Stochastic Stability of the Extended Kalman Filter: Extension to Intermittent Observations

Mehdi Abedinpour Fallah

Atousa Assadi

A Course Project for Stochastic control and decision theory (ECSE 506)

Department of Electrical Engineering
McGill University
Montreal, Quebec, Canada

April 2012

ABSTRACT

This report is based on two papers entitled as follows,

1. Stochastic stability of the discrete time extended kalman filter [1].
2. Stochastic stability of the extended Kalman filter with intermittent observations [2].

In particular, the estimation error behavior of the discrete-time extended Kalman filter for nonlinear systems is analyzed in a stochastic framework. It is shown that under certain assumptions the estimation error remains bounded. The results are then generalized to systems with intermittent observations. Furthermore, modelling the arrival of the observations as a random process, the effect of two different measurement models on the bounds for the error covariance matrices is studied.

TABLE OF CONTENTS

Chapter 1: Introduction

The Kalman filter, under Gaussian assumption, is the optimal state estimator for linear dynamic systems. Although originally devised for linear systems, nonlinear systems can also be addressed by the Kalman filter through some modifications to it as approximations to the optimal state estimator. The extended Kalman filter (EKF) is one of the most popular estimation techniques that has been largely investigated for state estimation of nonlinear systems [3],[4]. The EKF uses the standard Kalman filter equations to the first-order approximation of the nonlinear model about the last estimate. It is very sensitive to initialization, and filter divergence is inevitable if the arbitrary noise matrices have not been chosen appropriately [1]. The stability results for the usual Kalman filter are studied in [5]-**[7**].

Recently there has been an increased research attention for Networked Control Systems. A common feature of these systems is the presence of significant communication delays and data loss across the network. Therefore, it has become necessary to jointly address the issues of control and communication in these systems [8]-[13]. In contrast to traditional filtering problems, an important feature in networked systems is that the delivery of measurements to the estimator is not always reliable and losses of data may occur. This leads to estimation schemes which are required to handle missing data. For example, Figure 1 shows a structure where the arrival of an observation is modelled by a binary stochastic variable γ_t. If a measurement arrives after the t^{th} step, γ_t is set to 1, if no measurement arrives after the t^{th} step, γ_t is set to 0. The stability and convergence properties of the estimation process have been studied in the case of linear Kalman filtering with intermittent observations in [10], where it is shown that there exists a certain threshold of the packet loss rate above which the state estimation error diverges in the expected sense, i.e., the expected value of the error covariance matrix becomes unbounded as time goes to infinity. The lower and upper bounds of the threshold value is also provided.

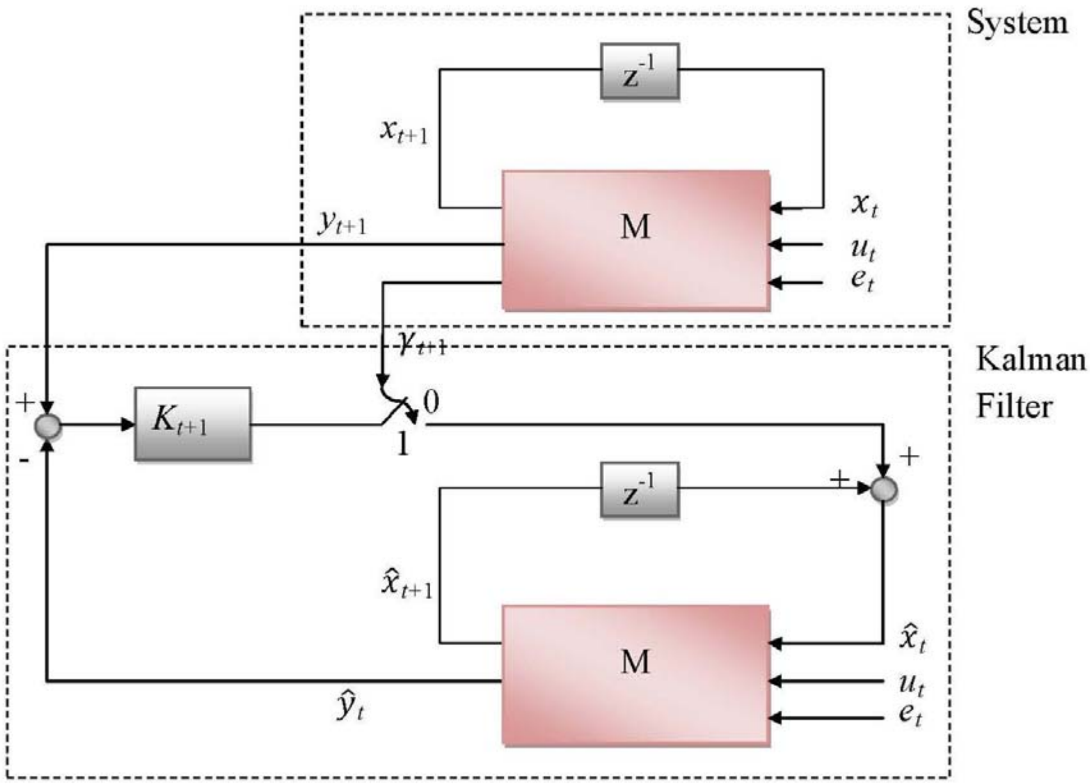

Figure 1. An estimation scheme for Kalman filtering with measurement loss

In this report, the estimation error behavior of the extended Kalman filter is studied. Due to stochasticity, the exponential stability of the nonlinear system is analyzed in the mean square error sense. Furthermore, these results are generalized to systems with intermittent observations and the impact of the data loss process model on the stability of the estimation process is discussed.

The rest of the report is organized as follows. Chapter 2 provides some preliminaries. Next chapter studies the stochastic stability of the discrete-time extended Kalman filter. In chapter 4 these results are generalized for the case of intermittent observations. A critical arrival probability in the case of a Bernoulli process and deterministic bounds for the error covariance matrices in the case of a maximum dropout interval are also derived. Finally, concluding remarks are made in chapter 5.

Chapter 2: Preliminaries

In this section we recall some auxiliary results for the state estimation problem of nonlinear stochastic discrete-time systems as well as their stochastic stability.

2.1 Nonlinear Control System

Consider a nonlinear control system of the type

$$\begin{aligned} x_{t+1} &= f(x_t, u_t, w_t) \\ y_t &= h(x_t, v_t) \end{aligned} \tag{1}$$

where $x_t \in \mathbb{R}^n$ is the state of the system, $u_t \in \mathbb{R}^d$ the control input and $y_t \in \mathbb{R}^m$ the measurements. The stochastic variables $w_t \in \mathbb{R}^s$ and $v_t \in \mathbb{R}^t$ denote the process noise and the measurement noise, respectively. They are both assumed to be uncorrelated white noise processes, and be independent from initial x_0. Note that if we assume linear additive white noise, and therefore, no nonlinear dependency between the state and the system noise, then we can represent the nonlinear control system as

$$\begin{aligned} x_{t+1} &= f(x_t, u_t) + G_t w_t \\ y_t &= h(x_t) + D_t v_t \end{aligned} \tag{2}$$

2.1.1 General Assumptions

In this report, we will refer to the following assumptions as the general assumptions.

1) f and h are continuously differentiable C^1 functions, and the following Jacobian matrices are found for every $t \geq 0$.

$$\begin{aligned} A_t &= \frac{\partial f}{\partial x}(\hat{x}_{t|t}, u_t, 0), \; G_t = \frac{\partial f}{\partial w}(\hat{x}_{t|t}, u_t, 0) \\ C_t &= \frac{\partial h}{\partial x}\left(\hat{x}_{t|t-1}, 0\right), \quad D_t = \frac{\partial h}{\partial v}\left(\hat{x}_{t|t-1}, 0\right) \end{aligned} \tag{3}$$

2) There are positive real numbers $\bar{a}, \bar{g}, \bar{c}, \bar{d}, \underline{q}, \underline{r} > 0$ such that the following bounds on various matrices are fulfilled for every $t \geq 0$.

$$\|A_t\| \le \bar{a} \qquad \|G_t\| \le \bar{g} \qquad \underline{q}\, I_n \le Q_t$$
$$\|C_t\| \le \bar{c} \qquad \|D_t\| \le \bar{d} \qquad \underline{r}\, I_m \le R_t \tag{4}$$

3) There are positive real numbers $\varepsilon_\varphi, \varepsilon_\chi, \delta_\varphi, \delta_\chi > 0$ such that the nonlinear functions φ, χ which are the remaining terms of the Taylor expansions,

$$f(x_t, u_t, w_t) = f(\hat{x}_{t|t}, u_t, 0) + A_t(x_t - \hat{x}_{t|t}) + G_t w_t + \varphi(x_t, \hat{x}_{t|t-1}, u_t, w_t)$$
$$h(x_t, v_t) = h(\hat{x}_{t|t-1}, 0) + C_t(x_t - \hat{x}_{t|t-1}) + D_t v_t + \chi(x_t, \hat{x}_{t|t-1}, v_t) \tag{5}$$

are bounded via

$$\|\varphi(x_t, \hat{x}_{t|t}, u_t, w_t)\|_2 \le \varepsilon_\varphi \|x_t - \hat{x}_{t|t}\|_2^2 \; with \; \forall \|x_t - \hat{x}_{t|t}\|_2 \le \delta_\varphi,$$
$$\|\chi(x_t, \hat{x}_{t|t}, u_t, w_t)\|_2 \le \varepsilon_\chi \|x_t - \hat{x}_{t|t-1}\|_2^2 \; with \; \forall \|x_t - \hat{x}_{t|t-1}\|_2 \le \delta_\chi \tag{6}$$

2.2 Extended Kalman Filter Formulations

There are two common formulations of the discrete-time extended Kalman filter in engineering literature: a two-step recursion with a relinearization between these two steps or a one-step formulation. In this section we review these formulations.

2.2.1 One-step EKF

The one-step extended Kalman filter formulation consists of the following coupled difference equations [11],

$$\hat{x}_{t+1} = f(\hat{x}_t, u_t) + K_t\big(y_t - h(\hat{x}_t)\big) \tag{7}$$

$$P_{t+1} = A_t P_t A_t^T + Q_t - K_t(C_t P_t C_t^T + R_t)\, K_t^T \tag{8}$$

$$A_t = \frac{\partial f}{\partial x}(\hat{x}_t,\, u_t), \;\; C_t = \frac{\partial h}{\partial x}(\hat{x}_t) \tag{9}$$

$$K_t = A_t P_t C_t^T (C_t P_t C_t^T + R_t)^{-1} \tag{10}$$

2.2.2 Two-step EKF

The two-step extended Kalman filter formulation consists of the following coupled difference equations [12],

$$\hat{x}_{t+1|t} = f(\hat{x}_{t|t}, u_t, 0) \tag{11}$$

$$\hat{x}_{t+1|t+1} = \hat{x}_{t+1|t} + K_{t+1}(y_{t+1} - h(\hat{x}_{t+1|t}, 0)) \tag{12}$$

$$P_{t+1|t} = A_t P_{t|t} A_t^T + Q_t \tag{13}$$

$$P_{t+1|t+1} = P_{t+1|t} - K_{t+1} C_{t+1} P_{t+1|t} \tag{14}$$

$$K_{t+1} = P_{t+1|t} C_{t+1}^T \left(C_{t+1} P_{t+1|t} C_{t+1}^T + R_{t+1} \right)^{-1} \tag{15}$$

Remark 2.1

These two formulations may have a different performance and transient behavior, but the convergence properties are the same [11].

2.3 Stochastic Boundedness

Consider the estimation error as

$$e_t = x_t - \hat{x}_t \tag{16}$$

For the analysis of the EKF error dynamics, let us recall two definitions:

Definition 2.1

The stochastic process e_t is said to be exponentially bounded in mean square, if there are real numbers $\beta, \nu > 0$ and $0 < \alpha < 1$ such that

$$\mathbb{E}[\|e_t\|^2] \leq \beta \|e_0\|^2 \alpha^t + \nu \tag{17}$$

holds for every $t \geq 0$.

Definition 2.2

The stochastic process e_t is said to be bounded with probability one, if

$$\sup_{t \geq 0} \|e_t\| < \infty \tag{18}$$

holds with probability one.

Next, some standard results concerning the boundedness of stochastic processes are given.

Proposition 2.1

Consider the nonlinear discrete-time system represented by

$$e_{t+1} = (A_t - K_t C_t) e_t + r_t + s_t \tag{19}$$

with

$$r_t = \varphi(x_t, \hat{x}_t, u_t) - K_t \chi(x_t, \hat{x}_t) \tag{20}$$

$$s_t = G_t w_t - K_t D_t v_t \tag{21}$$

Assume there is a stochastic process $V_t(e_t)$ and real numbers $\bar{v}, \underline{v}, \mu > 0$ and $0 < \alpha \leq 1$ such that

$$\underline{v} \| e_t \|^2 \leq V_t(e_t) \leq \bar{v} \| e_t \|^2 \tag{22}$$

and

$$\mathbb{E}[V_{t+1}(e_{t+1}) | e_t] - V_t(e_t) \leq \mu - \alpha V_t(e_t) \tag{23}$$

are satisfied for every solution of (19). Then the stochastic process e_t is exponentially bounded in mean square, i.e., we have

$$\mathbb{E}[\| e_t \|^2] \leq \frac{\bar{v}}{\underline{v}} \mathbb{E}[\| e_0 \|^2] (1 - \alpha)^t + \frac{\mu}{\underline{v} \alpha} \tag{24}$$

for every $t \geq 0$. Moreover, the stochastic process is bounded with probability one.

$$\underline{v} \| e_t \|^2 \leq V_t(e_t) \leq \bar{v} \| e_t \|^2 \tag{25}$$

Chapter 3: Stochastic Stability of the Extended Kalman Filter

In this chapter the estimation error boundedness is proved for the EKF, if certain conditions are satisfied. Moreover, the role of nonlinear observability in this context is discussed.

3.1 Boundedness of the Estimation Error for the EKF

Theorem 3.1

Consider the nonlinear stochastic systems given by (2) and the one-step EKF formulation as stated in section 2.1.1. Further to general assumptions 2.1.1, let the following assumptions hold:

1. A_t is nonsingular for every $t \geq 0$.
2. There exist real constants $\bar{p}, \underline{p}$>0 such that

$$\underline{p}\, I \leq P_t \leq \bar{p} I \tag{26}$$

If for some ε, $\delta > 0$ the initial estimation error satisfies

$$\|e_0\| \leq \delta \tag{27}$$

and the covariance matrices are bounded via

$$G_t G_t^T \leq \varepsilon^2 I, \qquad D_t D_t^T \leq \varepsilon^2 I \tag{28}$$

Then the estimation error e_t is exponentially bounded in mean square and bounded with probability one.

Proof. Noting that P_t is positive definite, we choose

$$V_t(e_t) = {e_t}^T {P_t}^{-1} e_t \tag{29}$$

The assumption (26) implies that

$$\frac{1}{\bar{p}} \|e_t\|^2 \leq V_t(e_t) \leq \frac{1}{\underline{p}} \|e_t\|^2 \tag{30}$$

Moreover, replacing for e_{t+1}using (19) we can write

$$V_{t+1}(e_{t+1}) = [(A_t - K_t C_t)e_t + r_t + s_t]^T {P_{t+1}}^{-1}[(A_t - K_t C_t)e_t + r_t + s_t] \tag{31}$$

Now, applying Lemma 7.1 we can write the following inequality

$$\begin{aligned} V_{t+1}(e_{t+1}) \le (1-\alpha)V_t(e_t) + {r_t}^T {P_{t+1}}^{-1}[2(A_t - K_t C_t)e_t + r_t]^T \\ + 2{s_t}^T {P_{t+1}}^{-1}[(A_t - K_t C_t)e_t + r_t] + {s_t}^T {P_{t+1}}^{-1} s_t \end{aligned} \tag{32}$$

Observe that by taking the conditional expectation $\mathbb{E}[V_{t+1}(e_{t+1})|e_t]$ and considering the white noise property, we have

$$\mathbb{E}\left[{s_t}^T {P_{t+1}}^{-1}((A_t - K_t C_t)e_t + r_t)|e_t\right] = 0 \tag{33}$$

because only $s_t = G_t w_t - K_t D_t v_t$ depends (linearly) on white noises w_t and v_t. Subsequently, applying Lemma 7.2 and Lemma 7.3 the remaining terms are expressed as the following inequality

$$\mathbb{E}[V_{t+1}(e_{t+1})|e_t] - V_t(e_t) \le -\alpha V_t(e_t) + \kappa_{nonl}\|e_t\|^3 + \kappa_{noise}\varepsilon \tag{34}$$

for $\|e_t\| \le \delta'$. On the other hand, defining

$$\delta = min\left(\frac{\alpha}{2\bar{p}\kappa_{nonl}}, \delta'\right) \tag{35}$$

and using (29) and (30) yields

$$\kappa_{nonl}\|e_t\|\|e_t\|^2 \le \frac{\alpha}{2\bar{p}}\|e_t\|^2 \le \frac{\alpha}{2}V_t(e_t) \tag{36}$$

for $\|e_t\| \le \delta$. Replacing (36) into (34), we get

$$\mathbb{E}[V_{t+1}(e_{t+1})|e_t] - V_t(e_t) \le -\frac{\alpha}{2}V_t(e_t) + \kappa_{noise}\varepsilon \tag{37}$$

for $\|e_t\| \le \delta$. Now we can apply Proposition 2.1 with $\|e_0\| \le \delta$, $\underline{v} = \frac{1}{\bar{p}}$, $\bar{v} = \frac{1}{\underline{p}}$, and $\mu = \kappa_{noise}\varepsilon$. Furthermore, choosing

$$\varepsilon = \frac{\alpha\tilde{\delta}^2}{2\bar{p}\kappa_{noise}} \tag{38}$$

with some $\tilde{\delta} \leq \|e_t\| \leq \delta$ we have

$$\kappa_{noise}\varepsilon \leq \frac{\alpha}{2\bar{p}}\|e_t\|^2 \leq \frac{\alpha}{2}V_t(e_t) \tag{39}$$

Therefore, the inequality

$$\mathbb{E}[V_{t+1}(e_{t+1})|e_t] - V_t(e_t) \leq -\frac{\alpha}{2}V_t(e_t) + \kappa_{noise}\varepsilon \leq 0 \tag{40}$$

guarantees the boundedness of the estimation error. This ends the proof. ■

Remark 3.1

This result states that if the nonlinearity is small then the EKF is stable if initialised close enough to the true initial value. The greater the deviation from linearity the better the initialisation needs to be. It is also noted that the proof presented here provides a technique for calculating conservative bounds for ε and δ. Moreover, simulation studies suggest that ε and δ can be significantly larger than these bounds in some situations.

3.2 The Significance of Nonlinear Observability for the EKF

Consider nonlinear autonomous systems of the following type

$$\begin{aligned} x_{t+1} &= f(x_t) \\ y_t &= h(x_t) + D_t v_t \end{aligned} \tag{41}$$

and recall the following observability rank condition [14].

Definition 3.1

A nonlinear autonomous system of the form (41) satisfies the nonlinear observability rank condition at $x_t \in R^n$, if the nonlinear observability matrix O has full rank n at x_t, where

$$O = \begin{bmatrix} \frac{\partial h}{\partial x}(x_t) \\ \frac{\partial h}{\partial x}(x_{t+1})\frac{\partial f}{\partial x}(x_t) \\ \vdots \\ \frac{\partial h}{\partial x}(x_{t+n-1})\frac{\partial f}{\partial x}(x_{t+n-2}) \dots \frac{\partial f}{\partial x}(x_t) \end{bmatrix} \tag{42}$$

Definition 3.2

The pair $(A_t,\ C_t)$ is said to satisfy the uniform observability condition [7], if there are real numbers $\overline{m}, \underline{m} > 0$ and an integer $l > 0$, such that

$$\underline{m}I \leq M_{t+l,t} = \sum_{j=t}^{t+l}(\Phi_{j,k}{}^T C_j{}^T C_j \Phi_{j,k}) \leq \overline{m}I. \tag{43}$$

With $\Phi_{t,t} = I$ and

$$\Phi_{j,t} = A_{j-1} \dots A_t \tag{44}$$

Theorem 3.2

Consider the nonlinear autonomous systems given by (41) and the one-step EKF formulation as stated in section 2.1.1. Let the following assumptions hold:

1. There is a compact subset $\mathcal{K}$ of R^{n} and the autonomous system satisfies the observability rank condition for every $x_{\mathrm{t}} \in \mathcal{K}$.
2. The nonlinear functions $f,\ h$ are twice continuously differentiable and $\frac{\partial f}{\partial x}(x_t) \neq 0$ holds for every $x_t \in \mathcal{K}$.
3. The sample paths of x_t are bounded with probability one, and $\mathcal{K}$ contains these sample paths as well as all points with distance smaller than δ_κ from these sample paths, where $\delta_\kappa > 0$ is a real number independent of t.
4. There are positive real numbers $\underline{\mathrm{q}}, \underline{\mathrm{r}} > 0$ such that the following bounds are fulfilled for every $\mathrm{t} \geq 0$.

$$\begin{aligned} \underline{\mathrm{q}}\,\mathrm{I_n} &\leq \mathrm{Q_t} \\ \underline{\mathrm{r}}\,\mathrm{I_m} &\leq \mathrm{R_t} \end{aligned} \tag{45}$$

If for some ε, $\delta > 0$ the initial estimation error satisfies

$$\|e_0\| \leq \delta \tag{46}$$

and the covariance matrix is bounded via

$$D_t D_t^T \leq \varepsilon^2 I \tag{47}$$

Then the estimation error e_t is exponentially bounded in mean square and bounded with probability one.

Proof. To prove this theory, we show that the conditions of Theorem 3.1 are implied by the conditions 1–4 of Theorem 3.2 as well as the observability results in Lemma 7.4 and Lemma 7.5. In particular, it can be seen immediately that conditions given by (45) also hold in Theorem 3.1. Moreover, since f, h are twice continuously differentiable for every $x_t \in \mathcal{K}$ and $\mathcal{K}$ is compact, it follows that the Hessian matrices of f_i and h_i are bounded with respect to the spectral norm of matrices, where f_i and h_i are the components of f and h, respectively. Consequently, the constants ε_φ and ε_χ in the following conditions (which hold for Theorem 3.1)

$$\begin{aligned} \|\varphi(x, \hat{x}, u)\| &\leq \varepsilon_\varphi \|x - \hat{x}\|^2 \; with \; \forall \|x - \hat{x}\| \leq \delta_\varphi, \\ \|\chi(x, \hat{x})\| &\leq \varepsilon_\chi \|x - \hat{x}\|^2 \; with \; \forall \|x - \hat{x}\| \leq \delta_\chi \end{aligned} \tag{48}$$

are given by

$$\begin{aligned} \varepsilon_\varphi &= \max_{1 \leq i \leq n} \sup_{x \in \mathcal{K}} \|\text{Hes } f_i(x)\|, \\ \varepsilon_\chi &= \max_{1 \leq i \leq m} \sup_{x \in \mathcal{K}} \|\text{Hes } h_i(x)\| \end{aligned} \tag{49}$$

For the remaining conditions of Theorem 3.1 it is sufficient to ensure these conditions one time-step in advance. We have to show that the boundedness with probability one of e_t and x_t implies the desired bounds on A_t, C_t and P_t. Then we obtain the boundedness of e_{t+1}. Repeating this procedure we get bounds on A_{t+1}, C_{t+1} and P_{t+1} and therefore on e_{t+2}. This strategy can be repeated to get the desired result. Since $n-1$ steps are required to set up the uniform observability condition, we treat the cases $0 \leq t < n$ and $t \geq n$ separately. Firstly, for the case $0 \leq t < n$, using the proof of Lemma 7.1 it follows that $P_{t+1} > 0$ if $P_t > 0$; because the Riccati difference equation determines the evolution for the error covariance, which is positive definite if $Q_t > 0$. Taking the minimum and maximum eigenvalue of P_t and the maximum singular value of A_t, C_t for $0 \leq t < n$ we obtain the bounds on A_t, $C_t$and P_t. Secondly, for the case $t \geq n$ we have to ensure that neither any eigenvalue of P_t converges to zero nor any of the matrices A_t, C_t and $P_t$diverges. The bounds on P_t follow from Lemma 7.4, Lemma 7.5 and applying the boundedness with probability one of e_i for $n \leq i < t$ in the region $\|e_i\| \leq \epsilon_{obs}$. Furthermore, the boundedness for A_t and C_t follows from the continuity of $\frac{\partial f}{\partial x}$ and $\frac{\partial h}{\partial x}$, the compactness of $\mathcal{K}$ and the fact that $\hat{x}_t \in \mathcal{K}$ with probability one, and using $\|x_t - \hat{x}_t\| \leq \delta_\kappa$. Considering these arguments, Theorem 3.1 can be applied by changing (35) with

$$\delta = min\left(\frac{\alpha}{2\bar{p}\kappa_{nonl}}, \delta_\kappa, \varepsilon_{obs}\right) \tag{50}$$

This ends the proof. ■

Remark 3.2

We observe that for autonomous systems the condition on the solution of the Riccati difference equation as imposed in Theorem 3.1, can be reduced to a nonlinear observability rank condition, which can be checked in advance.

Remark 3.3

The proof of Theorem 3.2 can be generalized to nonlinear autonomous systems with process noise, i.e.,

$$\begin{aligned} \tilde{x}_{t+1} &= f(\tilde{x}_t) + w_t \\ \tilde{y}_t &= h(\tilde{x}_t) + v_t \end{aligned} \tag{51}$$

If we assume that the solution $\tilde{x}_t$ for $t \geq 0$ is bounded with probability one sufficiently close to the nominal solution x_t for $t \geq 0$.

One can compare the results of Theorem 3.1and Theorem 3.2 with the stability results for the linear Kalman filter expressed below [5]-**[7]**.

Remark 3.4

Consider the following discrete-time linear dynamical system:

$$x_{t+1} = A_t x_t + G_t w_t$$
$$y_t = {C_t}^T x_t + v_t \tag{52}$$

where, w_t and v_t are independent, zero mean, white processes with $\mathbb{E}[w_t w_t^T] = I$, $\mathbb{E}[v_t v_t^T] = I$. (A nonunit covariance for w_t is absorbed in G_t and a nonunit covariance for v_t is absorbed by scaling y_t and ${C_t}^T$, as long as the covariance is nonsingular). Moreover, it is assumed that $\mathbb{E}[x_0 x_0^T] = P_0$, $\mathbb{E}[x_0] = m$, and x_0, w_t and v_t are independent. Additionally, A_t, G_t, and C_t are assumed to be bounded. Then we have the following results:

1. The pair (A_t, C_t) uniformly detectable is sufficient for the optimal Kalman filter error covariance to be bounded.
2. Furthermore, if the pair (A_t, G_t) is uniformly stabilizable, the Kalman filter is exponentially stable.
3. Uniform detectability of the pair (A_t, C_t) is sufficient for the existence of a bounded sequence K_t such that $(A_t - K_t {C_t}^T)\, x_t$ is exponentially stable.
4. If the pair (A_t, C_t) is uniformly detectable, the (Kalman) filter error covariance and one-step predictor error covariance are bounded.

Chapter 4: Stochastic Stability of the EKF with Intermittent Observation

The estimation error boundedness of the EKF with intermittent observations is studied in this chapter. Moreover, the role of the modeling approach for measurement process will be discussed in the context of the boundedness of the error covariance matrices.

4.1 Boundedness of the Estimation Error

In this section we will study the behavior of estimation error for the extended Kalman filter with intermittent observations.

Theorem 4.1

Suppose that there exists positive real constants $\underline{p}, \overline{p}, \underline{q}, \underline{r}$ such that:

$$\begin{gathered}\underline{p}I \leq P_{t+1|t+1} \leq P_{t+1|t} \leq \overline{p}I \\ \underline{q}I \leq Q_t \text{ and } \underline{r}I \leq R_t\end{gathered} \tag{53}$$

Considering the general assumptions, one can characterize the behavior of estimation error as follows:

$$\begin{gathered}\forall\kappa > 0, \exists\varepsilon > 0 \text{ such that } \mathbb{E}[w_t w_t^T] \leq \varepsilon^2 I \text{ and } \mathbb{E}[v_t v_t^T] \leq \varepsilon^2 I \\ \forall\kappa > 0, \exists\delta > 0 \text{ such that } \mathbb{E}\left[\left\|e_{1|0}\right\|^2\right] \leq \delta\end{gathered} \tag{54}$$

Based on (54) the estimation error $e_{t+1|t}$ is exponentially bounded in mean square and

$$\mathbb{E}\left[\left\|e_{t+1|t}\right\|^2\right] \leq \kappa \tag{55}$$

Proof. Let's assume that $V_t(e_{t|t-1}) = e_{t|t-1}^T P_{t|t-1}^{-1} e_{t|t-1}$. Substituting

$$e_{t+1|t} = A_t(I_t - \gamma_t K_t C_t)e_{t|t-1} + r_t + s_t \tag{56}$$

in $V_t(e_{t|t-1})$ will equal to:

$$\begin{aligned} V_{t+1}\left(e_{t+1|t}\right) = e_{t|t-1}^T (I - \gamma_t K_t C_t)^T A_t^T P_{t+1|t}^{-1} A_t (I - \gamma_t K_t C_t) e_{t|t-1} \\ + r_t^T P_{t+1|t}^{-1}\left[2A_t(I - \gamma_t K_t C_t)e_{t|t-1} + r_t\right] \\ + s_t^T P_{t+1|t}^{-1}[2A_t(I - \gamma_t K_t C_t)e_{t|t-1} + 2r_t + s_t] \end{aligned} \tag{57}$$

The first term in the above equation is estimated using Lemma 7.7. Furthermore, replacing C_t by $\tilde{C}_t = \gamma_t C_t$, one can notice that the effect of γ_t will vanish in the filter, error and Riccati equations. Since $\|\tilde{C}_t\| \leq \bar{c}$ one can follow the proof of Theorem 3.1and we establish:

$$\mathbb{E}\left[V_{t+1}\left(e_{t+1|t}\right)\right] \leq (1-\alpha)\mathbb{E}\left[V_t\left(e_{t|t-1}\right)\right] + \kappa_1 \left\|e_{t|t-1}\right\|_2^3 + \kappa_2 \varepsilon \tag{58}$$

κ_1 and κ_2 in equation (58) are dependent on the bounds $\delta_\chi, \delta_\varphi, \varepsilon_\chi$ and ε_φ from general assumptions and the parameters of the system $\bar{a}, \bar{g}, \bar{c}, \bar{d}, \underline{p}, \bar{p}, \underline{q}, \bar{q}, \underline{r}, \bar{r}$. If ε_χ and $\varepsilon_\phi$tend to zero then κ_1 and κ_2 will tend to zero too. From this point, we may proceed as the proof of (Satz IX.9) in **[**15]. ■

Remark 4.1

Proof of Theorem 4.1 states that as long as there is at least one measurement in a finite set of time steps, one can use arbitrary probabilities for measurements in state prediction. Otherwise, the boundedness of $P_{t+1|t+1}$ and $P_{t+1|t}$ is violated if for infinitely many $t \in \mathbb{N}$, $\mathbb{P}\{\gamma_t = 1\} \neq 1$.

4.2 Boundedness of the Error Covariance Matrices

As we discussed earlier, the error covariance matrices of the extended Kalman filter which meets the nonlinear observability rank condition are bounded for certain measurements. In this section we aim to study the boundedness of the error covariance matrices for the extended Kalman filter with intermittent observations.

For the randomly sampled system with the associated linearized system

$$\begin{aligned} x_{t+1} &= A_t x_t + G_t w_t \\ y_t &= C_t x_t + D_t v_t \end{aligned} \tag{59}$$

we assume that the measurements are only taken at those time steps t at which $\gamma_t = 1$. If the Jacobian bound $\overline{a} < 1$ and for all $t \in \mathbb{N}$, $\gamma_t = 0$, then the bounds of the error covariance matrices are independent from measurement process.

In this section the boundedness of the error covariance matrices for two different modeling approaches of intermittent observations will be studied: Bernoulli process and maximum drop out interval.

4.2.1 Boundedness of Error Covariance Matrices for Bernoulli Process

Considering a Bernoulli process, the intermittent observations are modeled using a binary random variable, γ_t, in each time step. γ_t determines the arrival of measurements after time t. If there is no measurement after time t, γ_t will be set to zero, otherwise γ_t will be 1. The boundedness of the error covariance matrices in this scenario follows Theorem 4.2.

Theorem 4.2

Let's assume that in this system the initial solution of the Riccati equation, $P_{1|0}$, is symmetric positive definite and there are real numbers $\underline{q}, \overline{q}, \underline{r}, \overline{r} > 0$ such that:

$$\underline{q}I \leq Q_t \leq \overline{q}I \text{ and } \underline{r}I \leq R_t \leq \overline{r}I \tag{60}$$

If $m = n$ and for all $t \in \mathbb{N}$, C_t is invertible and is bounded by $\underline{c}^{-1}$, i.e. $\|C_t^{-1}\| \leq \underline{c}^{-1}$. Under these assumptions and if $\gamma > 1 - \overline{a}^{-2}$, then there will exist positive value $\underline{p}, \overline{p}$ such that:

$$\begin{aligned} \underline{p}I \leq P_{t+1|t+1} \leq P_{t+1|t} \\ \mathbb{E}[P_{t+1|t+1}] \leq \mathbb{E}[P_{t+1|t}] \leq \overline{p}I \end{aligned} \tag{61}$$

Proof. Since C_t is invertible and the pair (A_t, C_t) is observable, (A_t, C_t) will be detectable. Then if for all $t \in \mathbb{N}$, $\gamma_t = 1$, the lower bound follows from **[7]**, [Corollary 5.2., p.29]. Substituting $P_{t|t}$ and K_t in $P_{t+1|t}$ and rearranging the terms $P_{t+1|t}$ can be rewritten as follow:

$$P_{t+1|t} = A_t \left(P_{t|t-1} - \gamma_t P_{t|t-1} C_t^T \left(C_t P_{t|t-1} C_t^T + R_t\right)^{-1} C_t P_{t|t-1}\right) A_t^T + Q_t \tag{62}$$

Using Lemma 7.8 and defining $A = C_t P_{t|t-1} C_t^T$ and $B = R_t$ we will obtain:

$$P_{t+1|t} \leq (1-\gamma_t) A_t P_{t|t-1} A_t^T + \gamma_t A_t C_t^{-1} R_t C_t^{-T} A_t^T + Q_t \tag{63}$$

One may also note that:

$$C_t^{-1} R_t C_t^{-T} \leq \overline{r} C_t^{-1} C_t^{-T} \leq \overline{r} \underline{c}^{-2} I_n \tag{64}$$

Using (64) and the fact that $Q_t \leq \overline{q} I$, inequality (63) will change to:

$$P_{t+1|t} \leq (1-\gamma_t) A_t P_{t|t-1} A_t^T + \gamma_t \frac{\overline{r}}{\underline{c}^2} A_t A_t^T + \overline{q} I \tag{65}$$

Using induction, one can show that for all $t \geq 1$ and $p = \max(\|P_{1|0}, \overline{a}^2 \gamma \overline{r} \underline{c}^{-2} + \overline{q}\|)$, $\mathbb{E}[P_{t+1|t}]$ is bounded, i.e.:

$$\mathbb{E}[P_{t+1|t}] \leq p \sum_{j=0}^{t-1} [(1-\gamma)\overline{a}^2]^j I_n \tag{66}$$

If $\gamma > 1 - \overline{a}^{-2}$, the right hand side of the inequality (66) converges. Therefore the upper bound of (61) exists. Stochastically independence of $P_{t|t-1}$ and γ_t can be used to simplify the inequality of (65) in induction.

$$\begin{aligned} \mathbb{E}[P_{2|1}] &\leq \mathbb{E}\left[(1-\gamma_1) A_1 P_{1|0} A_1^T + \gamma_1 \frac{\overline{r}}{\underline{c}^2} A_1 A_1^T + \overline{q} I\right] \\ &\leq (1-\gamma)\overline{a}^2 P_{1|0} + \gamma \frac{\overline{r}}{\underline{c}^2} \overline{a}^2 I_n + \overline{q} I_n \leq (1-\gamma)\overline{a}^2 p I_n + p I_n \end{aligned} \tag{67}$$

(66) holds for the basis of the induction. Now considering that (66) is true for $\mathbb{E}[P_{t|t-1}]$, one can calculate $\mathbb{E}[P_{t+1|t}]$:

$$
\begin{aligned}
\mathbb{E}[P_{t+1|t}] &\leq \mathbb{E}\left[(1-\gamma_t)A_t P_{t|t-1} A_t^T + \gamma_t \frac{\overline{r}}{\underline{c}^2} A_t A_t^T + \overline{q} I\right] \\
&\leq (1-\gamma)\mathbb{E}\left[\overline{a}^2 P_{t|t-1}\right] + \gamma \frac{\overline{r}}{\underline{c}^2} \overline{a}^2 I_n + \overline{q} I_n \\
&\leq (1-\gamma)\overline{a}^2 p \sum_{j=0}^{t-1} \left[(1-\gamma)\overline{a}^2\right]^j I_n + p I_n = p \sum_{j=0}^{t-1} \left[(1-\gamma)\overline{a}^2\right]^j I_n
\end{aligned} \tag{68}
$$

Therefore (66) is held. ■

Remark 4.2

The assumption of invertible C_t is quite restrictive but it is required to write the inequality (65). If C_t is not invertible then the best lower bound for $P_{t|t-1} C_t^T \left(C_t P_{t|t-1} C_t^T + R_t\right)^{-1} C_t P_{t|t-1}$ is going to be 0.

Remark 4.3

One of the drawbacks of using Bernoulli process for modeling the intermittent observation is that there is no guarantee that in any finite set of time steps, there will exists at least one measurement. Therefore if $\overline{a} \geq 1$ and $\gamma \neq 1$, deriving a deterministic bound will be impossible.

4.2.2 Boundedness of Error Covariance Matrices for Maximum Drop Out Interval

To overcome the problem of Bernoulli process, a new modeling approach proposed named maximum dropout interval. Maximum drop out interval assures that in a finite number of time steps at least one measurement will be taken by the system. Assuming that τ_i is the time step in which a measurement occurred, maximum loss rate of N will be defined as $\tau_{i+1} - \tau_i \leq N$. In this case one can extend the concept of uniform observability to this system.

Definition 4.1

The pair (A_t, C_t) is said to satisfy the modified uniform observability condition if positive real numbers $\underline{m}$ and $\overline{m}$ and positive integer value l exist such that the following modified observability gramian is bounded, i.e. $\underline{m}I_n \leq \widetilde{M}_{t+1,t} \leq \overline{m}I_n$.

$$\widetilde{M}_{t+1,t} = \sum_{j=t}^{t+1} \gamma_j \, \Phi_{j,t}^T C_j^T C_j \, \Phi_{j,t} \tag{69}$$

Theorem 4.3

If the linearized system defined by (59) satisfies the modified uniform observability condition and $P_{1|0} > 0$, then if there are real numbers $\underline{q}, \overline{q}, \underline{r}, \overline{r} > 0$ such that

$$\underline{q}I \leq Q_t \leq \overline{q}I \text{ and } \underline{r}I \leq R_t \leq \overline{r}I \tag{70}$$

the solutions of the Riccati equations will have deterministic bounds with $\underline{p}, \overline{p}$:

$$\underline{p}I \leq P_{t+1|t+1} \leq P_{t+1|t} \leq \overline{p}I \tag{71}$$

Proof. Consider the following associated system which is randomly sampled with measurement times τ_t and $\widetilde{w}_{\tau_t} = [w_{\tau_t}, \dots, w_{\tau_{t+1}-1}]^T$:

$$\begin{aligned} x_{\tau_{t+1}} &= \tilde{A}_{\tau_t} x_{\tau_t} + \tilde{G}_{\tau_t} \widetilde{w}_{\tau_t} \\ y_{\tau_t} &= C_{\tau_t} x_{\tau_t} + D_{\tau_t} v_{\tau_t} \end{aligned} \tag{72}$$

In this system $\tilde{A}_{\tau_t}$ and $\tilde{G}_{\tau_t}$ are defined as follow:

$$\tilde{A}_{\tau_t} = \prod_{j=\tau_t}^{\tau_{t+1}-1} A_j \tag{73}$$

$$\tilde{G}_{\tau_t} = \left[\left(\prod_{j=\tau_t}^{\tau_{t+1}-1} A_j \right) G_{\tau_t}, \dots, \tilde{A}_{\tau_{t+1}-1} G_{\tau_{t+1}-2}, G_{\tau_{t+1}-1} \right] \tag{74}$$

And the Cartesian products, $\widetilde{w}_{\tau_t}$, of white noise processes, w_j, are white noise processes. Also from maximum loss rate $\tau_{t+1} - \tau_t \leq N$ will be held. Then given the bounds in general assumption ($\|A_t\| \leq \bar{a}, \|G_t\| \leq \bar{g}$), one can establish $\|\tilde{A}_{\tau_t}\| < \max(1, \bar{a}^N)$ and $\|\tilde{G}_{\tau_t}\| < \max(\bar{g}, \bar{a}^{N-1}\bar{g})$. Furthermore, the uniform observability condition of the linearized system is inferred from the modified uniform observability condition of the linearized system (72). Therefore (72) is detectable. Using **[7]**, [Corollary 5.2., p.29] one can achieve (71). ■

Conclusions

In this report, the estimation error behavior of the extended Kalman filter was analyzed. In particular, it was shown that under certain conditions the estimation error remains bounded in mean square and bounded with probability one. These conditions include the requirements that the initial estimation error and the disturbing noise terms are small enough, the nonlinearities are not discontinuous, and the solution of the Riccati difference equation remains positive definite and bounded. For autonomous systems the condition on the solution of the Riccati difference equation is reduced to a nonlinear observability rank condition.

The results are then generalized to a setting, in which measurements may randomly be lost due to an unreliable communication channel between the sensor and the control unit of a nonlinear control system. One special feature of this result is that it holds for an arbitrary modelling of the intermittent observations. Moreover, two approaches for the modelling of the intermittent measurements were discussed: an i.i.d. Bernoulli process and a random process with a maximum dropout interval. Specifically, in the case of the Bernoulli model, a critical loss probability is derived which ensures the boundedness of the expectation value of the error covariance matrices. Additionally, by generalizing the concept of nonlinear observability to systems with intermittent observations, the existence of the deterministic bounds for the error covariance matrices in the case of a maximum dropout interval was shown.

Bibliography

[1] K. Reif, S. Günther, E. Yaz, and R. Unbehauen, "Stochastic stability of the discrete time extended kalman filter," *IEEE Trans. Autom. Control*, vol. 44, no. 4, pp. 714–728, Apr. 1999.

[2] S. Kluge, K. Reif, and M. Brokate, "Stochastic stability of the extended Kalman filter with intermittent observations," *IEEE Trans. Autom. Control*, 55(2), 514–518, Feb. 2010.

[3] R. G. Brown and P. Y. Hwang, *Introduction to Random Signals and Applied Kalman Filtering*. New York: Wiley, 1992.

[4] A. Gelb, Ed., *Applied Optimal Estimation*. Cambridge, MA: MIT Press, 1984.

[5] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. New York: Academic, 1972.

[6] P. S. Maybeck, *Stochastic models, estimation, and control*, ser. Mathematics in Science and Engineering, 1979, vol. 141.

[7] B. Anderson and J. B. Moore, "Detectability and stabilisability of time-varying discrete-time linear systems," SIAM J. Contr. Optim., vol. 19, pp. 20–32, 1981.

[8] Z. Wang, D. Ho, Y. Liu, and X. Liu, "Robust H_∞control for a class of nonlinear discrete time-delay stochastic systems with missing measurements," *Automatica*, vol. 45, pp. 684–691, 2009.

[9] M. Cloosterman, N. van der Wouw, W. Heemels, and H. Nijmeijer, "Stabilization of networked control systems with large delays and packet dropouts," in *Proc. Amer. Control Conf.*, Seattle, WA, Jun. 11–13, 2008, pp. 4991–4996.

[10] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Autom. Control*, vol. 59, no. 9, pp. 1453–1464, Sep. 2004.

[11] L. Ljung, "Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems," IEEE Trans. Automat. Contr., vol. AC-24, pp. 36–50, 1979.

[12] C. K. Chui and G. Chen, Kalman Filtering with Real-Time Applications. New York: Springer-Verlag, 1987.

[13] M. Huang, S. Dey, "Stability of Kalman filtering with Markovian packet losses", *Automatica*, Vol. 43, pp. 589-607, 2007.

[14] H. Nijmeijer, "Observability of autonomous discrete time nonlinear systems: A geometric approach," *Int. J. Contr.*, vol. 36, pp. 867–873, 1982.

[15] R. Unbehauen, Systemtheorie 2. Munich, Germany: R. Oldenbourg Verlag, 1998.

Appendix

Lemma 7.1

Consider that the conditions of Theorem 3.1 hold. Then there exist a real number $0 < \alpha < 1$ such that for $t \geq 0$, P_t^{-1} satisfies:

$$(A_t - K_t C_t)^T P_{t+1}^{-1} (A_t - K_t C_t) \leq (1 - \alpha) P_t^{-1} \tag{75}$$

Proof. Using definitions of P_{t+1} and K_t following equation is achieved:

$$P_{t+1} = A_t P_t A_t^T + Q_t - A_t P_t C_t^T K_t^T = (A_t - K_t C_t) P_t (A_t - K_t C_t)^T + Q_t + K_t C_t P_t (A_t - K_t C_t)^T \tag{76}$$

with the definition of K_t it can be shown that the result of the following equation is symmetric:

$$A_t^{-1}(A_t - K_t C_t) P_t = P_t - P_t C_t^T (C_t P_t C_t^T + R_t)^{-1} C_t P_t \tag{77}$$

Using matrix inversion lemma and knowing that $P_t^{-1} > 0$ the left hand side of (77) can be rewritten like:

$$A_t^{-1}(A_t - K_t C_t) P_t = (P_t^{-1} + C_t^T R_t^{-1} C_t)^{-1} > 0 \tag{78}$$

Using $P_t > 0$ and $R_t > 0$ and from the definition of Kalman gain, one can obtain:

$$A_t^{-1} K_t C_t = P_t C_t^T (C_t P_t C_t^T + R_t)^{-1} C_t \geq 0 \tag{79}$$

Considering $P_t = P_t^T$ and by combining (78) and (79) one will establish that:

$$K_t C_t P_t (A_t - K_t C_t)^T = A_t [A_t^{-1} K_t C_t][A_t^{-1}(A_t - K_t C_t) P_t]^T A_t^T \geq 0 \tag{80}$$

Substituting (80) into (76) leads to:

$$P_{t+1} \geq (A_t - K_t C_t) P_t (A_t - K_t C_t)^T + Q_t \tag{81}$$

From (78) one can imply that $(A_t - K_t C_t)^{-1}$ exists and:

$$P_{t+1} \geq (A_t - K_t C_t)[P_t + (A_t - K_t C_t)^{-1} Q_t (A_t - K_t C_t)^{-T}](A_t - K_t C_t)^T \tag{82}$$

By the definition of Kalman gain and considering general assumptions and the bounds on P_t following will be held:

$$\|K_t\| \leq \bar{a}\bar{p}\bar{c}\frac{1}{\underline{r}} \tag{83}$$

and

$$P_{t+1} \geq (A_t - K_tC_t)[P_t + \frac{q}{(\overline{a}+\overline{apc}^2/r)^2} I](A_t - K_tC_t)^T \tag{84}$$

Since $P_t \geq \underline{p}I$ and $A_t - K_tC_t$ is nonsingular, one can take the inverse of both sides. By multiplying $(A_t - K_tC_t)^T$ and $(A_t - K_tC_t)$ from left and right respectively and using $\underline{p}I \leq P_t \leq \overline{p}I$ the following will be held:

$$(A_t - K_tC_t)^T P_{t+1}^{-1}(A_t - K_tC_t) \leq \left[1 + \frac{\underline{q}}{\overline{p}(\overline{a}+\overline{apc}^2/r)^2}\right]^{-1} P_t^{-1} \tag{85}$$

which is inequality (75) with

$$1 - \alpha = \frac{1}{1+\frac{\underline{q}}{\overline{p}(\overline{a}+\overline{apc}^2/r)^2}} \tag{86}$$

Lemma 7.2

Considering the conditions of Theorem 3.1, there exist real numbers $\delta', \kappa_{nonl} > 0$ such that for $\|x_t - \hat{x}_t\| \leq \delta'$

$$r_t^T P_t^{-1}[2(A_t - K_tC_t)(x_t - \hat{x}_t) + r_t] \leq \kappa_{nonl}\|x_t - \hat{x}_t\|^3 \tag{87}$$

Proof. Under general assumptions and by considering $\underline{p}I \leq P_t \leq \overline{p}I$ and $C_tP_tC_t^T > 0$ and using the definition of K_t we will have:

$$\|K_t\| \leq \overline{apc}\frac{1}{\underline{r}} \tag{88}$$

Substituting (88) into the equation of r_t one will obtain:

$$\|r_t\| \leq \|\varphi(x_t, \hat{x}_t, u_t)\| + \bar{a}\bar{p}\bar{c}\frac{1}{\underline{r}}\|\chi(x_t, \hat{x}_t)\| \tag{89}$$

Choosing $\delta' = \min(\delta_\varphi, \delta_\chi)$, then for $\|x_t - \hat{x}_t\| \leq \delta'$:

$$\|r_t\| \leq \varepsilon_\varphi \|x_t - \hat{x}_t\|^2 + \bar{a}\bar{p}\bar{c}\frac{1}{\underline{r}}\varepsilon_\chi \|x_t - \hat{x}_t\|^2 = \varepsilon' \|x_t - \hat{x}_t\|^2 \tag{90}$$

in which

$$\varepsilon' = \varepsilon_\varphi + \bar{a}\bar{p}\bar{c}\frac{1}{\underline{r}}\varepsilon_\chi \tag{91}$$

Using (90) and under general assumptions for $\|x_t - \hat{x}_t\|^2 \leq \delta'$

$$r_t^T P_t^{-1}[2(A_t - K_t C_t)(x_t - \hat{x}_t) + r_t] \leq \varepsilon' \|x_t - \hat{x}_t\|^2 \frac{1}{\underline{p}}(2\left(\overline{a} + \overline{apc}\frac{1}{\underline{r}}\overline{c}\right)\|x_t - \hat{x}_t\| + \varepsilon'\delta'\|x_t - \hat{x}_t\|) \tag{92}$$

which is (87) with

$$\kappa_{nonl} = \varepsilon' \frac{1}{\underline{p}}(2\left(\overline{a} + \overline{apc}\frac{1}{\underline{r}}\overline{c}\right) + \varepsilon'\delta') \tag{93}$$

Lemma 7.3

Assuming that the conditions of Theorem 3.1 hold. Then there is a real number $\kappa_{noise} > 0$ which is independent of ε such that

$$\mathbb{E}\{s_t^T P_{t+1}^{-1} s_t\} \leq \kappa_{noise}\varepsilon \tag{94}$$

Proof. v_t and w_t are uncorrelated white noises. Therefore the expectation value of the crossterms that contains both v_t and w_t will become zero. Hence,

$$s_t^T P_{t+1}^{-1} s_t = w_t^T G_t^T P_{t+1}^{-1} G_t w_t + v_t^T D_t^T K_t^T P_{t+1}^{-1} K_t v_t \tag{95}$$

If one assume that the general assumptions hold and $C_t P_t C_t^T > 0$, then

$$\|K_t\| \leq \overline{apc}\frac{1}{\underline{r}} \tag{96}$$

Using (96) in (95) and considering $\underline{p}I \leq P_t \leq \overline{p}I$ yields:

$$s_t^T P_{t+1}^{-1} s_t \leq \frac{1}{\underline{p}} w_t^T G_t^T G_t w_t + \frac{\overline{a}^2 \overline{c}^2 \overline{p}^2}{\underline{p} \underline{r}^2} v_t^T D_t^T D_t v_t \tag{97}$$

Since left-hand side and right-hand side of (97) are scalars, one may take the trace on the right-hand side without changing its value

$$s_t^T P_{t+1}^{-1} s_t \leq \frac{1}{\underline{p}} tr(w_t^T G_t^T G_t w_t) + \frac{\overline{a}^2 \overline{c}^2 \overline{p}^2}{\underline{p} \underline{r}^2} tr(v_t^T D_t^T D_t v_t) \tag{98}$$

Simplifying this and taking the mean value we get:

$$\mathbb{E}\{s_t^T P_{t+1}^{-1} s_t\} \leq \frac{1}{\underline{p}} tr(G_t \mathbb{E}\{w_t w_t^T\} G_t^T) + \frac{\overline{a}^2 \overline{c}^2 \overline{p}^2}{\underline{p} \underline{r}^2} tr(D_t \mathbb{E}\{v_t v_t^T\} D_t^T) \tag{99}$$

In which D_t and G_t are deterministic matrices. Since v_t and w_t are standard vector-valued white noise processes, then $\mathbb{E}\{v_t v_t^T\} = I$ and $\mathbb{E}\{w_t w_t^T\} = I$ and

$$\mathbb{E}\{s_t^T P_{t+1}^{-1} s_t\} \leq \frac{1}{\underline{p}} tr(G_t G_t^T) + \frac{\overline{a}^2 \overline{c}^2 \overline{p}^2}{\underline{p} \underline{r}^2} tr(D_t D_t^T) \tag{100}$$

Using $G_t G_t^T \leq \varepsilon I$ and $D_t D_t^T \leq \varepsilon I$ for some $\delta, \varepsilon > 0$ following equations will held:

$$tr(G_t G_t^T) \leq \varepsilon tr(I) = q\varepsilon \tag{101}$$

$$tr(D_t D_t^T) \leq \varepsilon tr(I) = m\varepsilon \tag{102}$$

q and m are the number of the rows for G_t and D_t respectively. Setting

$$\kappa_{noise} = \frac{q}{\underline{p}} + \frac{\overline{a}^2 \overline{c}^2 \overline{p}^2 m}{\underline{p} \underline{r}^2} \tag{103}$$

yields

$$\mathbb{E}\{s_t^T P_{t+1}^{-1} s_t\} \leq \kappa_{noise} \varepsilon \tag{104}$$

which is the desired inequality (94).

Lemma 7.4

Assume that for $n \geq 0$, P_t is the solution of the following Riccati difference equation.

$$P_{t+1} = A_t P_t A_t^T + Q_t - K_t (C_t P_t C_t^T + R_t) K_t^T \tag{105}$$

Furthermore, consider that following conditions are held:

1) Positive real numbers $\underline{q}, \overline{q}, \underline{r}$ and $\overline{r}$ exist such that $\underline{q} I \leq Q_t \leq \overline{q} I$ and $\underline{r} I \leq Q_t \leq \overline{r} I$
2) The matrices A_t and C_t satisfy the uniform observability condition.
3) The initial P_0 of (105) is positive definite.

Then for every $n \geq 0$, P_t will be bounded, i.e. $\underline{p} I \leq P_t \leq \overline{p} I$.

Lemma 7.5

Let's assume $\mathcal{K} \subset R^n$ is a compact subset. If the following nonlinear system

$$\begin{aligned} x_{t+1} &= f(x_t) \\ y_t &= h(x_t) + D_t v_t \end{aligned} \tag{106}$$

satisfies the nonlinear observability condition for every $x_t \in \mathcal{K}$, then there exists a positive real number $\varepsilon_{obs} > 0$ such that $A_t = \frac{\partial f}{\partial x}(\hat{x}_t)$ and $C_t = \frac{\partial f}{\partial x}(\hat{x}_t)$ satisfy the uniform observability condition, provided that $\|x_t - \hat{x}_t\| \leq \varepsilon_{obs}$.

Lemma 7.6

If for matrices $A, B, C \in \mathbb{R}^{n \times n}$ we assume that B and C are symmetric positive definite and $C - ABA^T > 0$, then $B^{-1} - A^T C^{-1} A > 0$.

Proof. Since $C - ABA^T > 0$, then $(C - ABA^T)^{-1}$ exists and is symmetric positive definite. Knowing $B > 0$ the following will be held:

$$B + BA^T (C - ABA^T)^{-1} AB = (B^{-1} - A^T C^{-1} A)^{-1} > 0 \tag{107}$$

Therefore $(B^{-1} - A^T C^{-1} A)^{-1}$ is symmetric definite as well as its inverse.

Lemma 7.7

Further to general assumption, let's consider that the following assumption holds:

$$\exists \underline{p}, \bar{p} > 0 \; such \; that \; \underline{p} I \leq P_{t+1|t+1} \leq P_{t+1|t} \leq \overline{p} I \tag{108}$$

then there is a $\alpha \in \mathbb{R}$ which value is in interval (0,1) such that:

$$(I_n - \gamma_t K_t C_t)^T A_t^T P_{t+1|t}^{-1} A_t (I_n - \gamma_t K_t C_t) > (1 - \alpha) P_{t|t-1}^{-1} \tag{109}$$

Proof. One can observe that

$$P_{t+1|t} = A_t P_{t|t} A_t^T + Q_t > (1 + \frac{\underline{q}}{2\overline{a}^2 \overline{p}}) A_t P_{t|t} A_t^T \tag{110}$$

From

$$K_{t+1} = P_{t+1|t} C_{t+1}^T (C_{t+1} P_{t+1|t} C_{t+1}^T + R_{t+1})^{-1} \tag{111}$$

And the fact that $\gamma_t = \gamma_t^2$, one obtains

$$P_{t|t} = (I_n - \gamma_t K_t C_t) P_{t|t-1} (I_n - \gamma_t K_t C_t)^T + \gamma_t K_t R_t K_t^T \tag{112}$$

Substituting (112) in (111) and using $R_t > 0$ following inequality will be obtained:

$$P_{t+1|t} > \left(1 + \frac{\underline{q}}{2\overline{a}^2 \overline{p}}\right) A_t (I_n - \gamma_t K_t C_t) A_t^T P_{t|t-1} (I_n - \gamma_t K_t C_t)^T A_t^T \tag{113}$$

Let's define $A = A_t(I_n - \gamma_t K_t C_t)$, $B = 1 + \frac{\underline{q}}{2\overline{a}^2 \overline{p}} P_{t|t-1}$ and $C = P_{t+1|t}$. By choosing α to be equal to $\overline{q}(2\overline{a}^2 \overline{p} + \overline{q})^{-1}$ which is in interval (0,1) and applying

Lemma **7.6**, one can show that (109) holds.

Lemma 7.8

For symmetric positive definite matrices $A, B \in \mathbb{R}^n$ the following inequality holds:

$$(A+B)^{-1} > A^{-1} - A^{-1}BA^{-1} \tag{114}$$

Proof. Following will be established if one applies twice the matrix inversion lemma:

$$\begin{aligned}(A+B)^{-1} &= A^{-1} - A^{-1}[B^{-1} + A^{-1}]^{-1}A^{-1} = A^{-1} - A^{-1}[B - B(A+B)^{-1}B]A^{-1} \\ &= A^{-1} - A^{-1}BA^{-1} + A^{-1}B(A+B)^{-1}BA^{-1} > A^{-1} - A^{-1}BA^{-1}\end{aligned} \tag{115}$$