies the observability rank condition on a compact subset $K \subset \mathbb{R}^n$. If the estimates of the EKF are sufficiently close to the true state, then the linearized system along the estimated trajectory of the extended Kalman filter is uniformly observable; i.e., there exist $\gamma_1, \gamma_2, 0 < \gamma_1 \le \gamma_2 < \infty$ and $\delta_1 > 0$ such that*

$$\gamma_1 I \le O_e^T(\bar{X}_{n-1}, \hat{X}_{n-2}) O_e(\bar{X}_{n-1}, \hat{X}_{n-2}) \le \gamma_2 I \tag{4.6}$$

for all $\bar{x}_l \in K$ such that $|\bar{x}_l - x_l| \le \delta_1, l = 0, \cdots, n-1$, and all $\hat{x}_j \in K$ such that $|\hat{x}_j - x_j| \le \delta_1, j = 0, \cdots, n-2$, and for each $x_0 \in K$, where $x_{l+1} = f(x_l), l = 0, \cdots, n-2$.

Proof: Let $X_{n-1} := (x_0, x_1, \cdots, x_{n-1}) \in (\mathbb{R}^n)^n$ and $X_{n-2} := (x_0, x_1, \cdots, x_{n-2}) \in (\mathbb{R}^n)^{n-1}$. Now view $J_e(X_{n-1}, X_{n-2}) := O_e^T(X_{n-1}, X_{n-2}) O_e(X_{n-1}, X_{n-2})$ as a function on $\Omega := K^n \times K^{n-1}$. Let's denote the Euclidean norms on $\mathbb{R}^n$ and $(\mathbb{R}^n)^n \times (\mathbb{R}^n)^{n-1}$ as $|\cdot|, |||\cdot|||$, respectively, and the spectral norm of a matrix as $\|\cdot\|$. Since Ω is compact, J_e is uniformly continuous on Ω. Thus for all $\epsilon > 0$, there exists $\delta > 0$, independent of x_0, such that $\|J_e(\bar{X}_{n-1}, \hat{X}_{n-2}) - J_e(X_{n-1}, X_{n-2})\| \le \epsilon$ whenever $|||(\bar{X}_{n-1}, \hat{X}_{n-2}) - (X_{n-1}, X_{n-2})||| < \delta$. Define $\alpha_1 := \inf_{x_0 \in K} \lambda_{min}\{J_e(X_{n-1}, X_{n-2})\}$ and $\alpha_2 := \inf_{x_0 \in K} \lambda_{max}\{J_e(X_{n-1}, X_{n-2})\}$, where λ_{min} (λ_{max}, respectively) denotes minimum (maximum) eigenvalue. By compactness of K and continuity of J_e, $0 < \alpha_1 \le \alpha_2 < \infty$. Let $\epsilon = \alpha_1/2$ and δ_0 be the corresponding δ coming from the continuity of J_e. Then it follows that

$$\frac{\alpha_1}{2} I \le J_e(\bar{X}_{n-1}, \hat{X}_{n-2}) \le (\alpha_2 + \frac{\alpha_1}{2}) I$$

whenever $|||(\bar{X}_{n-1}, \hat{X}_{n-2}) - (X_{n-1}, X_{n-2})||| < \delta_0$. Thus the proposition holds with $\gamma_1 = \frac{\alpha_1}{2}, \gamma_2 = \alpha_2 + \frac{\alpha_1}{2}$, and $\delta_1 = \frac{\delta_0}{\sqrt{2n-1}}$. □

Remark 4.2 If one assumes that $[\frac{\partial f}{\partial x}(x_0)]^{-1}$ exists at each $x_0 \in K$, $K \subset \mathbb{R}^n$ compact, then, as long as $\frac{\partial f}{\partial x}$ is continuous, it follows that there exist $\nu_1, \nu_2, 0 < \nu_1 \le \nu_2 < \infty$, such that

$$\nu_1 I \le \frac{\partial f}{\partial x}(x_0)^T \frac{\partial f}{\partial x}(x_0) \le \nu_2 I.$$

Recall that this is important for linking (4.1) and (4.2).

Remark 4.3 Suppose that the system (3.1) satisfies the observability rank condition and that the output y is scalar valued. Then $\tilde{x} = H(x)$ is a local diffeomorphism. In the $\tilde{x}$- coordinates, the system (3.1) is transformed into a local, observer canonical form:

$$\begin{aligned} \tilde{x}_1(k+1) &= \tilde{x}_2(k) \\ &\vdots \\ \tilde{x}_{n-1}(k+1) &= \tilde{x}_n(k) \\ \tilde{x}_n(k+1) &= \phi(\tilde{x}_1(k), \cdots, \tilde{x}_n(k)) \\ y &= \tilde{x}_1(k). \end{aligned} \tag{4.7}$$

A simple computation shows that the linearized observability condition (4.2) is always satisfied for a system in the form (4.7); indeed, $O(k-M,k) \equiv I_n$ for $M = n-1$. This is in marked contrast to the situation analyzed in Proposition 4.1, and underlines the coordinate dependence of the extended Kalman filter in general, and the linearized observability condition (4.2) in particular.

5 Applicability of EKF as an Observer for Nonlinear Systems

In this section we seek to remove Assumption 3.1 by applying the EKF on a convex compact subset of the state space. Before we begin, a few notations are mentioned. Let $\mathcal{O}$ be a (not necessarily small) convex compact subset of R^n, $\sim \mathcal{O}$ the complement of $\mathcal{O}$, and $\epsilon > 0$ a positive constant. Define $d(x, \sim \mathcal{O}) = inf\{|x-y| : y \in \sim \mathcal{O}\}$, and $\mathcal{O}_\epsilon = \{x \in \mathcal{O} : d(x, \sim \mathcal{O}) \geq \epsilon\}$. Since $\mathcal{O}$ is compact, $||A|| := \sup_{x \in \mathcal{O}} ||\frac{\partial f}{\partial x}(x)||$ and $||Dh|| := \sup_{x \in \mathcal{O}} ||\frac{\partial h}{\partial x}(x)||$ are bounded. Let $a = max(1, ||A||)$ and

$$\begin{aligned} b_k &= (1 + ||\bar{P}_0|| ||Dh||^2 ||R^{-1}||^2) a^k \prod_{l=1}^{k} \{1 + ||Dh||^2 ||R^{-1}||^2 \\ &\quad \times [||A||^{2l} ||\bar{P}_0|| + ||N||^2 (||A||^{2(l-1)} + ||A||^{2(l-2)} + \cdots + 1)]\}. \end{aligned}$$

First we consider a sufficient condition to keep the estimates $\bar{x}_k$ and $\hat{x}_k$ near the true state x_k over a finite time period.

Lemma 5.1 *Consider the system (3.1) and its associated EKF (3.3) and (3.4). Suppose that the following conditions hold.*

1. $x_k \in \mathcal{O}_\epsilon$, *for some* $\epsilon > 0$, $0 \leq k \leq M$.
2. $|e_0| = |\bar{x}_0 - x_0| \leq \frac{\delta}{b_M}$ *for some* $0 < \delta \leq \epsilon/2$.

Then for $k = 0, 1, \cdots, M$,

$$|\bar{x}_k - x_k| \leq \delta \quad \text{and} \quad |\hat{x}_k - x_k| \leq \delta.$$

Proof: The proof is by induction. First, by assumption, $|\bar{x}_0 - x_0| \leq \delta$, thus $\bar{x}_0 \in \mathcal{O}_{\epsilon/2}$. Now

$$\begin{aligned} |\hat{x}_0 - x_0| &= |\bar{x}_0 + K_0(h(x_0) - h(\bar{x}_0)) - x_0| \\ &\leq |e_0| + \|K_0\| \cdot |\int_0^1 Dh(\bar{x}_0 + s(x_0 - \bar{x}_0))ds\, e_0| \\ &\leq (1 + \|K_0\|\,\|Dh\|)|e_0|, \end{aligned}$$

where we used the fact that, by convexity, $\bar{x}_0 + s(x_0 - \bar{x}_0) \in \mathcal{O}$ for $0 \leq s \leq 1$. Since $C_0 = \frac{\partial h}{\partial x}(\bar{x}_0)$, $\|K_0\| \leq \|\bar{P}_0\|\,\|Dh\|\,\|R^{-1}\|^2$. Thus

$$|\hat{x}_0 - x_0| \leq (1 + \|\bar{P}_0\|\,\|Dh\|^2\|R^{-1}\|^2)|e_0| \leq \delta.$$

For $k = 1$,

$$\begin{aligned} |\bar{x}_1 - x_1| &= |f(\hat{x}_0) - f(x_0)| \\ &= |\int_0^1 Df(x_0 + s(\hat{x}_0 - x_0))ds(\hat{x}_0 - x_0)| \\ &\leq \|A\|\,|\hat{x}_0 - x_0| \leq \|A\|(1 + \|\bar{P}_0\|\,\|Dh\|^2\|R^{-1}\|^2)|e_0| \leq \delta. \end{aligned}$$

In the same way as for $k = 0$,

$$|\hat{x}_1 - x_1| \leq (1 + \|K_1\|\,\|Dh\|)|e_1|.$$

Using the fact that $\|K_1\| \leq (\|A\|^2\|\bar{P}_0\| + \|N\|^2)\|Dh\|\,\|R^{-1}\|^2$,

$$|\hat{x}_1 - x_1| \leq b_1|e_0| \leq \delta.$$

Now suppose that $|\bar{x}_l - x_l| \leq \delta$, and $|\hat{x}_l - x_l| \leq \delta$ for $0 \leq l \leq k-1$. Then

$$\begin{aligned} |\bar{x}_k - x_k| &= |f(\hat{x}_{k-1}) - f(x_{k-1})| \\ &\leq \|A\|\,|\hat{x}_{k-1} - x_{k-1}| \\ &\leq \|A\|(1 + \|K_{k-1}\|\,\|Dh\|)|\bar{x}_{k-1} - x_{k-1}| \\ &\leq \|A\|(1 + \|K_{k-1}\|\,\|Dh\|)\|A\|(1 + \|K_{k-2}\|\,\|Dh\|) \\ &\quad \cdots \|A\|(1 + \|K_0\|\,\|Dh\|)|e_0|. \end{aligned}$$

Note also that for $1 \leq l \leq k-1$,

$$\begin{aligned} \|P_l\| &= \|[\bar{P}_l^{-1} + C_l^T(RR^T)^{-1}C_l]^{-1}\| \leq \|\bar{P}_l\|, \\ \|\bar{P}_l\| &= \|A_{l-1}P_{l-1}A_{l-1}^T + NN^T\| \\ &\leq \|A\|^2\|\bar{P}_{l-1}\| + \|N\|^2 \\ &\leq \|A\|^{2l}\|\bar{P}_0\| + \|N\|^2(\|A\|^{2(l-1)} + \|A\|^{2(l-2)} + \cdots + 1), \\ \|K_l\| &\leq \|\bar{P}_l\|\,\|Dh\|\,\|R^{-1}\|^2. \end{aligned}$$

Therefore, for $2 \le k \le M$,

$$|\bar{x}_k - x_k| \le b_{k-1}|e_0| \le \delta.$$

Also,

$$\begin{aligned} |\hat{x}_k - x_k| &\le (1 + \|K_k\| \cdot \|Dh\|)|\bar{x}_k - x_k| \\ &\le b_k|e_0| \le \delta. \end{aligned}$$

This completes the proof. □

Since we have conditions which keep the EKF estimates close to the true state, we can now use the results of Theorem 3.2, Proposition 4.1, and Lemma 5.1 to show the convergence of the EKF on a convex compact set without Assumption 3.1. It is only required that the system (3.1) satisfy the observability rank condition on a convex compact set $\mathcal{O}$, and that $[\frac{\partial f}{\partial x}(x)]^{-1}$ exist at each $x \in \mathcal{O}$.

Note that on a compact set $\mathcal{O} \subset R^n$, $|||D^2 f||| := \sup_{x \in \mathcal{O}} |||\frac{\partial^2 f}{\partial x^2}(x)|||$ and $|||D^2 h||| := \sup_{x \in \mathcal{O}} |||\frac{\partial^2 h}{\partial x^2}(x)|||$ are bounded, and Assumption 3.1.4 holds for all $x, y \in \mathcal{O}$. Let $\alpha_1 = \|N\|^2(1 + \|A\|^2 + \|A\|^4 + \cdots + \|A\|^{2(n-2)})$, $\alpha_2 =$ minimum eigenvalue of NN^T, $a = max(1, \|A\|)$, and

$$\begin{aligned} \beta_k &= (1 + \|\bar{P}_0\| \, \|Dh\|^2) a^k \prod_{l=1}^{k} \{1 + \|Dh\|^2 \\ &\quad \times [\|A\|^{2l} \|\bar{P}_0\| + \|N\|^2(\|A\|^{2(l-1)} + \|A\|^{2(l-2)} + \cdots + 1)]\}. \end{aligned}$$

Theorem 5.2 *Consider the system (3.1) and its associated EKF (3.3) and (3.4). Suppose that the system (3.1) satisfies the observability rank condition on a convex compact set $\mathcal{O}$, and that $[\frac{\partial f}{\partial x}(x)]^{-1}$ exists at each $x \in \mathcal{O}$. Let $\delta_1 > 0$ be a constant which satisfies the inequality (4.6) for some $0 < \gamma_1 \le \gamma_2$. Let $p = (\gamma_2 + 1/\alpha_2)$, $q = a^2(\alpha_1 + 1/\gamma_1) + \|N\|^2$. Let $\delta_2 > 0$ be such that $\varphi((pq)^{1/2}\delta_2, |||D^2 f|||, |||D^2 h|||) \le -\gamma$ for some $\gamma > 0$, where φ is defined in Section 3, and let M be the smallest integer which satisfies*

$$\begin{aligned} &[1 + (q\|A\|^2 + \|N\|^2)\|Dh\|^2]\|A\|(1 + q\|Dh\|^2) \\ &\times (1 - \tfrac{\gamma}{p})^{M/2}(pq)^{1/2} < 1. \end{aligned}$$

Suppose further that $x_k \in \mathcal{O}_\epsilon$, $k \ge 0$, for some $\epsilon > 0$, and that $|e_0| \le \frac{\delta}{\beta_{n+M-1}}$ with $\delta = min(\epsilon/2, \delta_1, \delta_2)$. Then we have the following results:

1. $|\bar{x}_k - x_k| \le \delta$ *and* $|\hat{x}_k - x_k| \le \delta \quad \forall k \ge 0$.

2. *The linearized system around $\bar{x}_k$ and $\hat{x}_k$, i.e., $z_{k+1} = \frac{\partial f}{\partial x}(\hat{x}_k)z_k$, $y_k = \frac{\partial h}{\partial x}(\bar{x}_k)z_k$, satisfies the observability condition (4.2) for $k \geq n-1$. Thus there exist $q < \infty, p < \infty$ such that $\|\bar{P}_k\| \leq q$, and $\|P_k^{-1}\| \leq p \;\; \forall k \geq n-1$.*

3. *The error is bounded by δ and converges to zero, i.e., for $k \leq n-1$, $|e_k| \leq \delta$, and for $k > n-1$, $|e_k| \leq min(\delta, (1-\frac{\gamma}{p})^{(k-n+1)/2}(pq)^{1/2}\delta)$.*

Proof: Since the assumptions satisfy the sufficient condition which bounds $\bar{x}_k$ and $\hat{x}_k$ near x_k for $k = 0, \cdots, n+M-1$, it follows that for $k = 0, \cdots, n+M-1$,

$$|\bar{x}_k - x_k| \leq \delta \leq \epsilon/2 \quad \text{and} \quad |\hat{x}_k - x_k| \leq \delta \leq \epsilon/2.$$

Therefore, the EKF (3.3), (3.4) satisfies the observability condition (4.2) with $R = I$; i.e., for $n-1 \leq k \leq n+M-1$,

$$\gamma_1 I \leq \sum_{i=k-n+1}^{k-1} \Phi^T(i, k-n+1)C_i^T C_i \Phi(i, k-n+1) \leq \gamma_2 I. \tag{5.1}$$

Since N is nonsingular and $\mathcal{O}$ is compact, it follows clearly that for $n-1 \leq k \leq n+M-1$,

$$\alpha_1 I \geq \sum_{i=k-M}^{k-1} \Phi(k, i+1)NN^T\Phi^T(k, i+1) \geq \alpha_2 I. \tag{5.2}$$

Hence by Deyst and Price [9], for $n-1 \leq k \leq n+M-1$,

$$\|P_k\| \leq \alpha_1 + 1/\gamma_1 \quad \text{and} \quad \|P_k^{-1}\| \leq p,$$

thereby giving the following bounds for $n-1 \leq k \leq n+M-1$,

$$1/p \leq \|P_k\| \leq \|\bar{P}_k\| \leq q \quad \text{and} \quad 1/q \leq \|\bar{P}_k^{-1}\| \leq \|P_k^{-1}\| \leq p.$$

Since $|e_{n-1}| \leq \delta$, we have

$$\varphi((pq)^{1/2}|e_{n-1}|, |||D^2 f|||, |||D^2 h|||) \leq -\gamma.$$

Accordingly, though the EKF is applied from $k=0$, we have the convergence only after $k = n-1$, i.e.,

$$|e_l| \leq (pq)^{1/2}(1-\frac{\gamma}{p})^{(l-n+1)/2}|e_{n-1}|, \quad l \geq n-1.$$

Now the remaining part is shown by induction. That is,

$$\begin{aligned} |\bar{x}_{n+M} - x_{n+M}| &\leq \|A\|\,|\hat{x}_{n+M-1} - x_{n+M-1}| \\ &\leq \|A\|(1 + \|K_{n+M-1}\|\,\|Dh\|)|e_{n+M-1}| \\ &\leq \|A\|(1 + q\|Dh\|^2)(pq)^{1/2}(1-\frac{\gamma}{p})^{M/2}|e_{n-1}| \leq \delta. \end{aligned}$$

Recall that $R = I$ is used as a design variable. Also,

$$\begin{aligned} |\hat{x}_{n+M} - x_{n+M}| &\leq (1 + \|K_{n+M}\| \, \|Dh\|)|e_{n+M}| \\ &\leq [1 + (q\|A\|^2 + \|N\|^2)\|Dh\|^2]\|A\|(1 + q\|Dh\|^2) \\ &\quad \times (1 - \frac{\gamma}{p})^{M/2}(pq)^{1/2}|e_{n-1}| \leq \delta. \end{aligned}$$

In addition, we have

$$\bar{x}_{n+M} \in \mathcal{O}_{\epsilon/2} \quad \text{and} \quad \hat{x}_{n+M} \in \mathcal{O}_{\epsilon/2}.$$

Thus the conditions (5.1) and (5.2) are also met for $k = n + M$. Hence $\|P_{n+M}\| \leq \alpha_1 + 1/\gamma_1$ and $\|P_{n+M}^{-1}\| \leq p$. Therefore by induction it holds that for $k \geq n + M$,

1. $|\bar{x}_k - x_k| \leq \delta \leq \epsilon/2, \quad |\hat{x}_k - x_k| \leq \delta \leq \epsilon/2.$
2. $\|\bar{P}_k\| \leq q, \quad \|P_k^{-1}\| \leq p.$
3. $|e_k| \leq \delta(pq)^{1/2}(1 - \frac{\gamma}{p})^{(k-n+1)/2}.$ □

Remark 5.3

(a) In order to satisfy the observability condition, it is necessary to keep the estimates $\bar{x}_k$ and $\hat{x}_k$ near x_k for $0 \leq k \leq n+M-1$, thus requiring a good initial guess.

(b) We also need to have a converging period $(n-1 \leq k \leq n+M-1)$ for the EKF in order to build up the observability condition; after this, the recursions proceed automatically.

Remark 5.4 The above results hold wherever the initial guess is close enough to the true state. In other words, we have convergence of the observation error on an open neighborhood of the diagonal of the product space of the true state and the estimate, which includes the origin. This kind of observer is termed *quasi-local* [11]. Note that most results on local observers are only valid on an open neighborhood of the origin [8, 16, 17].

6 Conclusions

Motivated by the fact that the EKF can be used as a parameter estimator, we have analyzed in detail how the EKF works when it is used as an observer for general discrete-time nonlinear systems. First, we gave a new proof of the fact that the Kalman filter is a global observer for linear (discrete-time) time-varying systems. Based on this proof, we were able to show that the EKF is a quasi-local asymptotic observer for discrete-time nonlinear systems. It was shown that, in order to obtain the convergence, it is generally necessary either to have a good initial guess or to have a weak nonlinearity in the sense that $|||D^2f|||$ and $|||D^2h|||$ should be sufficiently small. It was also shown that, in order to *establish* the boundedness of the error covariances, which is necessary for convergence of the error, an observability condition must be imposed on the linearization of the nonlinear system along the estimated trajectory. This observability condition for the linearization was then related to the observability properties of the underlying nonlinear system. With this relation, it was proven that the uniform asymptotic convergence of the observation error is achieved whenever the nonlinear system satisfies an observability rank condition and the states stay within a convex compact domain. This last result also provides a theoretical foundation for this classic, approximate nonlinear filter.

Acknowledgement

The authors gratefully acknowledge the anonymous reviewer's helpful suggestions, which improved the clarity of the paper.

References

[1] B.D.O. Anderson and J.B. Moore. *Optimal Filtering.* Englewood Cliffs, New Jersey: Prentice-Hall, 1979.

[2] B.D.O. Anderson and J.B. Moore. Detectability and stabilizability of time-varying discrete-time linear systems, *SIAM J. Cont. and Optim.*, **19(1)** (1981), 20-32.

[3] M. Aoki. *Optimization of Stochastic Systems.* New York: Academic Press, 1967.

[4] J.S. Baras, A. Bensoussan and M.R. James. Dynamic observers as asymptotic limits of recursive filters: Special cases, *Siam J. Appl. Math.* **48(5)** (Oct. 1988), 1147-1158.

[5] J.S. Baras and P.S. Krishnaprasad. Dynamic observers as asymptotic limits of recursive filters, in *Proc. 21st IEEE Conf. Decision Contr.*, Orlando, FL, pp. 1126-1127, Dec. 1982.

[6] A.E. Bryson, Jr. and Y.C. Ho. *Applied Optimal Control.* New York: Hemisphere, 1975.

[7] S.W. Chan, G.C. Goodwin and K.S. Sin. Convergence properties of the Riccati difference equation in optimal filtering of nonstabilizable systems, *IEEE Trans. Automat. Contr.* **29(2)**, 110-118.

[8] S.-T. Chung and J.W. Grizzle. Sampled-data observer error linearization, *Automatica* **26(6)** (Nov. 1990), 997-1007.

[9] J.J. Deyst, Jr. and C.F. Price. Conditions for asymptotic stability of the discrete minimum-variance linear estimator, *IEEE Trans. Automat. Contr.*, **13(6)**, 702-705.

[10] John M. Fitts. *On the Global Observability of Nonlinear Systems,* Ph.D. Dissertation, Univ. of California, Los Angeles, 1970.

[11] J.W. Grizzle and P.E. Moraal. Newton, Observers and Nonlinear Discre-time Control, in *Proc. 29th IEEE Conf. Decision Contr.*, Hawaii (Dec. 1990), 760-767.

[12] O. Hijab. Asymptotic Baysean estimation of a first order equation with small diffusion, *Annals. Probab.* **12** (1984), 890-902.

[13] M.R. James. Asymptotic nonlinear filtering and large deviations with application to observer design, SRC Report Ph D 88-1, University of Maryland, 1988.

[14] A.H. Jazwinski. *Stochastic Processes and Filtering Theory.* New York: Academic Press, 1970.

[15] R.E. Kalman and J.E. Bertram. Control system analysis and design via the second method of Lyapunov, II discrete-time systems, *Trans. ASME, J. Basic Engrg.*, (June 1960), 394-400.

[16] A.J. Krener and A. Isidori. Linearization by output injection and nonlinear observers, *Syst. & Contr. Lett.* **3** (1983), 47-52.

[17] A.J. Krener and W. Respondeck. Nonlinear observer with linearizable error dynamics, *SIAM J. Control* **23(2)** (1985), 197-216.

[18] H. Kwakernaak and R. Sivan. *Linear Optimal Control Systems.* New York: Wiley-Interscience, 1972.

[19] V. Lakshmikantham and D. Trigiante. *Theory of Difference Equations: Numerical Methods and Applications.* Boston: Academic, 1988.

[20] Frank L. Lewis. *Optimal Control.* New York: John Wiley & Sons, 1986.

[21] Frank L. Lewis. *Optimal Estimation.* New York: John Wiley & Sons, 1986.

[22] L. Ljung. Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems, *IEEE Trans. Automat. Contr.* **AC-24(1)** (Feb. 1979), 36-50.

[23] E.A. Misawa and J.K. Hedrick. Nonlinear observers— a state-of-the-art survey, *ASME Journal of Dynamic System Measurement, and Control* **111** (Sept. 1989), 344-352.

[24] H. Nijmeijer. Observability of autonomous discrete time non-linear systems: a geometric approach, *Int. J. Control* **36(5)** (1982), 867-874.

[25] C. de Souza, M.R. Gevers and G.C. Goodwin. Riccati equations in optimal filtering of nonstabilizable systems, having singular state transition matrices, *IEEE Trans. Automat. Contr.* **31(9)** (1986), 831-838.

[26] Y. Song. *Estimation and Control in Discrete-Time Nonlinear Systems,* Ph.D. Dissertation, Dept. of Aerospace Engineering, University of Michigan, 1992.

[27] Y. Song and J. W. Grizzle. The extended Kalman filter as a local asymptotic observer for nonlinear discrete-time systems, in *Proc. 1992 American Control Conference,* Chicago, IL, (June 1992), 3365-3369.

[28] Special issue on applications of Kalman filtering, *IEEE Trans. Automat. Contr.* **28(3)** (March 1983).

DEPARTMENT OF AERONAUTICAL ENGINEERING, HANKUK AVIATION UNIVERSITY GOYANGSHI, KYUNGGIDO, KOREA

DEPT. OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE, UNIVERSITY OF MICHIGAN, ANN ARBOR, MI 48109, USA

Communicated by Clyde F. Martin