Stability Analysis of Kalman Filter by Orthonormalized Compressed Measurement

Hyung Keun Lee and Jang Gyu Lee

School of Electrical Engineering and Computer Science and Automatic Control Research Center, Seoul National University

Abstract: In this paper, we propose the concept of orthonormalized compressed measurement for the stability analysis of discrete linear time-varying Kalman filters. Compared with the previous studies that deals with the homogeneous portion of Kalman filters, the proposed Lyapunov method directly deals with the stochastically-driven system. The orthonormalized compressed measurement provides information on the *a priori* state estimate of the Kalman filter at the k-th step that is propagated from the *a posteriori* state estimate at the previous block of time. Since the complex multiple-step propagations of a candidate Lyapunov function by process and measurement noises can be simplified to a one-step Lyapunov propagation by the orthonormalized compressed measurement, a stochastic radius of attraction can be derived that would be practically difficult to obtain by the conventional multiple-step Lyapunov method.

Keywords: Kalman filter, stability, Lyapunov, orthonormalized compressed measurement

I. Introduction

As an optimal estimator of stochastic Linear Time-Varying (LTV) systems in least mean square sense, the Kalman filter (KF) provides the best estimate of system states for the system where the measurement and process noises are wide sense stationary white Gaussian. Because of its simplicity for digital implementation and guaranteed optimality, the KF has been widely applied to various engineering problems. However, engineers who work with the KF frequently experience that its stability is not always guaranteed in spite of its optimality. The stability of the LTV KF has been a subject of active research. Earlier researches on the subject were concerned with the Lyapunov stability theory applied to the homogeneous portion of the filter to establish zero-input stability criteria [1-4]. Additionally, remarks on Bounded-Input Bounde-Output (BIBO) stability are provided for the stability of stochastically-driven system.

Since the stability of stochastically-driven systems is difficult to analyze, a detailed and practical Lyapunov analysis result has not been appeared in the literature. The non-singularity of the observability grammians for a time-varying system plays a central role in showing a strict decrease of a candidate Lyapunov function. It is obtained only by summing the multiple information matrices where each information matrix of a measurement is usually singular. Thus, the selected Lyapunov function should be propagated in multiple steps. In our experience, any attempt to treat input noise terms directly in the multiple-step propagation of a candidate Lyapunov function is intractably complex. To extend the previous works on the KF stability analysis avoiding complex intermediate propagations of a stochastically-driven Lyapunov function, we present the concept of Orthonormalized Compressed Measurement (OCM) which is a special form of stacking measurements.

The procedure to obtain the OCM is summarized in Fig. 1. In Fig. 1, the single stacked measurement vector $Z^{s}_{k/k-N+1}$ is obtained if we merely stack all the measurement $\{z_j\}$ in a specified time interval where $j = k - N + 1, k - N + 2, \cdots, k$. The de-correlated stacked measurement $Z^{\perp}_{k/k-N+1}$ is formed to eliminate the correlation between the stacked measurement $Z^{s}_{k/k-N+1}$ and the estimation error of KF. To reduce the large dimension of $Z^{\perp}_{k/k-N+1}$, a pseudo-inverse is pre-multiplied to $Z^{\perp}_{k/k-N+1}$ resulting in the OCM to obtain $Z^{n}_{k/k-N+1}$.

The OCM provides information on the *a priori* state estimate at the k-th step that is propagated from *a posteriori* state estimate at the ($k-N$)-th step. The proposed OCM is advantageous in analysis because (i) the measurement coefficient matrix is a simple identity matrix, (ii) the noise term is uncorrelated with the *a priori* state estimate at the k-th step that is propagated from *a posteriori* state estimate at the ($k-N$)-th step, (iii) the dimension is equal to the system dimension, and (iv) it is closely related with the observability grammian so as to be utilized for proving the stability of stochastically-driven systems.

This paper is organized as follows. In Section II, we will describe the system model, the filter model, and the error model. Five Lemmas will be introduced. The five Lemmas provide us an idea on how to handle the measurements without any information loss. In Section III, three concepts of equivalent measurements, i.e., the stacked measurement, the de-correlated measurement, and the OCM, are introduced and formulated. In Section IV, several Lemmas are introduced to clarify useful inequalities between several important error covariance matrices. The main theorem regarding the stability condition and the stochastic radius of attraction for LTV KF is sought. In Section V, a concluding remark is given.

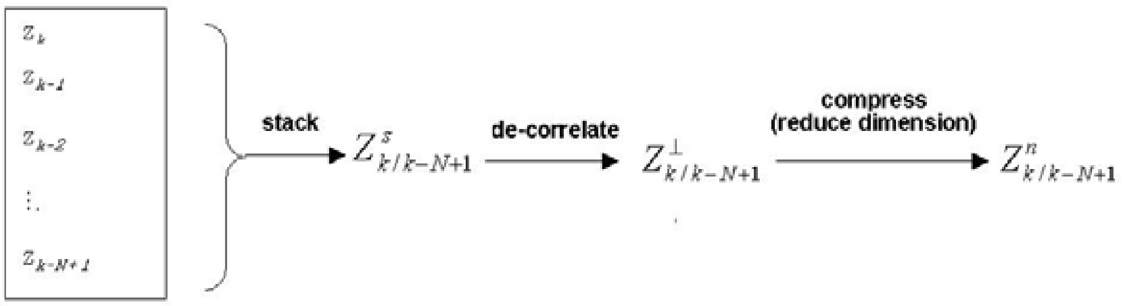

Fig. 1 Equivalent measurements for stability analysis

II. Preservation of Information

We consider the following discrete LTV system driven by white Gaussian noises,

$$x_{k+1} = F_{k+1/k} x_k - G_k w_k\,,\quad w_k \sim \left(O, qI\right),\quad q > 0\,, \tag{1}$$

and the direct measurement,

$$y_k = h_k x_k - v_k\,,\quad v_k \sim \left(O, rI\right),\quad r > 0\,, \tag{2}$$

where

$x_k \in \mathbf{R}^n$: system state vector at the k -th step

$F_{k+1/k} \in \mathbf{R}^{n\times n}$: state transition matrix from the k -th step to the ($k+1$)-th step

$G_k \in \mathbf{R}^{n\times l}$: coefficient matrix for the process noise at the k -th step

$w_k \in \mathbf{R}^l$: process noise at the k -th step

$q \in \mathbf{R}$: strength of each process noise

I : identity matrix of appropriate dimension

O : zero matrix of appropriate dimension

$y_k \in \mathbf{R}^m$: direct measurement vector at the k -th step

$h_k \in \mathbf{R}^{m\times n}$: measurement coefficient matrix at the k -th step

$v_k \in \mathbf{R}^m$: measurement noise at the k -th step

$r \in \mathbf{R}$: strength of each measurement noise

We assume that v_k and w_k are zero-mean and white sequences such that,

$$E[v_j(v_k)] = O\,,\quad \forall j \neq k,\quad j,k = 1, 2, 3, \cdots$$

$$E[w_j(w_k)] = O\,,\quad \forall j \neq k,\quad j,k = 1, 2, 3, \cdots$$

$$E[w_j(v_k)] = O\,,\quad j,k = 1, 2, 3, \cdots. \tag{3}$$

For the discrete-time system model given by Eqs. (1)-(3), the well-known KF performs the following steps.

Time Propagation :

$$\hat{x}_{k+1/k} = F_{k+1/k}\hat{x}_k$$

$$e_{k+1/k} = F_{k+1/k} e_k + G_k w_k$$

$$P_{k+1/k} = F_{k+1/k} P_k F_{k+1/k}{}^T + G_k q G_k{}^T \tag{4}$$

Measurement Update :

$$K_k = P_{k/k-1}h_k^T\left(h_kP_{k/k-1}h_k^T + r_k\right)^{-1}$$

$$\hat{x}_k = \hat{x}_{k/k-1} - K_k z_k^*$$

$$e_k = (I - K_k h_k)e_{k/k-1} - K_k v_k$$

$$P_k = (I - K_k h_k)P_{k/k-1}(I - K_k h_k)^T + K_k r_k K_k^T \quad (5)$$

where

$\hat{x}_{k+1/k}$: optimal estimation of x_{k+1} using measurements upto the k -th step (*a priori* estimation)

$\hat{x}_k = \hat{x}_{k/k}$: optimal estimation of x_k using measurements upto the k -th step (*a posteriori* estimation)

$e_{k+1/k} = \hat{x}_{k+1/k} - x_{k+1}$: *a priori* estimation error

$e_k = \hat{x}_k - x_k$: *a posteriori* estimation error

$P_{k+1/k}$: error covariance matrix of $e_{k+1/k}$

$P_k = P_{k/k}$: error covariance matrix of e_k

$z_k^* = h_k\hat{x}_{k/k-1} - y_k = h_k e_{k/k-1} + v_k$: indirect measurement vector

With the description of models and KF, we introduce five Lemmas that are utilized to prove the main theorem of stability. Given *a priori* estimation and a measurement of a Gaussian random variable, Lemma 2-1 states that any pre-multiplication of a non-singular square matrix with the original measurement will not alter the *a posteriori* estimation. Lemma 2-2 states that any measurement vector whose noise is correlated with the *a priori* estimation error can be represented by an equivalent measurement that is uncorrelated with the *a priori* estimation error. Lemma 2-3 states that the pre-multiplication of the Kalman gain with the original measurement does not alter the *a posteriori* estimation and makes the measurement coefficient matrix as the identity matrix when the dimension of the original measurement vector is larger than the system dimension. Lemma 2-4 states that even though the system states change by time-propagation, the measurement information with respect to the system states can be maintained by considering the exact effects of time-propagation. Proofs of most Lemmas are obvious. Only Lemma 2-5 will be proven. Throughout the Lemmas, the symbol $Est[x]$ is used to denote an estimation of x where the optimality is not guaranteed and the symbol $E(x)$ is used to denote the probabilistic expectation of x.

Lemma 2-1: Equivalence of Measurements [5]

Suppose that we are given *a priori* estimation $\hat{x}^- \in \mathbf{R}^{n\times 1}$ of a Gaussian random variable $x \in \mathbf{R}^{n\times 1}$ with $e^- \in \mathbf{R}^{n\times 1}$, $M \in \mathbf{R}^{n\times n}$, $z \in \mathbf{R}^{m\times 1}$, $H \in \mathbf{R}^{m\times n}$, $R \in \mathbf{R}^{m\times m}$, and $S \in \mathbf{R}^{m\times n}$ such that

$$\hat{x}^- = Est[x]$$

$$M = E[(e^-)(e^-)^T] \tag{6}$$

where

$$e^- := \hat{x}^- - x$$

$$z = He^- + v, \quad v \sim \left(O_{m\times 1}, R\right)$$

$$S = E[v(e^-)^T]. \tag{7}$$

Then, no information is lost in obtaining the *a posteriori* estimation $\hat{x}^+$ of x by applying a non-singular transformation $C \in \mathbf{R}^{\mathbf{m}\times\mathbf{m}}$ to z if we properly consider the changes in H, R, and S, i.e., z is equivalent to $\bar{z}$ where

$$\bar{z} := Cz = \bar{H}e^- + \bar{v}, \quad \bar{v} \sim \left(O_{m\times 1}, \bar{R}\right)$$

$$\bar{H} := CH$$

$$\bar{R} := E[\bar{v}\bar{v}^T] = CRC^T$$

$$\bar{S} := E[\bar{v}(e^-)^T] = CS. \tag{8}$$

Lemma 2-2: De-Correlation of a Measurement [2, 5]

Suppose that we are given *a priori* estimation $\hat{x}^- \in \mathbf{R}^{n\times 1}$ of a Gaussian random variable $x \in \mathbf{R}^{n\times 1}$ with $e^- \in \mathbf{R}^{n\times 1}$, $M \in \mathbf{R}^{n\times n}$, $z \in \mathbf{R}^{m\times 1}$, $H \in \mathbf{R}^{m\times n}$, $R \in \mathbf{R}^{m\times m}$, and $S \in \mathbf{R}^{m\times n}$ satisfying Eqs. (6) and (7). In addition, the following condition is satisfied.

$$R - SM^{-1}S^T > O. \tag{9}$$

Then, no information is lost in obtaining the *a posteriori* estimation $\hat{x}^+$ of x by removing cross-correlation between v of z from e^- if we properly consider the changes in H and R, i.e., z is equivalent to $\bar{z}$ where

$$\bar{z} := \bar{H}e^{-} + \bar{v}\,,\ \ \bar{v} \sim \left(O_{n\times1}, \bar{R}\right)$$

$$\bar{H} := H + SM^{-1}$$

$$\bar{R} := R - SM^{-1}S^{T}$$

$$E[\bar{v}e^{T}] = O_{m\times n}\,. \tag{10}$$

Lemma 2-3: Preservation of Information in Dimension Reduction [2, 5]

Suppose that we are given *a priori* estimation $\hat{x}^{-} \in \mathbf{R}^{n\times1}$ of a Gaussian random variable $x \in \mathbf{R}^{n\times1}$ with $e^{-} \in \mathbf{R}^{n\times1}$, $M \in \mathbf{R}^{n\times n}$, $z \in \mathbf{R}^{m\times1}$, $H \in \mathbf{R}^{m\times n}$, $R \in \mathbf{R}^{m\times m}$, and $S \in \mathbf{R}^{m\times n}$ satisfying Eqs. (6) and (7). In addition, the following conditions are satisfied.

$$m \ge n\,,\ \ H^{T}H > O\,,\ \ S = O \tag{11}$$

Then, no information is lost in obtaining the *a posteriori* estimation $\hat{x}^{+}$ of x by applying the unique transformation $C^{*} \in \mathbf{R}^{n\times m}$ to z if we properly consider the changes in H, R, and S, i.e., z is equivalent to $\bar{z}$ where

$$\bar{z} := C^{*}z = \bar{H}e^{-} + \bar{v}\,,\ \ \bar{v} \sim \left(O_{n\times1}, \bar{R}\right)$$

$$C^{*} := (H^{T}R^{-1}H)^{-1}H^{T}R^{-1}\,,$$

$$\bar{H} := I_{n\times n}$$

$$\bar{R} := E[vv^{T}] = (H^{T}R^{-1}H)^{-1}. \tag{12}$$

Lemma 2-4: Time-Propagated Measurement [2, 5]

Suppose that we are given *a priori* estimation $\hat{x}_{k}^{-}$ of a Gaussian random variable x_{k} with an estimation error e_{k}^{-} at the k-th step with M_{k}, z_{k}, R_{k}, and S_{k} such that

$$\hat{x}_{k}^{-} = Est[x_{k}]$$

$$z_{k} = H_{k}e_{k}^{-} + v_{k}\,,\ \ v_{k} \sim \left(O, R_{k}\right)$$

$$e_k^- := \hat{x}_k^- - x_k$$

$$S_k = E[v_k (e_k^-)^T] \tag{13}$$

where

$$R_k - S_k M_k^{-1} S_k^T > O$$

$$M_k := E[e_k^- (e_k^-)^T]. \tag{14}$$

In addition, suppose that the system states are propagated in time by Eq. (1) and its estimation is changed by the following equation.

$$\hat{x}_{k+1}^- = F_{k+1/k} \hat{x}_k^- \tag{15}$$

where

$$w_k \sim (O, q_k), \quad E[w_k v_k^T] = O, \quad E[w_k e_k^-] = O \tag{16}$$

Then, no information is lost in using z_k to estimate x_{k+1} by the following interpretation.

$$\bar{z}_{k+1} := z_k = \bar{H}_{k+1} e_{k+1}^- + \bar{v}_{k+1} \tag{17}$$

where

$$e_{k+1}^- := \hat{x}_{k+1}^- - x_{k+1}$$

$$\bar{H}_{k+1} := H_k (F_{k+1/k})^{-1}$$

$$\bar{R}_{k+1} := E[\bar{v}_{k+1} (\bar{v}_{k+1})^T] = R_k + \bar{H}_{k+1} G_k q_k G_k{}^T (\bar{H}_{k+1})^T$$

$$\bar{S}_{k+1} := E[\bar{v}_{k+1} (e_{k+1}^-)^T] = S_k (F_{k+1/k})^T - \bar{H}_{k+1} G_k q_k G_k{}^T . \tag{18}$$

<Proof>

By Eqs. (1) and (15), the time-propagation of estimation error from e_k^- to e_{k+1}^- is derived as

$$e_{k+1}^- = F_{k+1/k} e_k^- + G_k w_k . \tag{19}$$

According to Eqs. (13) and (19), the given z_k satisfies the following relationship with e_{k+1}^-.

$$z_k = H_k (F_{k+1/k})^{-1} e_{k+1}^- + v_k - H_k (F_{k+1/k})^{-1} G_k w_k \tag{20}$$

By defining $\bar{v}_{k+1}$ as

$$\bar{v}_{k+1} := v_k - H_k (F_{k+1/k})^{-1} G_k w_k \,, \tag{21}$$

we obtain the result. In addition to the interpretation of z_k by $\bar{z}_{k+1}$ in Eqs. (17) and (18), various equivalent interpretations can be obtained by applying Lemmas 2-1, 2-2, and 2-3.

III. Equivalent Batch Measurements

Given $\hat{x}_{k-N}$, P_{k-N}, $\{w_j\}_{k-N \le j<k}$, $\{z_j\}_{k-N<j\le k}$, and $\{v_j\}_{k-N+1<j\le k}$ for a discrete-time system model represented by Eqs. (1)-(3), the most common method for obtaining the optimal *a posteriori* estimate $\hat{x}_k$ of x_k is to implement, step by step, the recursive KF algorithm by Eqs. (4) and (5). Given the Lemmas regarding the preservation of information, it is possible to obtain the same optimal *a posteriori* estimate $\hat{x}_k$ of x_k in a different way given $\hat{x}_{k-N/k-N}$, $P_{k-N/k-N}$, $\{w_j\}_{k-N\le j<k}$, $\{z_j\}_{k-N<j\le k}$, and $\{v_j\}_{k-N+1<j\le k}$.

Suppose that the *a priori* estimation of $\{x_j\}_{k-N<j\le k}$ was performed by only the following multiple-step time-propagations, without utilizing the measurements $\{z_j\}_{k-N<j\le k}$.

$$\hat{x}_{k/k-N} = F_{k/k-N}\hat{x}_{k-N/k-N}$$

$$\begin{aligned} e_{k/k-N} &= F_{k/k-N}e_{k-N/k-N} + W_{k-1/k-N} \\ &= F_{k/k-N}e_{k-N/k-N} + (G^o_{k-1/k-N})^T w^o_{k-1/k-N} \end{aligned}$$

$$M_k = F_{k/k-N}P_{k-N}F_{k/k-N}{}^T + \Xi_{k-1/k-N} \tag{23}$$

where

$$W_{k-1/k-N} := \sum_{j=k-N}^{k-1} F_{k/j+1}G_j w_j = (G^o_{k-1/k-N})^T w^o_{k-1/k-N}$$

$$W_{k-1/k-N} \sim \left(O, \Xi_{k-1/k-N}\right)$$

$$\Xi_{k-1/k-N} := G^o_{k-1/k-N} Q^o_{k-1/k-N} (G^o_{k-1/k-N})^T$$

$$G^o_{k-1/k-N} := \left[G_{k-1} \;\vdots\; F_{k/k-1}G_{k-2} \;\vdots\; F_{k/k-2}G_{k-3} \;\vdots\; \cdots \;\vdots\; F_{k/k-N+1}G_{k-N} \right]$$

$$w^o_{k-1/k-N} := \begin{bmatrix} w_{k-1} \\ \cdots \\ w_{k-2} \\ \cdots \\ w_{k-3} \\ \cdots \\ \vdots \\ \cdots \\ w_{k-N} \end{bmatrix},$$

$$w^o_{k-1/k-N} \sim \left(O, Q^o_{k-1/k-N}\right)$$

$$Q^o_{k-1/k-N} = qI > 0\,. \tag{24}$$

Note that $\Xi_{k-1/k-N}$ in Eq. (24) is a controllability grammian matrix by the process noise from the ($k-N$)-th step to the ($k-1$)-th step. Suppose that we are also given the measurements from the ($k-N+1$)-th step to the k-th step expressed by a stacked measurement vector $Z^s_{k/k-N+1}$ in addition to the *a priori* estimation $\hat{x}_{k/k-N}$ of x_{k-N}.

$$Z^s_{k/k-N+1} := \begin{bmatrix} z_k \\ \cdots \\ z_{k-1} \\ \cdots \\ z_{k-2} \\ \cdots \\ \vdots \\ \cdots \\ z_{k-N+1} \end{bmatrix}$$

$$z_j := h_j \hat{x}_{j/k-N+1} - y_j = h_j e_{j/k-N+1} + v_j \tag{25}$$

By the first three theorems given in Section II, it is possible to find an $n \times Nm$ optimal transformation matrix $C^*_{k/k-N+1}$ from an $Nm \times 1$ stacked measurement vector $Z^s_{k/k-N+1}$ to an $n \times 1$ single equivalent measurement vector $Z^n_{k/k-N+1}$ without any information loss,

$$Z^n_{k/k-N+1} = C^*_{k/k-N+1}\, Z^s_{k/k-N+1}\,. \tag{26}$$

If we exactly trace the changes of the measurement coefficient matrix, error covariance matrix, and cross-correlation matrix of $Z^n_{k/k-N+1}$, it is possible to obtain the optimal *a posteriori* estimate $\hat{x}_k$ of x_k by a batch form as

$$\hat{x}_k = \hat{x}_{k/k-N} - K^n_{k/k-N+1} Z^n_{k/k-N+1} = \hat{x}_{k/k-N} - K^n_{k/k-N+1} C^*_{k/k-N+1} Z^s_{k/k-N+1} \tag{27}$$

where $K^n_{k/k-N+1}$ denotes the optimal gain related with $Z^n_{k/k-N+1}$. For this purpose, three kinds of equivalent batch measurements are presented. They are classified as stacked measurement, de-correlated measurement, and OCM according to their corresponding characteristics. By combining any one of the presented equivalent measurements with the *a priori* estimation $\hat{x}_{k/k-N}$, we can obtain the optimal *a posteriori* estimation $\hat{x}_k$ of x_k periodically at intermittent times $k = jN$, $j = 0, 1, 2, 3, \cdots$ according to the Lemmas in Section II.

To notationally discriminate between various vectors and matrices that are related with several concepts of equivalent measurement, the superscript s will be used for a stacked measurement, the superscript $\perp$ will be used for a de-correlated stacked measurement, and the superscript n will be used for an OCM. To align all the measurements obtained from the ($k-N+1$)-th step to the k-th step with respect to $e_{j/k-N}$, the following relationship will be used,

$$z_j = h_j F_{j/k} e_{k/k-N} - \sum_{a=j}^{k-1} h_j F_{j/a+1} G_a w_a + v_j \,. \tag{28}$$

3.1 Stacked Measurement

A single stacked measurement vector $Z^s_{k/k-N+1}$ is obtained if we merely stack each measurement obtained from the ($k-N+1$)-th step to the k-th step. In this case, the stacked measurement $Z^s_{k/k-N+1}$ can be represented by the following single vector equation,

$$\begin{aligned} Z^s_{k/k-N+1} &= H^s_{k/k-N+1} e_{k/k-N} + v^o_{k/k-N+1} - H^o_{k-1/k-N} w^o_{k-1/k-N} \\ &= H^s_{k/k-N+1} e_{k/k-N} + v^s_{k/k-N+1} \end{aligned} \tag{29}$$

where

$$Z^s_{k/k-N+1} := \begin{bmatrix} z_k \\ \hdashline z_{k-1} \\ \hdashline z_{k-2} \\ \hdashline \vdots \\ \hdashline z_{k-N+1} \end{bmatrix}, \qquad H^s_{k/k-N+1} := \begin{bmatrix} h_k \\ \hdashline h_{k-1} F_{k-1/k} \\ \hdashline h_{k-2} F_{k-2/k} \\ \hdashline \vdots \\ \hdashline h_{k-N+1} F_{k-N+1/k} \end{bmatrix}$$

$$v^o_{k/k-N+1} := \begin{bmatrix} v_k \\ \hdashline v_{k-1} \\ \hdashline v_{k-2} \\ \hdashline \vdots \\ \hdashline v_{k-N+1} \end{bmatrix}, \qquad v^o_{k/k-N+1} \sim \left(O, R^o_{k/k-N+1}\right)$$

$$R^o_{k/k-N+1} = \begin{bmatrix} rI_{m\times m} & O & \cdots & O \\ O & rI_{m\times m} & & O \\ \vdots & & \ddots & \vdots \\ O & O & \cdots & rI_{m\times m} \end{bmatrix} = rI_{mN\times mN}$$

$$H^o_{k/k-N+1} := \begin{bmatrix} O & O & \cdots & O & O \\ h_{k-1}F_{k-1/k}G_{k-1} & O & \cdots & O & O \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ h_{k-N+2}F_{k-N+2/k}G_{k-1} & h_{k-N+2}F_{k-N+2/k-1}G_{k-2} & \cdots & O & O \\ h_{k-N+1}F_{k-N+1/k}G_{k-1} & h_{k-N+1}F_{k-N+1/k-1}G_{k-2} & \cdots & h_{k-N+1}F_{k-N+1/k-N+2}G_{k-N+1} & O \end{bmatrix}$$

$$v^s_{k/k-N+1} := v^o_{k/k-N+1} - H^o_{k-1/k-N} w^o_{k-1/k-N}$$

$$v^s_{k/k-N+1} \sim \left(O, R^s_{k/k-N+1}\right)$$

$$\begin{aligned} R^s_{k/k-N+1} &:= E\left[\left(v^s_{k/k-N+1}\right)\left(v^s_{k/k-N+1}\right)^T\right] \\ &= R^o_{k/k-N+1} + H^o_{k-1/k-N} Q^o_{k-1/k-N} (H^o_{k-1/k-N})^T \end{aligned}$$

$$\begin{aligned} S^s_{k/k-N+1} &:= E\left[\left(v^s_{k/k-N+1}\right)\left(e_{k/k-N}\right)^T\right] \\ &= -H^o_{k-1/k-N} Q^o_{k-1/k-N} (G^o_{k-1/k-N})^T . \end{aligned} \tag{30}$$

As shown in Eqs. (29) and (30), the measurement noise $v^s_{k/k-N+1}$ of the stacked measurement $Z^s_{k/k-N+1}$ is composed of two terms, i.e., the noise term $v^o_{k/k-N+1}$, which is uncorrelated with the *a priori* estimation error $e_{k/k-N}$, and the noise term $H^o_{k-1/k-N} w^o_{k-1/k-N}$, which is correlated with the *a priori* estimation error $e_{k/k-N}$. Since a sequence of process noise $\{w_j\}_{k-N\le j<k}$ generates a correlation between the stacked measurement noise $v^s_{k/k-N+1}$ and the *a priori* estimation error $e_{k/k-N}$ of $\hat{x}_{k/k-N} := E[x_k \mid y_1, y_2, \cdots, y_{k-N}]$ as shown in Eq. (30), the optimal *a posteriori* estimation $\hat{x}_k := E[x_k \mid y_1, y_2, \cdots, y_k]$ of x_k using all the measurement up to the k-th step can be obtained by a somewhat complicated computation:

$$K^s_{k/k-N+1} = \left[M_k (H^s_{k/k-N+1})^T + (S^s_{k/k-N+1})^T\right] \cdot \begin{bmatrix} (H^s_{k/k-N+1}) M_k (H^s_{k/k-N+1})^T \\ + (S^s_{k/k-N+1}) + (S^s_{k/k-N+1})^T + (R^s_{k/k-N+1}) \end{bmatrix}^{-1}$$

$$P_k = M_k - K^s_{k/k-N+1}\left[(H^s_{k/k-N+1}) M_k + (S^s_{k/k-N+1})\right]$$

$$\hat{x}_k = \hat{x}_{k/k-N} - K^s_{k/k-N+1} Z^s_{k/k-N+1}\,. \tag{31}$$

By performing above computation, the *a posteriori* estimation error e_k of $\hat{x}_k$ satisfies

$$e_k = \left[I - K^s_{k/k-N+1} H^s_{k/k-N+1}\right] e_{k/k-N} - K^s_{k/k-N+1} v^s_{k/k-N+1}\,. \tag{32}$$

3.2 De-Correlated Stacked Measurement

As shown in Eqs. (29) and (30), the correlation between the stacked measurement noise $v^s_{k/k-N+1}$ and the *a priori* estimation error $e_{k/k-N}$ makes the optimal measurement update somewhat complicated. To circumvent this complication, an equivalent measurement whose error is uncorrelated with $e_{k/k-N}$ is desirable. The de-correlated stacked measurement $Z^{\perp}_{k/k-N+1}$ is formed as

$$Z^{\perp}_{k/k-N+1} = Z^s_{k/k-N+1} = H^{\perp}_{k/k-N+1} e_{k/k-N} + v^{\perp}_{k/k-N+1} \tag{33}$$

where

$$H^{\perp}_{k/k-N+1} := H^s_{k/k-N+1} + S^s_{k/k-N+1} (M_k)^{-1} = \begin{bmatrix} h_k \\ \hdashline h_{k-1} F^{\perp}_{k-1/k} \\ \hdashline h_{k-2} F^{\perp}_{k-2/k} \\ \hdashline \vdots \\ \hdashline h_{k-N} F^{\perp}_{k-N+1/k} \end{bmatrix}$$

$$v^{\perp}_{k/k-N+1} = -S^s_{k/k-N+1} (M_k)^{-1} e_{k/k-N} + v^s_{k/k-N+1}$$

$$\begin{aligned} S^{\perp}_{k/k-N+1} &:= E\left[\left(v^{\perp}_{k/k-N+1}\right)\left(e_{k/k-N}\right)^T\right] \\ &= -S^s_{k/k-N+1} + S^s_{k/k-N+1} = O \end{aligned}$$

$$\begin{aligned} R^{\perp}_{k/k-N+1} &:= E\left[\left(v^{\perp}_{k/k-N+1}\right)\left(v^{\perp}_{k/k-N+1}\right)^T\right] \\ &= R^s_{k/k-N+1} - S^s_{k/k-N+1} M_k^{-1} \left(S^s_{k/k-N+1}{}^T\right) \end{aligned}$$

$$F^{\perp}_{l/k} := F_{l/k}[I - \Xi_{k-1/l} (M_k)^{-1}], \quad l < k$$

$$\Xi_{k/l} := \sum_{j=l}^{k} F_{k+1/j+1} G_j q_j \left(F_{k+1/j+1} G_j\right). \tag{34}$$

Since the measurement noise $v_{k/k-N+1}^{\perp}$ of the de-correlated measurement $Z_{k/k-N+1}^{\perp}$ is uncorrelated with $e_{k/k-N}$ as shown in Eq. (34), it will simplify the form of measurement update equations of Eqs. (31) and (32) as follows.

$$\begin{aligned}
K_{k/k-N+1}^{\perp} &= M_k (H_{k/k-N+1}^{\perp})^T \left[(H_{k/k-N+1}^{\perp}) M_k (H_{k/k-N+1}^{\perp})^T + (R_{k/k-N+1}^{\perp})\right]^{-1} \\
P_k &= M_k - K_{k/k-N+1}^{\perp} H_{k/k-N+1}^{\perp} M_k \\
\hat{x}_k &= \hat{x}_{k/k-N} - K_{k/k-N+1}^{\perp} Z_{k/k-N+1}^{\perp}.
\end{aligned} \tag{35}$$

By performing the computation of Eq. (35), the *a posteriori* estimation error e_k of $\hat{x}_k$ satisfies

$$e_{k/k} = \left[I - K_{k/k-N+1}^{\perp} H_{k/k-N+1}^{\perp}\right] e_{k/k-N} - K_{k/k-N+1}^{\perp} v_{k/k-N+1}^{\perp}. \tag{36}$$

3.3 Orthonormalized Compressed Measurement

Though the de-correlated stacked measurement $Z_{k/k-N+1}^{\perp}$ obtained in Subsection 3.2 simplifies the optimal measurement update, its dimension Nm is rather large and the corresponding measurement coefficient matrix $H_{k/k-N+1}^{\perp}$ in Eq. (34) is still complicated. Under the assumption that the modified observability grammian $(H_{k/k-N+1}^{\perp})^T H_{k/k-N+1}^{\perp}$ is non-singular, a pseudo-inverse can be pre-multiplied to $Z_{k/k-N+1}^{\perp}$ for further simplification. Then, by Lemma 2-3, this pre-multiplication generates no information loss in estimating x_k using $\hat{x}_{k/k-N}$ and $Z_{k/k-N+1}^{\perp}$. As a result, an OCM and its error statistics are obtained as follows.

$$\begin{aligned}
Z_{k/k-N+1}^{n} &:= (H_{k/k-N+1}^{\perp})^{+} Z_{k/k-N+1}^{\perp} \\
&= e_{k/k-N} + v_{k/k-N+1}^{n}
\end{aligned} \tag{37}$$

where

$$\begin{aligned}
(H_{k/k-N+1}^{\perp})^{+} &:= \left[(H_{k/k-N+1}^{\perp})^T (R_{k/k-N+1}^{\perp})^{-1} (H_{k/k-N+1}^{\perp})\right]^{-1} (H_{k/k-N+1}^{\perp})^T (R_{k/k-N+1}^{\perp})^{-1} \\
v_{k/k-N+1}^{n} &:= (H_{k/k-N+1}^{\perp})^{+} v_{k/k-N+1}^{\perp} \\
v_{k/k-N+1}^{n} &\sim \left(O, R_{k/k-N+1}^{n}\right)
\end{aligned}$$

$$R^n_{k/k-N+1} = (H^{\perp}_{k/k-N+1})^{+} R^{\perp}_{k/k-N+1} \left[H^{\perp}_{k/k-N+1}\right]^T$$
$$= \left[(H^{\perp}_{k/k-N+1})^T (R^{\perp}_{k/k-N+1})^{-1} (H^{\perp}_{k/k-N+1})\right]^{-1}$$
$$S^n_{k/k-N+1} := E\left[\left(v^n_{k/k-N+1}\right)\left(e_{k/k-N}\right)^T\right] = O\,. \tag{38}$$

The resulting OCM $Z^n_{k/k-N+1}$ has the following analytic advantages:

i) the measurement coefficient matrix is a simple identity matrix;

ii) the measurement noise is not correlated with $e_{k/k-N}$;

iii) the dimension is minimal, i.e., equal to the system dimension;

iv) it is closely related with the observability grammian.

By the OCM $Z^n_{k/k-N+1}$, the computation of the optimal estimate $\hat{x}_{k/k}$ of x_k using $\hat{x}_{k/k-N}$ and $\{z_j\}_{k-N<j\le k}$ can be performed as follows.

$$K^n_{k/k-N+1} = M_k\left[M_k + R^n_{k/k-N+1}\right]^{-1}$$
$$P_k = M_k - K^n_{k/k-N+1} M_k$$
$$\hat{x}_{k/k} = \hat{x}_{k/k-N} - K^n_{k/k-N+1} Z^n_{k/k-N+1}\,. \tag{39}$$

By performing Eq. (39), the *a posteriori* estimation error e_k of $\hat{x}_k$ satisfies

$$e_{k/k} = \left[I - K^n_{k/k-N+1}\right] e_{k/k-N} - K^n_{k/k-N+1} v^n_{k/k-N+1}\,. \tag{40}$$

IV. Analysis by a Stochastically Driven Lyapunov function

An analysis of LTV KF by a stochastically-driven Lyapunov function is now given. For analysis, we introduce two observability grammian concepts: the scaled observability grammian and the orthogonalized observability grammian. The scaled observability grammian $\Theta_{k/k-N+1}$ is defined by

$$\Theta_{k/k-N+1} := (H^s_{k/k-N+1})^T (H^s_{k/k-N+1})\,. \tag{41}$$

Since we assumed for brevity that the error covariance matrix of each measurement is constant, the conventional-sense observability grammian can be obtained by scaling $\Theta_{k/k-N+1}$. The orthogonalized observability grammian $\Theta^{\perp}_{k/k-N+1}$ is defined as

$$\Theta^{\perp}_{k/k-N+1} := (H^{\perp}_{k/k-N+1})^T (R^{\perp}_{k/k-N+1})^{-1} (H^{\perp}_{k/k-N+1}) . \quad (42)$$

As shown, the orthogonalized observability grammian $\Theta^{\perp}_{k/k-N+1}$ is closely related with the OCM since its inverse, if it exists, is the error covariance matrix of the OCM. Compared with the degree of observability, the degree of controllability is represented only by the conventional-sense controllability grammian $\Xi_{k-1/k-N}$ defined in Eq. (24).

For simplicity of analysis, we consider only the system that is uniformly controllable and uniformly observable. Thus, there exists a positive integer N and positive constants $\underline{\theta}$, $\overline{\theta}$, $\underline{\xi}$, and $\overline{\xi}$ such that

$$O < \underline{\theta} I \le \Theta_{k/k-N+1} \le \overline{\theta} I , \ \forall k \ge N ,$$

$$O < \underline{\xi} I \le \Xi_{k-1/k-N} \le \overline{\xi} I , \ \forall k \ge N . \quad (43)$$

In addition, we assume that, once N is fixed, the system matrices are bounded. Thus, there exist positive real constants $\underline{f}$, $\bar{f}$, $\overline{g}$, and $\overline{h}$ that satisfy the following inequalities.

$$O < \underline{f} I \le F_{k/j} \le \bar{f} I , \quad k \ge N , \quad j = k-N, k-N+1, \cdots, k-1$$

$$\|G_k\|_\infty \le \overline{g} , \qquad \|h_k\|_\infty \le \overline{h} , \qquad k = 1,2,3,\cdots \quad (44)$$

Lemma 4-1: Boundedness of Error Covariance Matrices [2-4]

Assuming that $P_0 > O$, the given system is uniformly completely controllable and observable satisfying Eq. (43), and the system matrices are bounded satisfying Eq. (44). Then, the solutions of the discrete Riccatti recursion satisfying

$$M_k = F_{k/k-N} P_{k-N} F_{k/k-N}{}^T + \Xi_{k-1/k-N}$$

$$P_k{}^{-1} = M_k{}^{-1} + (H^{\perp}_{k/k-N+1})^T (R^{\perp}_{k/k-N+1})^{-1} (H^{\perp}_{k/k-N+1})^{-1} \quad (45)$$

are bounded by the following inequalities.

$$O < \underline{p} I \le P_k \le \overline{p} I , \ \forall k = N, 2N, 3N, \cdots$$

$$O < \underline{m} I \le M_k \le \overline{m} I , \ \forall k = 2N, 3N, \cdots \quad (46)$$

where

$$\underline{p} := \frac{r\underline{\xi}}{r+\bar{\theta}\underline{\xi}} > 0\,, \quad \bar{p} := \frac{r}{\underline{\theta}} + \bar{\xi}$$

$$\underline{m} := \underline{f}^2\,\underline{p} + \underline{\xi} = \frac{\left[(1+\underline{f}^2)r + \bar{\theta}\underline{\xi}\right]\underline{\xi}}{r+\bar{\theta}\underline{\xi}} > 0$$

$$\bar{m} := \bar{f}^2\,\bar{p} + \bar{\xi} = \frac{\bar{f}^2 r + (1+\bar{f}^2)\underline{\theta}\bar{\xi}}{\underline{\theta}}\,, \tag{47}$$

which means that the solutions of the discrete Riccati recursion M_k and P_k remain bounded.

□

Remark: Lemma 4-1 is the modification of previous study results [2-4] regarding the boundedness of the discrete Riccati recursion that utilizes the optimality of P_k. By this lemma, various coefficient matrices that will be used in the later Lyapunov analysis are shown to be bounded.

Lemma 4-2:

If the system matrices are bounded satisfying Eq. (44), then

$$\bar{\sigma}\left(H^o_{k/k-N+1}\right) \le h_o\,, \tag{48}$$

where

$$h_o := 0.7N\bar{h}\bar{f}\bar{g}\,. \tag{49}$$

<Proof>

$$\left\|H^o_{k/k-N+1}\right\|_\infty \le \left\| \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ \left\|h_{k-1}F_{k-1/k}G_{k-1}\right\|_\infty & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \left\|hF_{k-N+2/k}G_{k-1}\right\|_\infty & \left\|hF_{k-N+2/k-1}G_{k-2}\right\|_\infty & \cdots & 0 & 0 \\ \left\|hF_{k-N+1/k}G_{k-1}\right\|_\infty & \left\|hF_{k-N+1/k-1}G_{k-2}\right\|_\infty & \cdots & \left\|hF_{k-N+1/k-N+2}G_{k-N+1}\right\|_\infty & 0 \end{bmatrix} \right\|_\infty$$
$$\le \bar{h}\bar{f}\bar{g} \left\| \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \cdots & 0 & 0 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \right\|_\infty \le 0.7N\bar{h}\bar{f}\bar{g} \tag{50}$$

□

Remark: $H^o_{k/k-N+1}$ used in Eq. (50) represents the correlation between the stacked measurement $Z^s_{k/k-N+1}$ and the *a priori* estimation error $e_{k/k-N}$. Lemma 4-2 shows that the maximum singular value of

$H^{o}_{k/k-N+1}$ is bounded by a finite positive number although it increases with the block size N of the stacked measurement.

Lemma 4-3: Inequality between Error Covariance of Equivalent Measurements

If $P_{k-N} > O$, (51)

then,

$$R^{o}_{k/k-N+1} < R^{\perp}_{k/k-N+1} = R^{s}_{k/k-N+1} - (S^{s}_{k/k-N+1})M_k^{-1}(S^{s}_{k/k-N+1})^T \le R^{s}_{k/k-N+1}. \quad (52)$$

<Proof>

i) By definition, it is obvious that

$$R^{\perp}_{k/k-N+1} = R^{s}_{k/k-N+1} - (S^{s}_{k/k-N+1})M_k^{-1}(S^{s}_{k/k-N+1})^T \le R^{s}_{k/k-N+1}. \quad (53)$$

ii) If $P_{k-N} > O$, since $Q^{o}_{k-1/k-N} > O$ by definition, we have

$$\Rightarrow M_k = F_{k/k-N}P_{k-N}F_{k/k-N}{}^T + G^{o}_{k-1/k-N}Q^{o}_{k-1/k-N}(G^{o}_{k-1/k-N})^T > G^{o}_{k-1/k-N}Q^{o}_{k-1/k-N}(G^{o}_{k-1/k-N})^T$$

$$\Rightarrow Q^{o}_{k-1/k-N} + Q^{o}_{k-1/k-N}(G^{o}_{k-1/k-N})^T\left[M_k - G^{o}_{k-1/k-N}Q^{o}_{k-1/k-N}(G^{o}_{k-1/k-N})^T\right]^{-1}G^{o}_{k-1/k-N}Q^{o}_{k-1/k-N} > O$$

$$\Rightarrow \left[\left(Q^{o}_{k-1/k-N}\right)^{-1} - (G^{o}_{k-1/k-N})^T M_k^{-1}(G^{o}_{k-1/k-N})\right]^{-1} > O$$

$$\Rightarrow \left(Q^{o}_{k-1/k-N}\right)^{-1} - (G^{o}_{k-1/k-N})^T M_k^{-1}(G^{o}_{k-1/k-N}) > O$$

$$\Rightarrow Q^{o}_{k-1/k-N} - Q^{o}_{k-1/k-N}(G^{o}_{k-1/k-N})^T M_k^{-1}(G^{o}_{k-1/k-N})Q^{o}_{k-1/k-N} > O$$

$$\Rightarrow \begin{array}{l} R^{o}_{k/k-N+1} + \{ H^{o}_{k-1/k-N}Q^{o}_{k-1/k-N}(H^{o}_{k-1/k-N})^T \\ \quad - H^{o}_{k-1/k-N}Q^{o}_{k-1/k-N}(G^{o}_{k-1/k-N})^T M_k^{-1}G^{o}_{k-1/k-N}Q^{o}_{k-1/k-N}(H^{o}_{k-1/k-N})^T \} > R^{o}_{k/k-N+1} \end{array}$$

$$\Rightarrow R^{s}_{k/k-N+1} - H^{o}_{k-1/k-N}Q^{o}_{k-1/k-N}(G^{o}_{k-1/k-N})^T M_k^{-1}G^{o}_{k-1/k-N}Q^{o}_{k-1/k-N}(H^{o}_{k-1/k-N})^T > R^{o}_{k/k-N+1}$$

$$\Rightarrow R^{s}_{k/k-N+1} - (S^{s}_{k/k-N+1})M_k^{-1}(S^{s}_{k/k-N+1})^T > R^{o}_{k/k-N+1}$$

$$\Rightarrow R^{\perp}_{k/k-N+1} > R^{o}_{k/k-N+1}. \quad (54)$$

□

Remark: Lemma 4-3 shows that the error covariance matrix $R^{\perp}_{k/k-N+1}$ of the de-correlated stacked measurement is bounded above and below regardless of largeness of process noise $\{w_j\}_{k-N\le j<k}$.

Lemma 4-4: Error Covariance Inequality of Two Dependent Gaussian Random Variables

Suppose we are given two Gaussian random vectors W and $(Z-He)$ with the following co-distribution.

$$\begin{bmatrix} W \\ \hline Z-He \end{bmatrix} \sim \left(O, \left[\begin{array}{c|c} \Xi & S^T \\ \hline S & R \end{array}\right] \right) \tag{55}$$

Then, the following inequalities hold.

$$\begin{aligned} & R - S\Xi^{-1}S^T > O \\ & \Xi - S^T R^{-1} S > O \end{aligned} \tag{56}$$

<Proof>

By definition of error covariance matrices, it is obvious that

$$\left[\begin{array}{c|c} \Xi & S^T \\ \hline S & R \end{array}\right] > O. \tag{57}$$

Since

$$\Xi > O, \quad R > O, \tag{58}$$

the joint error covariance matrix appearing in Eq. (55) can be decomposed as

$$\begin{aligned} \left[\begin{array}{c|c} \Xi & S^T \\ \hline S & R \end{array}\right] &= \left[\begin{array}{c|c} I & O \\ \hline S\Xi^{-1} & I \end{array}\right] \left[\begin{array}{c|c} \Xi & O \\ \hline O & R - S\Xi^{-1}S^T \end{array}\right] \left[\begin{array}{c|c} I & \Xi^{-1}S^T \\ \hline O & I \end{array}\right] \\ &= \left[\begin{array}{c|c} I & S^T R^{-1} \\ \hline O & R \end{array}\right] \left[\begin{array}{c|c} \Xi - S^T R^{-1} S & O \\ \hline O & R \end{array}\right] \left[\begin{array}{c|c} I & O \\ \hline R^{-1}S & I \end{array}\right]. \end{aligned} \tag{59}$$

Thus, if the inequalities of Eq. (56) do not hold, Eq. (57) does not hold, which results in a contradiction.

□

Remark: W in Lemma 4-4 represents the accumulated process noise, Ξ represents the controllability grammian, $(Z-He)$ represents the measurement error, and R represents the error covariance matrix of an equivalent measurement that is gathered within a fixed time-interval.

Lemma 4-5: Lower Bound for Intermediate Observability Grammian

Assume that the following matrix inequality holds.

$$S^T R^{-1} S < M \tag{60}$$

where $M = M^T > O$ and $R = R^T > O$. Then, for any non-negative constant ε, the following matrix inequality holds.

$$(H + SM^{-1})^T R^{-1} (H + SM^{-1}) \geq \frac{\varepsilon}{1+\varepsilon} H^T R^{-1} H - \varepsilon M^{-1} \tag{61}$$

where all the matrix-dimensions are assumed appropriate.

<Proof>

Since R is positive definite and symmetric, it can be decomposed as follows.

$$R = R^{\frac{1}{2}} R^{\frac{1}{2}}, \qquad R^{\frac{1}{2}} > O \tag{62}$$

Define X as

$$X := \frac{1}{\sqrt{1+\varepsilon}} R^{-\frac{1}{2}} H + \sqrt{1+\varepsilon} R^{-\frac{1}{2}} S M^{-1}. \tag{63}$$

By the definition of error covariance matrix, it is obvious that

$$\begin{aligned} X^T X &= \frac{1}{1+\varepsilon} H^T R^{-1} H + (1+\varepsilon) M^{-1} (S^T R^{-1} S) M^{-1} \\ &\quad + H^T R^{-1} S M^{-1} + M^{-1} S^T R^{-1} H. \end{aligned} \tag{64}$$

Since $X^T X > O$, we have

$$\begin{aligned} M^{-1}(S^T R^{-1} S) M^{-1} &+ H^T R^{-1} S M^{-1} + M^{-1} S^T R^{-1} H \\ &\geq -\frac{1}{1+\varepsilon} H^T R^{-1} H - \varepsilon M^{-1} (S^T R^{-1} S) M^{-1}. \end{aligned} \tag{65}$$

Utilizing Eqs. (60) and (65), we finally have the result.

□

Remark: $(H + SM^{-1})^T R^{-1} (H + SM^{-1})$ in Lemma 4-5 represents an intermediate observability grammian that is necessary to seek the lower bound of the orthogonalized observability grammian. $H^T R^{-1} H$ represents the observability grammian by the original concept. Thus, Lemma 4-5 provides us with information on how the orthogonalized observability grammian can be positive definite.

Lemma 4-6: Lower Bound for Orthogonalized Observability Grammian

Assume that Eq. (43), which means uniform complete controllability and observability, is satisfied. If there exists an integer N and a positive constant a that satisfies the following condition

$$(H^s_{k/k-N+1})^T (R^s_{k/k-N+1})^{-1} (H^s_{k/k-N+1}) \geq (\Xi_{k-1/k-N})^{-1} + aI \,, \tag{66}$$

the orthogonalized observability grammian $\Theta^{\perp}_{k/k-N+1}$ is bounded below by the following inequality.

$$\begin{aligned} &(H^{\perp}_{k/k-N+1})^T (R^{\perp}_{k/k-N+1})^{-1} (H^{\perp}_{k/k-N+1}) \geq \frac{1}{\bar{\gamma}} I \\ &\bar{\gamma} := \frac{2(2 + a\underline{\xi})}{(a\underline{\xi})^2} > 0 \end{aligned} \tag{67}$$

where $\underline{\xi}$ is a constant bound for the controllability grammian shown in Eq. (43).

<Proof>

Since

$$M_k > \Xi_{k-1/k-N} \,, \tag{68}$$

the following inequality holds by adopting the result of Lemma 4-5.

$$\begin{aligned} &(H^{\perp}_{k/k-N+1})^T (R^s_{k/k-N+1})^{-1} (H^{\perp}_{k/k-N+1}) \\ &\qquad \geq \frac{\varepsilon}{1+\varepsilon} (H^s_{k/k-N+1})^T (R^s_{k/k-N+1})^{-1} (H^s_{k/k-N+1}) - \varepsilon (\Xi_{k-1/k-N})^{-1} \\ &\qquad \geq \frac{\varepsilon a}{1+\varepsilon} I - \frac{\varepsilon^2}{1+\varepsilon} (\Xi_{k-1/k-N})^{-1} \\ &\qquad \geq \frac{\varepsilon (a - \varepsilon / \underline{\xi})}{1+\varepsilon} I \end{aligned} \tag{69}$$

Thus, if we select ε as

$$\varepsilon = \frac{a\underline{\xi}}{2} \,, \tag{70}$$

the following inequality results.

$$(H^{\perp}_{k/k-N+1})^T (R^s_{k/k-N+1})^{-1} (H^{\perp}_{k/k-N+1}) \geq \frac{(a\underline{\xi})^2}{2(2 + a\underline{\xi})} I > O \tag{71}$$

Since $R^s_{k/k-N+1} \geq R^{\perp}_{k/k-N+1}$, as shown by Lemma 4-3, the result is obtained.

□

Remark: The OCM introduced in Section III becomes full column-rank if $(H^{\perp}_{k/k-N+1})^T (H^{\perp}_{k/k-N+1})$ is non-singular. Lemma 4-6 means that a full column-rank OCM exists if Eq. (66) is satisfied. The physical meaning of Eq. (66) is that the information gathered by the measurements, excluding the *a priori* estimate, is greater than the information gathered by propagating the initially-perfect estimate with process noises.

Lemma 4-7: Upper Bound for Orthogonalized Observability Grammian

Assume that the given system is uniformly completely controllable and uniformly completely observable satisfying Eq. (43) and that the system matrices are bounded satisfying Eq. (44). Then, the orthogonalized observability grammian $\Theta^{\perp}_{k/k-N+1}$ is upper-bounded by

$$(H^{\perp}_{k/k-N+1})^T (R^{\perp}_{k/k-N+1})^{-1} (H^{\perp}_{k/k-N+1}) \le \frac{1}{\underline{\gamma}} I$$

$$\underline{\gamma} := \frac{1}{2r}\left[\overline{\theta} + \overline{m}\left(\frac{1}{\underline{m}}\right)^2 (r + qh_o^2)\right]^{-1} > 0 \, . \tag{72}$$

<Proof>

By Eq. (30), Lemma 4-2, and Lemma 4-3, the following inequality holds.

$$R^s_{k/k-N+1} \le (r + qh_o^2) I \tag{73}$$

Due to Eq. (73) and Lemma 4-4, the following inequality also holds.

$$\frac{1}{r + qh_o^2} (S^s_{k/k-N+1})^T (S^s_{k/k-N+1}) \le (S^s_{k/k-N+1})^T (R^s_{k/k-N+1})^{-1} (S^s_{k/k-N+1}) \le M_k \tag{74}$$

By Eqs. (52) and Lemma 4-1, we obtain

$$(S^s_{k/k-N+1})^T (S^s_{k/k-N+1}) \le \overline{m}(r + qh_o^2) I \, . \tag{75}$$

Since $R^o_{k/k-N+1} < R^{\perp}_{k/k-N+1}$, as shown in Lemma 4-3, it can be shown that

$$
\begin{aligned}
&(H^{\perp}_{k/k-N+1})^T (R^{\perp}_{k/k-N+1})^{-1} (H^{\perp}_{k/k-N+1}) \\
&\qquad \le (H^{\perp}_{k/k-N+1})^T (R^{o}_{k/k-N+1})^{-1} (H^{\perp}_{k/k-N+1}) \\
&\qquad = r(H^{s}_{k/k-N+1} + S^{s}_{k/k-N+1} M_k^{-1})^T (H^{s}_{k/k-N+1} + S^{s}_{k/k-N+1} M_k^{-1}) \\
&\qquad = r\begin{bmatrix} (H^{s}_{k/k-N+1})^T (H^{s}_{k/k-N+1}) + M_k^{-1} (S^{s}_{k/k-N+1})^T H^{s}_{k/k-N+1} \\ + (H^{s}_{k/k-N+1})^T S^{s}_{k/k-N+1} M_k^{-1} + M_k^{-1} (S^{s}_{k/k-N+1})^T S^{s}_{k/k-N+1} M_k^{-1} \end{bmatrix} \\
&\qquad \le \frac{r(\varepsilon+1)}{\varepsilon} (H^{s}_{k/k-N+1})^T (H^{s}_{k/k-N+1}) \\
&\qquad\qquad + r(\varepsilon+1) M_k^{-1} (S^{s}_{k/k-N+1})^T S^{s}_{k/k-N+1} M_k^{-1}.
\end{aligned}
\tag{76}
$$

Letting $\varepsilon = 1$, we have

$$
\begin{aligned}
&(H^{\perp}_{k/k-N+1})^T (R^{\perp}_{k/k-N+1})^{-1} (H^{\perp}_{k/k-N+1}) \\
&\qquad \le 2r\left[\overline{\theta} + \overline{m}\left(\frac{1}{\underline{m}}\right)^2 (r + q h_o^2) \right].
\end{aligned}
\tag{77}
$$

Lemma 4-8: Bounds for Orthogonalized Observability Grammian

Assume that the given system is uniformly completely controllable and uniformly completely observable satisfying Eq. (43). If there exists an integer N and a positive constant a that satisfies

$$
(H^{s}_{k/k-N+1})^T (R^{s}_{k/k-N+1})^{-1} (H^{s}_{k/k-N+1}) \ge (\Xi_{k-1/k-N})^{-1} + aI \,, \tag{78}
$$

then the error covariance matrix of the OCM is bounded by

$$
\underline{\gamma} I \le R^{n}_{k/k-N+1} \le \overline{\gamma} I \tag{79}
$$

where the positive bounding constants $\overline{\gamma}$ and $\underline{\gamma}$ are defined in Lemma 4-6 and Lemma 4-7, respectively.

For the stability analysis of the stochastic Lyapunov method, we make use of the following two definitions and one lemma for the boundedness of stochastic process.

Definition 4-1: Exponential Boundedness in Mean Square [6-8]

The stochastic process ζ_n is said to be exponentially bounded in mean square, if there are real numbers η, $v > 0$, and $0 < \vartheta < 1$, such that

$$E\left\{\left\|\zeta_n\right\|^2\right\} \le \eta \left\|\zeta_0\right\| \vartheta^n + \nu \tag{80}$$

holds for every $n \ge 0$.

Definition 4-2: Boundedness with Probability One [6-8]

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

$$\sup_{n \ge 0} \left\|\zeta_n\right\| < \infty \tag{81}$$

holds with probability one.

For later use, we recall some standard results about the boundedness of stochastic processes.

Lemma 4-8 [8]

Assume there is a stochastic process $V_n(e_n)$ as well as real numbers $\underline{v}, \bar{v}, \mu > 0$ and $0 < \alpha \le 1$ such that

$$\underline{v}\left\|e_n\right\|^2 \le V_n(e_n) \le \bar{v}\left\|e_n\right\|^2 \tag{82}$$

and

$$E\left\{V_{n+1}\left(e_{n+1}\right) \mid e_n\right\} - V_n\left(e_n\right) \le \mu - \alpha V_n\left(e_n\right) \tag{83}$$

are fulfilled for every solution e_n of Eqs. (4) and (5). Then the stochastic process is exponentially bounded in mean square, i.e., we have

$$E\left\{\left\|e_n\right\|^2\right\} \le \frac{\bar{v}}{\underline{v}} E\left\{\left\|e_0\right\|^2\right\}(1-\alpha)^n + \frac{\mu}{\underline{v}} \sum_{i=1}^{n-1} (1-\alpha)^i \tag{84}$$

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

□

Since all the preliminaries are completed prepared, we can now state the main result of this paper.

Theorem: Stochastic Radius of Attraction of LTV KF

Consider the discrete LTV system of Eqs. (1)-(3) which is uniformly completely controllable and uniformly completely observable satisfying Eq. (43), and the system matrices are bounded satisfying Eq.

(44). Given conditions i) and ii),

i) there exists an integer N and positive constants $\bar{\gamma}$ and $\underline{\gamma}$ by which

$$O < \underline{\gamma} I \le R^n_{k/k-N+1} \le \bar{\gamma} I \quad \text{holds for all} \quad k \ge N , \tag{85}$$

and

ii) there exists an integer N and a positive constant a that satisfies Eq. (66),

then if condition i) or ii) holds, the estimation error of the KF by Eqs. (4) and (5) is exponentially bounded in mean square and bounded with probability one with the radius of contraction π as

$$\pi = \sqrt{\bar{\gamma}\,\bar{p}\,\underline{p}\left(\frac{\bar{\xi}}{\underline{m}} + 2\frac{\bar{p}}{\underline{\gamma}}\right)} \tag{86}$$

where the positive constants $\bar{p}$, $\underline{p}$, $\underline{m}$, $\bar{\xi}$, $\bar{\gamma}$, and $\underline{\gamma}$ are defined in Eqs. (43), (47), (67), and (72).

<Proof>

The following matrices

$$M_k := P_{k/k-N}$$

$$K^n_{k/k-N+1} := M_k (M_k + R^n_{k/k-N+1})^{-1} = P_k (R^n_{k/k-N+1})^{-1}$$

$$\begin{aligned} U_k &:= -K^n_{k/k-N+1} e_{k/k-N} - K^n_{k/k-N+1} v^n_{k/k-N+1} \\ &= -\left[P_k (R^n_{k/k-N+1})^{-1}\right] e_{k/k-N} - \left[P_k (R^n_{k/k-N+1})^{-1}\right] v^n_{k/k-N+1} \end{aligned}$$

$$\Gamma_{k/k-N} := F_{k/k-N}{}^{-1}$$

$$N_k := \left(F_{k/k-N}{}^T M_k{}^{-1} F_{k/k-N}\right)^{-1} \tag{87}$$

satisfy the relationships

$$P_k{}^{-1} = M_k{}^{-1} + (R^n_{k/k-N+1})^{-1}$$

$$P_k M_k{}^{-1} = I - P_k (R^n_{k/k-N+1})^{-1} = I - K^n_{k/k-N+1}$$

$$\begin{aligned} e_k &= P_k M_k{}^{-1} e_{k/k-N} - K_{k/k-N+1} v^n_{k/k-N+1} \\ &= e_{k/k-N} + \left(P_k M_k{}^{-1} - I\right) e_{k/k-N} - K_{k/k-N+1} v^n_{k/k-N+1} \end{aligned}$$

$$e_{k/k-N} = M_k P_k^{-1} e_k + M_k (R^n_{k/k-N+1})^{-1} v^n_{k/k-N+1}$$

$$N_k^{-1} = F_{k/k-N}{}^T M_k^{-1} F_{k/k-N} \,. \tag{88}$$

We also have

$$e_k = e_{k/k-N} + U_k = \left(I - K^n_{k/k-N+1}\right) e_{k/k-N} - K^n_{k/k-N+1} v^n_{k/k-N+1} \,. \tag{89}$$

Defining a Lyapunov candidate $V_k(e_k)$ as

$$V_k(e_k) := e_k{}^T P_k^{-1} e_k \,, \tag{90}$$

then, by the properties of Eqs. (87)–(89), $V_k(e_k)$ satisfies

$$\begin{aligned}
V_k(e_k) &= e_{k/k-N}{}^T M_k^{-1} e_{k/k-N} - e_k{}^T (R^n_{k/k-N+1})^{-1} e_k + 2 e_k{}^T P_k^{-1} e_k \\
&\quad - e_k{}^T M_k^{-1} e_k - e_{k/k-N}{}^T M_k^{-1} e_{k/k-N} \\
&= e_{k/k-N}{}^T M_k^{-1} e_{k/k-N} - e_k{}^T (R^n_{k/k-N+1})^{-1} e_k + 2 e_k{}^T M_k^{-1} e_{k/k-N} \\
&\quad - 2 e_k{}^T P_k^{-1} K^n_{k/k-N+1} v^n_{k/k-N+1} - e_k{}^T M_k^{-1} e_k - e_{k/k-N}{}^T M_k^{-1} e_{k/k-N} \\
&= e_{k/k-N}{}^T M_k^{-1} e_{k/k-N} - e_k{}^T (R^n_{k/k-N+1})^{-1} e_k \\
&\quad - \left(e_k - e_{k/k-N}\right)^T M_k^{-1} \left(e_k - e_{k/k-N}\right) - 2 e_k{}^T (R^n_{k/k-N+1})^{-1} v^n_{k/k-N+1} \\
&= e_{k/k-N}{}^T M_k^{-1} e_{k/k-N} - e_k{}^T (R^n_{k/k-N+1})^{-1} e_k \\
&\quad - \left(e_k - e_{k/k-N}\right)^T M_k^{-1} \left(e_k - e_{k/k-N}\right) - 2 e_k{}^T (R^n_{k/k-N+1})^{-1} v^n_{k/k-N+1} \\
&= e_{k/k-N}{}^T M_k^{-1} e_{k/k-N} - e_k{}^T (R^n_{k/k-N+1})^{-1} e_k - U_k{}^T M_k^{-1} U_k \\
&\quad - 2 e_{k/k-N}{}^T \left(I - K^n_{k/k-N+1}\right)^T (R^n_{k/k-N+1})^{-1} v^n_{k/k-N+1} \\
&\quad + 2 (v^n_{k/k-N+1})^T (R^n_{k/k-N+1})^{-1} P_k (R^n_{k/k-N+1})^{-1} v^n_{k/k-N+1} \\
&= \left(F_{k/k-N} e_{k-N} + W_{k-1/k-N}\right)^T M_k^{-1} \left(F_{k/k-N} e_{k-N} + W_{k-1/k-N}\right) \\
&\quad - e_k{}^T (R^n_{k/k-N+1})^{-1} e_k - U_k{}^T M_k^{-1} U_k \\
&\quad - 2 (F_{k/k-N} e_{k-N} + W_{k-1/k-N})^T \left(I - K^n_{k/k-N+1}\right)^T (R^n_{k/k-N+1})^{-1} v^n_{k/k-N+1} \\
&\quad + 2 (v^n_{k/k-N+1})^T (R^n_{k/k-N+1})^{-1} P_k (R^n_{k/k-N+1})^{-1} v^n_{k/k-N+1} \\
&\le e_{k-N}{}^T P_{k-N}^{-1} e_{k-N} \\
&\quad - e_k{}^T (R^n_{k/k-N+1})^{-1} e_k - U_k{}^T M_k^{-1} U_k \\
&\quad + 2 e_{k-N}{}^T F_{k/k-N}{}^T M_k^{-1} W_{k-1/k-N} \\
&\quad - 2 e_{k-N}{}^T F_{k/k-N}{}^T \left(I - K^n_{k/k-N+1}\right)^T (R^n_{k/k-N+1})^{-1} v^n_{k/k-N+1} \\
&\quad - 2 W_{k-1/k-N}{}^T \left(I - K^n_{k/k-N+1}\right)^T (R^n_{k/k-N+1})^{-1} v^n_{k/k-N+1} \\
&\quad + 2 (v^n_{k/k-N+1})^T (R^n_{k/k-N+1})^{-1} P_k (R^n_{k/k-N+1})^{-1} v^n_{k/k-N+1} \\
&\quad + W_{k-1/k-N}{}^T M_k^{-1} W_{k-1/k-N}
\end{aligned}$$

$$
\begin{aligned}
\leq\ & V_{k-N}(e_{k-N}) \\
& - e_k{}^T (R^n_{k/k-N+1})^{-1} e_k \\
& - U_k{}^T M_k{}^{-1} U_k \\
& + W_{k-1/k-N}{}^T M_k{}^{-1} W_{k-1/k-N} \\
& + 2(v^n_{k/k-N+1})^T (R^n_{k/k-N+1})^{-1} P_k (R^n_{k/k-N+1})^{-1} v^n_{k/k-N+1} \\
& + 2e_{k-N}{}^T F_{k/k-N}{}^T M_k{}^{-1} W_{k-1/k-N} \\
& - 2e_{k-N}{}^T F_{k/k-N}{}^T \left(I - K^n_{k/k-N+1}\right)^T (R^n_{k/k-N+1})^{-1} v^n_{k/k-N+1} \\
& - 2W_{k-1/k-N}{}^T \left(I - K^n_{k/k-N+1}\right)^T (R^n_{k/k-N+1})^{-1} v^n_{k/k-N+1}
\end{aligned}
$$

$$
\begin{aligned}
\leq\ & V_{k-N}(e_{k-N}) \\
& - e_k{}^T (R^n_{k/k-N+1})^{-1} e_k \\
& + W_{k-1/k-N}{}^T M_k{}^{-1} W_{k-1/k-N} \\
& + 2(v^n_{k/k-N+1})^T (R^n_{k/k-N+1})^{-1} P_k (R^n_{k/k-N+1})^{-1} v^n_{k/k-N+1} \\
& + C(k)
\end{aligned}
\tag{91}
$$

where $C(k)$ in Eq. (91) is defined as

$$
\begin{aligned}
C(k) := {} & 2e_{k-N}{}^T F_{k/k-N}{}^T M_k{}^{-1} W_{k-1/k-N} \\
& - 2e_{k-N}{}^T F_{k/k-N}{}^T \left(I - K^n_{k/k-N+1}\right)^T (R^n_{k/k-N+1})^{-1} v^n_{k/k-N+1} \\
& - 2W_{k-1/k-N}{}^T \left(I - K^n_{k/k-N+1}\right)^T (R^n_{k/k-N+1})^{-1} v^n_{k/k-N+1.}
\end{aligned}
\tag{92}
$$

By Lemma 4-1 and Lemma 4-8, it can be shown that

$$
e_k{}^T (R^n_{k/k-N+1})^{-1} e_k \geq \frac{1}{\bar{\gamma}} \|e_k\|^2 \geq \frac{1}{\bar{\gamma}\underline{p}} e_k{}^T P_k{}^{-1} e_k = \frac{1}{\bar{\gamma}\underline{p}} V_k(e_k)\ . \tag{93}
$$

Substituting Eq. (93) for Eq. (91), we have

$$
\begin{aligned}
V_k(e_k) \leq {} & V_{k-N}(e_{k-N}) \\
& - \frac{1}{\bar{\gamma}\underline{p}} V_k(e_k) \\
& + W_{k-1/k-N}{}^T M_k{}^{-1} W_{k-1/k-N} \\
& + 2(v^n_{k/k-N+1})^T (R^n_{k/k-N+1})^{-1} P_k (R^n_{k/k-N+1})^{-1} v^n_{k/k-N+1} \\
& + C(k).
\end{aligned}
\tag{94}
$$

Rearranging and scaling the above equation, we obtain

$$
\begin{aligned}
V_k(e_k) \le \; & \frac{\bar{\gamma}\,\underline{p}}{1+\bar{\gamma}\,\underline{p}} V_{k-N}(e_{k-N}) \\
& + \frac{\bar{\gamma}\,\underline{p}}{1+\bar{\gamma}\,\underline{p}} W_{k-1/k-N}{}^{T} M_k{}^{-1} W_{k-1/k-N} \\
& + 2\frac{\bar{\gamma}\,\underline{p}}{1+\bar{\gamma}\,\underline{p}} (v^n_{k/k-N+1})^T (R^n_{k/k-N+1})^{-1} P_k (R^n_{k/k-N+1})^{-1} v^n_{k/k-N+1} \\
& + \frac{\bar{\gamma}\,\underline{p}}{1+\bar{\gamma}\,\underline{p}} C(k).
\end{aligned}
\tag{95}
$$

Taking the conditional expectation $E[V_k(e_k)|e_{k-N}]$ and removing the products of uncorrelated terms by expectation, we have

$$
\begin{aligned}
E[V_k(e_k)|e_{k-N}] - V_{k-N}(e_{k-N}) \le & -\frac{1}{1+\bar{\gamma}\,\underline{p}} V_{k-N}(e_{k-N}) \\
& + \frac{\bar{\gamma}\,\underline{p}}{1+\bar{\gamma}\,\underline{p}} \cdot \frac{\bar{\xi}}{\underline{m}} \\
& + 2\frac{\bar{\gamma}\,\underline{p}}{1+\bar{\gamma}\,\underline{p}} \cdot \frac{\bar{p}}{\underline{\gamma}}.
\end{aligned}
\tag{96}
$$

To summarize, we have the inequality

$$
E[V_k(e_k)|e_{k-N}] - V_{k-N}(e_{k-N}) \quad \le \quad -\alpha V_{k-N}(e_{k-N}) + \beta
\tag{97}
$$

where

$$
\alpha := \frac{1}{1+\bar{\gamma}\,\underline{p}} > 0, \qquad \beta := \frac{\bar{\gamma}\,\underline{p}}{1+\bar{\gamma}\,\underline{p}} \left(\frac{\bar{\xi}}{\underline{m}} + 2\frac{\bar{p}}{\underline{\gamma}} \right).
\tag{98}
$$

In addition, from

$$
\beta \; < \; \frac{\alpha}{\bar{p}} \pi^2 \quad \le \quad \frac{\alpha}{\bar{p}} \|e_{k-N}\|_2^2 \quad \le \quad \alpha V_{k-N}(e_{k-N})
\tag{99}
$$

we find that the radius of attraction is

$$
\pi = \sqrt{\frac{\bar{p}\beta}{\alpha}} = \sqrt{\bar{\gamma}\,\bar{p}\,\underline{p} \left(\frac{\bar{\xi}}{\underline{m}} + 2\frac{\bar{p}}{\underline{\gamma}} \right)}
\tag{100}
$$

□

V. Conclusion

In order to simplify the stability analysis of the discrete linear time-varying Kalman filter, we presented a stochastic Laypunov method by an orthonormalized compressed measurement. As a result, a stochastic radius of attraction was derived. For the derivation of the stochastic radius of attraction, five Lemmas were introduced to explain preservation of information. Utilizing five Lemmas, three concepts of equivalent measurements were introduced. All the measurements in a specified time interval were stacked to a vector form a stacked measurement. A de-correlation process is applied to the stacked measurement to generate a de-correlated measurement. Finally, the large dimension of the de-correlated measurement is reduced by a weighted pseudo-inverse resulting in an orthonormalized compressed measurement. Afterwards, various matrix inequalities were explained to show the boundness of error covariance matrices and grammians. Finally, a stochastically-driven Lyapunov method is applied to derive the radius of attraction. During the derivation, it was shown that the complex multiple-step propagations a stochastic Lyapunov function candidate driven by measurement and process noises can be simplified to a one-step Lyapunov propagation by the orthonormalized compressed measurement. Since the unique analysis procedure in this study considers the effects of the noise terms correctly, it will also help to understand the physical meanings of grammians and the stability of linear time-varying Kalman filters more better.

Acknowdegements

This work has been supported by the Automatic Control Research Center (ACRC) and Automation Systems Research Institute (ASRI) of Seoul National University and by the Agency for Defence Development (ADD).

Reference

[1] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems", *ASME J. Basic Eng.* Vol. 82, pp. 35-45, March, 1960

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

[3] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Eaglewood Cliffs, NJ: Prentice-Hall, 1979

[4] T. P. McGarty, *Stochastic Systems and State Estimation*, New York : Wiley, 1973

[5] P. S. Maybeck, *Stochastic Models, Estimation and Control*, Vol. 1, Academic Press, 1979

[6] R. G. Agniel and E. I. Jury, "Almost sure boundedness of randomly sampled systems", *SIAM J. Contr*., vol. 9, pp. 372-384, 1971

[7] T. J. Tarn and Y. Rasis, "Observers for nonlinear stochastic systems", *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 441-448, 1976

[8] K. Reif, S. Günthe, E. Yaz, R. Unbehauen, "Stochastic Stability of the Discrete-Time Extended Kalman Filter", *IEEE Trans. Automat. Contr*., Vol. 44, No. 4, pp. 714-728, April, 1999