

Stability Analysis of Kalman Filter by Orthonormalized Compressed Measurement

Hyung Keun Lee and Jang Gyu Lee

Abstract - In this paper, we propose the concept of orthonormalized compressed measurement for the stability analysis of discrete linear time-varying Kalman filters. Unlike previous studies that deal 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 with 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 impractically difficult to obtain by the conventional multiple-step Lyapunov method.

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

1. 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 find 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 research on the subject has been 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 bounded-output (BIBO) stability are provided for the stability of stochastically-driven systems.

Since the stability of stochastically driven systems is difficult to analyze, no detailed and practical Lyapunov analysis result has 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. This non-singularity 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

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

Manuscript received: March 15, 2002 accepted: July 29, 2002

Hyung Keun Lee and Jang Gyu Lee are with the School of Electrical Engineering and Computer Science and the Automatic Control Research Center, Seoul National University Korea.

propagation of a candidate Lyapunov function is intractably complex. To extend previous work on 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 for obtaining the OCM is summarized in Fig. 1. The single stacked measurement vector $Z_{k/k-N+1}^s$ is obtained if we merely stack all the measurements $\{z_j\}$ at a specified time interval where $j = k - N + 1, k - N + 2, \dots, k$. The de-correlated stacked measurement $Z_{k/k-N+1}^d$ is formed to eliminate the correlation between the stacked measurement $Z_{k/k-N+1}^s$ and the estimation error of the KF. To reduce the large dimension of $Z_{k/k-N+1}^d$, a pseudo-inverse is pre-multiplied to $Z_{k/k-N+1}^d$, resulting in the OCM for obtaining $Z_{k/k-N+1}^n$.

The OCM provides information on the *a priori* state estimate at the k -th step that is propagated from a *a posteriori* state estimate at the $(k-N)$ -th step. The proposed OCM is advantageous in analysis in four ways: (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 *a posteriori* state estimate at the $(k-N)$ -th step, (iii) the dimension is equal to the system dimension, and (iv) the OCM is closely related with the observability grammian and can be utilized for proving the stability of stochastically-driven systems.

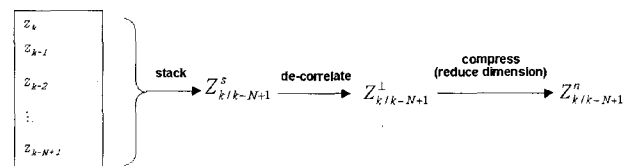


Fig. 1 Equivalent measurements for stability analysis

2. Preservation of information

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

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

and the direct measurement,

$$y_k = h_k x_k - v_k, \quad v_k \sim (O, rI), \quad r > 0, \quad (2)$$

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

$$\begin{aligned} E[v_j(v_k)] &= O, \quad \forall j, k, \quad j, k = 1, 2, 3, \dots \\ E[w_j(w_k)] &= O, \quad \forall j \neq k, \quad j, k = 1, 2, 3, \dots \\ E[w_j(v_k)] &= O, \quad j, k = 1, 2, 3, \dots \end{aligned} \quad (3)$$

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

$$\begin{aligned} \hat{x}_{k+1/k} &= F_{k+1/k} \hat{x}_k \\ e_{k+1/k} &= F_{k+1/k} e_k + G_k w_k \\ P_{k+1/k} &= F_{k+1/k} P_k F_{k+1/k}^T + G_k q G_k^T \end{aligned} \quad (4)$$

and the measurement update step,

$$\begin{aligned} K_k &= P_{k/k-1} h_k^T (h_k P_{k/k-1} h_k^T + r_k)^{-1} \\ \hat{x}_k &= \hat{x}_{k/k-1} - K_k z_k^* \\ e_k &= (I - K_k h_k) e_{k/k-1} - K_k v_k \\ P_k &= (I - K_k h_k) P_{k/k-1} (I - K_k h_k)^T + K_k r_k K_k^T \end{aligned} \quad (5)$$

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 : 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.

With the description of the models and the KF, we introduce five Lemmas 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 equal to 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 here. Throughout the Lemmas, $Est[x]$ is used to denote an estimation of x where the optimality is not guaranteed and $E(x)$ 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 $e^- \in \mathbf{R}^{n \times 1}$, $M \in \mathbf{R}^{m \times n}$, $z \in \mathbf{R}^{m \times 1}$, $H \in \mathbf{R}^{m \times n}$, $R \in \mathbf{R}^{m \times m}$, and $S \in \mathbf{R}^{m \times n}$ such that

$$\begin{aligned} \hat{x}^- &= Est[x] \\ M &= E[(e^-)(e^-)^T] \end{aligned} \quad (6)$$

where

$$\begin{aligned} e^- &= \hat{x}^- - x \\ z &= H e^- + v, \quad v \sim (O_{m \times 1}, R) \\ S &= E[v(e^-)^T]. \end{aligned} \quad (7)$$

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

$$\begin{aligned} \bar{z} &:= Cz = \bar{H} e^- + \bar{v}, \quad \bar{v} \sim (O_{m \times 1}, \bar{R}) \\ \bar{H} &:= CH \\ \bar{R} &:= E[\bar{v} \bar{v}^T] = CRC^T \\ \bar{S} &:= E[\bar{v} (e^-)^T] = CS. \end{aligned} \quad (8)$$

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 $e^- \in \mathbf{R}^{n \times 1}$, $M \in \mathbf{R}^{m \times n}$, $z \in \mathbf{R}^{m \times 1}$, $H \in \mathbf{R}^{m \times n}$, $R \in \mathbf{R}^{m \times m}$, and $S \in \mathbf{R}^{m \times n}$ satisfying Eqs. (6) and (7). In addition, $R - SM^{-1}S^T > O$ is satisfied. Then, no information is lost in obtaining the *a posteriori* estimation \hat{x}^+ of x by removing the 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

$$\begin{aligned}\bar{z} &:= \bar{H}e^- + \bar{v}, \bar{v} \sim (O_{n \times 1}, \bar{R}) \\ \bar{H} &:= H + SM^{-1} \\ \bar{R} &:= R - SM^{-1}S^T \\ E[\bar{v}\bar{v}^T] &= O_{m \times m}.\end{aligned}\quad (9)$$

Lemma 2-3: Preservation of information in dimension reduction [2, 5]

Suppose that we are given *a priori* estimation $\hat{x}^- \in \mathbf{R}^{n \times 1}$ 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, $m \geq n$, $H^T H > O$, $S = O$ are satisfied. 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

$$\begin{aligned}\bar{z} &:= C^* z = \bar{H}e^- + \bar{v}, \bar{v} \sim (O_{n \times 1}, \bar{R}) \\ C^* &:= (H^T R^{-1} H)^{-1} H^T R^{-1} \\ \bar{H} &:= I_{n \times n} \\ \bar{R} &:= E[vv^T] = (H^T R^{-1} H)^{-1}.\end{aligned}\quad (10)$$

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

$$\begin{aligned}\hat{x}_k^- &= Est[x_k] \\ z_k &= H_k e_k^- + v_k, v_k \sim (O, R_k) \\ e_k^- &:= \hat{x}_k^- - x_k \\ S_k &= E[v_k (e_k^-)^T]\end{aligned}\quad (11)$$

where

$$\begin{aligned}R_k - S_k M_k^{-1} S_k^T &> O \\ M_k &:= E[e_k^- (e_k^-)^T].\end{aligned}\quad (12)$$

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

$$\hat{x}_{k+1}^- = F_{k+1/k} \hat{x}_k^- \quad (13)$$

where

$$w_k \sim (O, q_k), E[w_k v_k^T] = O, E[w_k e_k^-] = O. \quad (14)$$

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

$$\bar{z}_{k+1} := z_k = \bar{H}_{k+1} e_{k+1}^- + \bar{v}_{k+1} \quad (15)$$

where

$$\begin{aligned}e_{k+1}^- &:= \hat{x}_{k+1}^- - x_{k+1} \\ \bar{H}_{k+1} &:= H_k (F_{k+1/k})^{-1} \\ \bar{R}_{k+1} &:= E[\bar{v}_{k+1} (\bar{v}_{k+1})^T] = R_k + \bar{H}_{k+1} G_k q_k G_k^T (\bar{H}_{k+1})^T\end{aligned}$$

$$\bar{S}_{k+1} := E[\bar{v}_{k+1} (e_{k+1}^-)^T] = S_k (F_{k+1/k})^T - \bar{H}_{k+1} G_k q_k G_k^T. \quad (16)$$

Proof: By Eqs. (1) and (13), the time-propagation of estimation error from e_k^- to e_{k+1}^- is derived as

$$e_{k+1}^- = F_{k+1/k} e_k^- + G_k w_k. \quad (17)$$

According to Eqs. (11) and (17), the given z_k satisfies the following relationship with e_{k+1}^-

$$z_k = H_k (F_{k+1/k})^{-1} e_{k+1}^- + v_k - H_k (F_{k+1/k})^{-1} G_k w_k. \quad (18)$$

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

$$\bar{v}_{k+1} := v_k - H_k (F_{k+1/k})^{-1} G_k w_k, \quad (19)$$

we obtain the result. In addition to the interpretation of Z_k by \bar{Z}_{k+1} in Eqs. (15) and (16), various equivalent interpretations can be obtained by applying Lemmas 2-1, 2-2, and 2-3.

3. Equivalent batch measurements

Given \hat{x}_{k-N} , P_{k-N} , $\{w_j\}_{k-N \leq j < k}$, $\{z_j\}_{k-N < j \leq k}$, and $\{v_j\}_{k-N+1 < j \leq k}$ 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 $\hat{x}_{k-N/k-N}$, $P_{k-N/k-N}$, $\{w_j\}_{k-N \leq j < k}$, $\{z_j\}_{k-N < j \leq k}$, and $\{v_j\}_{k-N+1 < j \leq k}$.

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

$$\hat{x}_{k/k-N} = F_{k/k-N} \hat{x}_{k-N/k-N} \quad (20)$$

$$\begin{aligned}e_{k/k-N} &= F_{k/k-N} e_{k-N/k-N} + W_{k-1/k-N} \\ &= F_{k/k-N} e_{k-N/k-N} + (G_{k-1/k-N}^o)^T w_{k-1/k-N}\end{aligned}$$

$$M_k = F_{k/k-N} P_{k-N} F_{k/k-N}^T + \Xi_{k-1/k-N} \quad (21)$$

where

$$W_{k-1/k-N} = \sum_{j=k-N}^{k-1} F_{k/j} G_j w_j = (G_{k-1/k-N}^o)^T w_{k-1/k-N}$$

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

$$\Xi_{k-1/k-N} = G_{k-1/k-N}^o Q_{k-1/k-N}^o (G_{k-1/k-N}^o)^T$$

$$G_{k-1/k-N}^o := [G_{k-1} \mid F_{k/k-1} G_{k-2} \mid F_{k/k-2} G_{k-3} \mid \cdots \mid F_{k/k-N+1} G_{k-N}]$$

$$w_{k-1/k-N}^o := \begin{bmatrix} w_{k-1} \\ \vdots \\ w_{k-2} \\ \vdots \\ w_{k-3} \\ \vdots \\ w_{k-N} \end{bmatrix}$$

$$w_{k-1/k-N}^o \sim (O, Q_{k-1/k-N}^o)$$

$$Q_{k-1/k-N}^o = qI > 0. \quad (22)$$

Note that $\Xi_{k-1/k-N}$ in Eq. (22) 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 $Z_{k/k-N+1}^s$ in addition to the *a priori* estimation $\hat{x}_{k/k-N}$ of x_{k-N} .

$$Z_{k/k-N+1}^s := \begin{bmatrix} z_k \\ \vdots \\ z_{k-1} \\ \vdots \\ z_{k-2} \\ \vdots \\ z_{k-N+1} \end{bmatrix} \quad (23)$$

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

By the first three theorems given in Section 2, it is possible to find an $n \times Nm$ optimal transformation matrix $C_{k/k-N+1}^s$ from an $Nm \times 1$ stacked measurement vector $Z_{k/k-N+1}^s$ to an $n \times 1$ single equivalent measurement vector $Z_{k/k-N+1}^n$ without any information loss:

$$Z_{k/k-N+1}^n = C_{k/k-N+1}^s Z_{k/k-N+1}^s. \quad (24)$$

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

$$\begin{aligned} \hat{x}_k &= \hat{x}_{k/k-N} - K_{k/k-N+1}^n Z_{k/k-N+1}^n \\ &= \hat{x}_{k/k-N} - K_{k/k-N+1}^n C_{k/k-N+1}^s Z_{k/k-N+1}^s \end{aligned} \quad (25)$$

where $K_{k/k-N+1}^n$ denotes the optimal gain related with $Z_{k/k-N+1}^n$. 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 $k = jN, j = 0, 1, 2, 3, \dots$ according to the Lemmas in Section 2.

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:

$$z_j = h_j F_{j/k-N} e_{k/k-N} - \sum_{a=j}^{k-1} h_j F_{j/a+1} G_a w_a + v_j. \quad (26)$$

3.1 Stacked measurement

A single stacked measurement vector $Z_{k/k-N+1}^s$ 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 $Z_{k/k-N+1}^s$ can be represented by the following single vector equation,

$$\begin{aligned} Z_{k/k-N+1}^s &= H_{k/k-N+1}^s e_{k/k-N} + v_{k/k-N+1}^o - H_{k-1/k-N}^o W_{k-1/k-N}^o \\ &= H_{k/k-N+1}^s e_{k/k-N} + v_{k/k-N+1}^o \end{aligned} \quad (27)$$

where

$$Z_{k/k-N+1}^s := \begin{bmatrix} z_k \\ \vdots \\ z_{k-1} \\ \vdots \\ z_{k-2} \\ \vdots \\ z_{k-N+1} \end{bmatrix} \quad H_{k/k-N+1}^s := \begin{bmatrix} h_k \\ \vdots \\ h_{k-1} F_{k-1/k} \\ \vdots \\ h_{k-2} F_{k-2/k} \\ \vdots \\ h_{k-N+1} F_{k-N+1/k} \end{bmatrix}$$

$$v_{k/k-N+1}^o := \begin{bmatrix} v_k \\ \vdots \\ v_{k-1} \\ \vdots \\ v_{k-2} \\ \vdots \\ v_{k-N+1} \end{bmatrix} \quad v_{k/k-N+1}^o \sim (O, R_{k/k-N+1}^o)$$

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

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

$$\begin{aligned} v_{k/k-N+1}^s &:= v_{k/k-N+1}^o - H_{k-1/k-N}^o W_{k-1/k-N}^o \\ v_{k/k-N+1}^s &\sim (O, R_{k/k-N+1}^s) \\ R_{k/k-N+1}^s &:= E\left\{ \left(v_{k/k-N+1}^s \right) \left(v_{k/k-N+1}^s \right)^T \right\} \\ &= R_{k/k-N+1}^o + H_{k-1/k-N}^o Q_{k-1/k-N}^o \left(H_{k-1/k-N}^o \right)^T \\ S_{k/k-N+1}^s &:= E\left\{ \left(v_{k/k-N+1}^s \right) \left(e_{k/k-N} \right)^T \right\} \\ &= -H_{k-1/k-N}^o Q_{k-1/k-N}^o \left(G_{k-1/k-N}^o \right)^T. \end{aligned} \quad (28)$$

As shown in Eqs. (27) and (28), the measurement noise $v_{k/k-N+1}^o$ of the stacked measurement $Z_{k/k-N+1}^s$ is composed of two terms, i.e., the noise term $v_{k/k-N+1}^o$, which is uncorrelated with the *a priori* estimation error $e_{k/k-N}$, and the noise term $H_{k-1/k-N}^o W_{k-1/k-N}^o$, which is correlated with the *a priori* estimation error $e_{k/k-N}$. Since a sequence of process noise $\{w_j\}_{k-N \leq j < k}$ generates a correlation between the stacked measurement noise $v_{k/k-N+1}^o$ and the *a priori* estimation error $e_{k/k-N}$ of $\hat{x}_{k/k-N} = E[x_k | y_1, y_2, \dots, y_{k-N}]$ as shown in Eq. (28), the optimal *a posteriori* estimation $\hat{x}_k = E[x_k | y_1, y_2, \dots, y_k]$ of x_k using all the measurement up to the k -th step can be obtained by a somewhat complicated computation:

$$\begin{aligned} K_{k/k-N+1}^s &= \\ & \left[M_k \left(H_{k/k-N+1}^s \right)^T + \left(S_{k/k-N+1}^s \right)^T \right] \left[\left(H_{k/k-N+1}^s \right) M_k \left(H_{k/k-N+1}^s \right)^T \right. \\ & \quad \left. + \left(S_{k/k-N+1}^s \right) + \left(S_{k/k-N+1}^s \right)^T + \left(R_{k/k-N+1}^s \right) \right]^{-1} \\ P_k &= M_k - K_{k/k-N+1}^s \left[\left(H_{k/k-N+1}^s \right) M_k + \left(S_{k/k-N+1}^s \right) \right] \end{aligned}$$

$$\hat{x}_k = \hat{x}_{k/k-N} - K_{k/k-N+1}^s Z_{k/k-N+1}^s e_{k/k-N} \quad (29)$$

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

$$e_k = [I - K_{k/k-N+1}^s H_{k/k-N+1}^s] e_{k/k-N} - K_{k/k-N+1}^s v_{k/k-N+1}^s \quad (30)$$

3.2 De-correlated stacked measurement

As shown in Eqs. (27) and (28), the correlation between the stacked measurement noise $v_{k/k-N+1}^s$ 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 $Z_{k/k-N+1}^\perp$ is formed as

$$Z_{k/k-N+1}^\perp = Z_{k/k-N+1}^s = H_{k/k-N+1}^\perp e_{k/k-N} + v_{k/k-N+1}^\perp \quad (31)$$

where

$$H_{k/k-N+1}^\perp = H_{k/k-N+1}^s + S_{k/k-N+1}^s (M_k)^{-1} = \begin{bmatrix} h_k \\ \dots \\ h_{k-1} F_{k-1/k}^\perp \\ \dots \\ h_{k-2} F_{k-2/k}^\perp \\ \dots \\ h_{k-N} F_{k-N+1/k}^\perp \end{bmatrix}$$

$$v_{k/k-N+1}^\perp = -S_{k/k-N+1}^s (M_k)^{-1} e_{k/k-N} + v_{k/k-N+1}^s$$

$$S_{k/k-N+1}^s := E\left\{ (v_{k/k-N+1}^s)(e_{k/k-N})^T \right\}$$

$$= -S_{k/k-N+1}^s + S_{k/k-N+1}^s = O$$

$$R_{k/k-N+1}^\perp := E\left\{ (v_{k/k-N+1}^\perp)(v_{k/k-N+1}^\perp)^T \right\}$$

$$= R_{k/k-N+1}^s - S_{k/k-N+1}^s (M_k)^{-1} (S_{k/k-N+1}^s)^T$$

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

$$\Xi_{k/l} := \sum_{j=l}^k F_{k+1/j+1} G_j q_j (F_{k+1/j+1} G_j)^T \quad (32)$$

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. (32), it will simplify the form of measurement update equations of Eqs. (29) and (30) as follows.

$$K_{k/k-N+1}^\perp = M_k (H_{k/k-N+1}^\perp)^T \left[(H_{k/k-N+1}^\perp)^T M_k (H_{k/k-N+1}^\perp)^T + (R_{k/k-N+1}^\perp) \right]^{-1}$$

$$P_k = M_k - K_{k/k-N+1}^\perp H_{k/k-N+1}^\perp M_k$$

$$\hat{x}_k = \hat{x}_{k/k-N} - K_{k/k-N+1}^\perp Z_{k/k-N+1}^\perp \quad (33)$$

By performing the computation of Eq. (33), the *a posteriori* estimation error e_k of \hat{x}_k satisfies

$$e_{k/k} = [I - K_{k/k-N+1}^\perp H_{k/k-N+1}^\perp] e_{k/k-N} - K_{k/k-N+1}^\perp v_{k/k-N+1}^\perp \quad (34)$$

3.3 Orthonormalized compressed measurement

Though the de-correlated stacked measurement $Z_{k/k-N+1}^\perp$ obtained in Section 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. (32) 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.

$$Z_{k/k-N+1}^n := (H_{k/k-N+1}^\perp)^+ Z_{k/k-N+1}^\perp$$

$$= e_{k/k-N} + v_{k/k-N+1}^n \quad (35)$$

where

$$(H_{k/k-N+1}^\perp)^+ := \left[(H_{k/k-N+1}^\perp)^T (R_{k/k-N+1}^\perp)^{-1} (H_{k/k-N+1}^\perp) \right]^{-1}$$

$$\left((H_{k/k-N+1}^\perp)^T (R_{k/k-N+1}^\perp)^{-1} \right)^{-1}$$

$$v_{k/k-N+1}^n := (H_{k/k-N+1}^\perp)^+ v_{k/k-N+1}^\perp$$

$$v_{k/k-N+1}^n \sim (O, R_{k/k-N+1}^n)$$

$$R_{k/k-N+1}^n = (H_{k/k-N+1}^\perp)^+ R_{k/k-N+1}^\perp \left[(H_{k/k-N+1}^\perp) \right]^T$$

$$= \left[(H_{k/k-N+1}^\perp)^T (R_{k/k-N+1}^\perp)^{-1} (H_{k/k-N+1}^\perp) \right]^{-1}$$

$$S_{k/k-N+1}^n := E\left\{ (v_{k/k-N+1}^n)(e_{k/k-N})^T \right\} = O \quad (36)$$

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

- i) its measurement coefficient matrix is a simple identity matrix;
- ii) its measurement noise is uncorrelated with $e_{k/k-N}$;
- iii) its dimension is minimal, i.e., equal to the system dimension;
- iv) it is closely related with the observability grammian.

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

$$K_{k/k-N+1}^n = M_k \left[M_k + R_{k/k-N+1}^n \right]^{-1}$$

$$P_k = M_k - K_{k/k-N+1}^n M_k$$

$$\hat{x}_{k/k} = \hat{x}_{k/k-N} - K_{k/k-N+1}^n Z_{k/k-N+1}^n \quad (37)$$

By performing Eq. (37), the *a posteriori* estimation error e_k of \hat{x}_k satisfies

$$e_{k/k} = [I - K_{k/k-N+1}^n] e_{k/k-N} - K_{k/k-N+1}^n v_{k/k-N+1}^n \quad (38)$$

4. 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

$$\Theta_{k/k-N+1} = (H_{k/k-N+1}^s)^T (H_{k/k-N+1}^s) \quad (39)$$

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 $\Theta_{k/k-N+1}^\perp$ is defined as

$$\Theta_{k/k-N+1}^\perp := (H_{k/k-N+1}^\perp)^T (R_{k/k-N+1}^\perp)^{-1} (H_{k/k-N+1}^\perp). \quad (40)$$

As shown, the orthogonalized observability grammian $\Theta_{k/k-N+1}^\perp$ 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. (22).

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 $\underline{\theta}$, $\bar{\theta}$, $\underline{\xi}$, and $\bar{\xi}$ such that

$$\begin{aligned} 0 < \underline{\theta} I \leq \Theta_{k/k-N+1} \leq \bar{\theta} I, \quad \forall k \geq N, \\ 0 < \underline{\xi} I \leq \Xi_{k-1/k-N} \leq \bar{\xi} I, \quad \forall k \geq N. \end{aligned} \quad (41)$$

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

$$\begin{aligned} 0 < \underline{f} I \leq F_{k/j} \leq \bar{f} I, \quad k \geq N, \quad j = k-N, k-N+1, \dots, k-1 \\ \|G_k\|_\infty \leq \bar{g}, \quad \|h_k\|_\infty \leq \bar{h}, \quad k = 1, 2, 3, \dots \end{aligned} \quad (42)$$

Lemma 4-1: Boundedness of error covariance matrices [2~4]

Assuming that $P_0 > 0$, the given system is uniformly completely controllable and observable satisfying Eq. (41) and the system matrices are bounded satisfying Eq. (42). Then, the solutions of the discrete Riccati recursion satisfying

$$\begin{aligned} M_k &= F_{k/k-N} P_{k-N} F_{k/k-N}^T + \Xi_{k-1/k-N} \\ P_k^{-1} &= M_k^{-1} + (H_{k/k-N+1}^\perp)^T (R_{k/k-N+1}^\perp)^{-1} (H_{k/k-N+1}^\perp)^{-1} \end{aligned} \quad (43)$$

are bounded by the inequalities

$$\begin{aligned} 0 < \underline{p} I \leq P_k \leq \bar{p} I, \quad \forall k = N, 2N, 3N, \dots \\ 0 < \underline{m} I \leq M_k \leq \bar{m} I, \quad \forall k = 2N, 3N, \dots \end{aligned} \quad (44)$$

where

$$\begin{aligned} \underline{p} &:= \frac{r \underline{\xi}}{r + \bar{\theta} \underline{\xi}} > 0, \quad \bar{p} := \frac{r}{\underline{\theta}} + \bar{\xi} \\ \underline{m} &:= \underline{f}^2 \underline{p} + \underline{\xi} = \frac{[(1 + \underline{f}^2)r + \bar{\theta} \underline{\xi}] \underline{\xi}}{r + \bar{\theta} \underline{\xi}} > 0 \\ \bar{m} &:= \bar{f}^2 \bar{p} + \bar{\xi} = \frac{\bar{f}^2 r + (1 + \bar{f}^2) \bar{\theta} \bar{\xi}}{\underline{\theta}}, \end{aligned} \quad (45)$$

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. (42), then

$$\sigma(H_{k/k-N+1}^\circ) \leq h_o, \quad (46)$$

where

$$h_o := 0.7 N \bar{h} \bar{f} \bar{g}. \quad (47)$$

Proof:

$$\begin{aligned} \|H_{k/k-N+1}^\circ\|_\infty &\leq \left\| \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ \|h_{k-1} F_{k-1/k} G_{k-1}\|_\infty & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \|h F_{k-N+2/k} G_{k-1}\|_\infty & \|h F_{k-N+2/k-1} G_{k-2}\|_\infty & \dots & 0 & 0 \\ \|h F_{k-N+1/k} G_{k-1}\|_\infty & \|h F_{k-N+1/k-1} G_{k-2}\|_\infty & \dots & \|h F_{k-N+1/k-N+2} G_{k-N+1}\|_\infty & 0 \end{bmatrix} \right\|_\infty \\ &\leq \bar{h} \bar{f} \bar{g} \left\| \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \dots & 0 & 0 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix} \right\|_\infty \leq 0.7 N \bar{h} \bar{f} \bar{g} \end{aligned} \quad (48)$$

Remark: $H_{k/k-N+1}^\circ$ used in Eq. (48) represents the correlation between the stacked measurement $Z_{k/k-N+1}^s$ and the *a priori* estimation error $e_{k/k-N}$. Lemma 4-2 shows that the maximum singular value of $H_{k/k-N+1}^\circ$ 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

$$P_{k-N} > 0, \quad (49)$$

then,

$$\begin{aligned} R_{k/k-N+1}^\circ < R_{k/k-N+1}^\perp = \\ R_{k/k-N+1}^s - (S_{k/k-N+1}^s) M_k^{-1} (S_{k/k-N+1}^s)^T \leq R_{k/k-N+1}^s. \end{aligned} \quad (50)$$

Proof: i) By definition, it is obvious that

$$\begin{aligned} R_{k/k-N+1}^\perp = \\ R_{k/k-N+1}^s - (S_{k/k-N+1}^s) M_k^{-1} (S_{k/k-N+1}^s)^T \leq R_{k/k-N+1}^s. \end{aligned} \quad (51)$$

ii) If $P_{k-N} > 0$ and since $Q_{k-1/k-N}^\circ > 0$ by definition, we have

$$\begin{aligned} M_k &= F_{k/k-N} P_{k-N} F_{k/k-N}^T + G_{k-1/k-N}^\circ Q_{k-1/k-N}^\circ (G_{k-1/k-N}^\circ)^T \\ &> G_{k-1/k-N}^\circ Q_{k-1/k-N}^\circ (G_{k-1/k-N}^\circ)^T \\ &\Rightarrow Q_{k-1/k-N}^\circ + Q_{k-1/k-N}^\circ (G_{k-1/k-N}^\circ)^T \\ &\quad [M_k - G_{k-1/k-N}^\circ Q_{k-1/k-N}^\circ (G_{k-1/k-N}^\circ)^T]^{-1} G_{k-1/k-N}^\circ Q_{k-1/k-N}^\circ > 0 \\ &\Rightarrow [Q_{k-1/k-N}^\circ]^{-1} - (G_{k-1/k-N}^\circ)^T M_k^{-1} (G_{k-1/k-N}^\circ) > 0 \\ &\Rightarrow [Q_{k-1/k-N}^\circ]^{-1} - (G_{k-1/k-N}^\circ)^T M_k^{-1} (G_{k-1/k-N}^\circ) > 0 \\ &\Rightarrow Q_{k-1/k-N}^\circ - Q_{k-1/k-N}^\circ (G_{k-1/k-N}^\circ)^T M_k^{-1} (G_{k-1/k-N}^\circ) Q_{k-1/k-N}^\circ > 0 \end{aligned}$$

$$\begin{aligned}
&\Rightarrow R_{k/k-N+1}^o + \{ H_{k-1/k-N}^o Q_{k-1/k-N}^o (H_{k-1/k-N}^o)^T \\
&\quad - H_{k-1/k-N}^o Q_{k-1/k-N}^o (G_{k-1/k-N}^o)^T \cdot \\
&\quad M_k^{-1} G_{k-1/k-N}^o Q_{k-1/k-N}^o (H_{k-1/k-N}^o)^T \} > R_{k/k-N+1}^o \\
&\Rightarrow R_{k/k-N+1}^s - H_{k-1/k-N}^o Q_{k-1/k-N}^o (G_{k-1/k-N}^o)^T \cdot \\
&\quad M_k^{-1} G_{k-1/k-N}^o Q_{k-1/k-N}^o (H_{k-1/k-N}^o)^T > R_{k/k-N+1}^o \\
&\Rightarrow R_{k/k-N+1}^s - (S_{k/k-N+1}^s)^T M_k^{-1} (S_{k/k-N+1}^s)^T > R_{k/k-N+1}^o \\
&\Rightarrow R_{k/k-N+1}^s > R_{k/k-N+1}^o. \tag{52}
\end{aligned}$$

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

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.

$$\begin{bmatrix} W \\ Z - He \end{bmatrix} \sim \left(O, \begin{bmatrix} \Xi & S^T \\ S & R \end{bmatrix} \right) \tag{53}$$

Then, the following inequalities hold.

$$\begin{aligned}
R - S \Xi^{-1} S^T &> O \\
\Xi - S^T R^{-1} S &> O \tag{54}
\end{aligned}$$

Proof: By definition of error covariance matrices, it is obvious that

$$\begin{bmatrix} \Xi & S^T \\ S & R \end{bmatrix} > O. \tag{55}$$

Since

$$\Xi > O, R > O, \tag{56}$$

the joint error covariance matrix appearing in Eq. (53) can be decomposed as

$$\begin{aligned}
\begin{bmatrix} \Xi & S^T \\ S & R \end{bmatrix} &= \begin{bmatrix} I & O \\ S \Xi^{-1} & I \end{bmatrix} \begin{bmatrix} \Xi & O \\ O & R - S \Xi^{-1} S^T \end{bmatrix} \begin{bmatrix} I & \Xi^{-1} S^T \\ O & I \end{bmatrix} \\
&= \begin{bmatrix} I & S^T R^{-1} \\ O & R \end{bmatrix} \begin{bmatrix} \Xi - S^T R^{-1} S & O \\ O & R \end{bmatrix} \begin{bmatrix} I & O \\ R^{-1} S & I \end{bmatrix}. \tag{57}
\end{aligned}$$

Thus, if the inequalities of Eq. (54) fail to hold, Eq. (55) fails to hold, which results in a contradiction.

Remark: In Lemma 4-4, W 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.

$$S^T R^{-1} S < M \tag{58}$$

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

$$(H + SM^{-1})^T R^{-1} (H + SM^{-1}) \geq \frac{\varepsilon}{1 + \varepsilon} H^T R^{-1} H - \varepsilon M^{-1} \tag{59}$$

where all the matrix-dimensions are assumed appropriate.

Proof: Since R is positive definite and symmetric, it can be decomposed as

$$R = R^{\frac{1}{2}} R^{\frac{1}{2}}, \quad R^{\frac{1}{2}} > O. \tag{60}$$

Define X as

$$X := \frac{1}{\sqrt{1 + \varepsilon}} R^{\frac{1}{2}} H + \sqrt{1 + \varepsilon} \frac{1}{2} S M^{-1}. \tag{61}$$

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

$$\begin{aligned}
X^T X &= \frac{1}{1 + \varepsilon} H^T R^{-1} H + (1 + \varepsilon) M^{-1} (S^T R^{-1} S) M^{-1} \\
&\quad + H^T R^{-1} S M^{-1} + M^{-1} S^T R^{-1} H. \tag{62}
\end{aligned}$$

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

$$\begin{aligned}
M^{-1} (S^T R^{-1} S) M^{-1} + H^T R^{-1} S M^{-1} + M^{-1} S^T R^{-1} H \\
\geq -\frac{1}{1 + \varepsilon} H^T R^{-1} H - \varepsilon M^{-1} (S^T R^{-1} S) M^{-1}. \tag{63}
\end{aligned}$$

Utilizing Eqs. (58) and (63), we finally have the result.

Remark: In Lemma 4-5, $(H + SM^{-1})^T R^{-1} (H + SM^{-1})$ represents an intermediate observability grammian 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. (41), which means uniform complete controllability and observability, is satisfied. If there exists an integer N and a positive constant a that satisfies

$$(H_{k/k-N+1}^s)^T (R_{k/k-N+1}^s)^{-1} (H_{k/k-N+1}^s) \geq (\Xi_{k-1/k-N})^{-1} + aI, \tag{64}$$

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

$$\begin{aligned}
(H_{k/k-N+1}^\perp)^T (R_{k/k-N+1}^\perp)^{-1} (H_{k/k-N+1}^\perp) &\geq \frac{1}{\bar{\gamma}} I \\
\bar{\gamma} &:= \frac{2(2 + a\xi)}{(a\xi)^2} > 0 \tag{65}
\end{aligned}$$

where ξ is a constant bound for the controllability grammian shown in Eq. (41).

Proof: Since

$$M_k > \Xi_{k-1/k-N}, \tag{66}$$

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

$$\begin{aligned}
& (H_{k/k-N+1}^\perp)^T (R_{k/k-N+1}^s)^{-1} (H_{k/k-N+1}^\perp) \\
& \geq \frac{\varepsilon}{1+\varepsilon} (H_{k/k-N+1}^s)^T (R_{k/k-N+1}^s)^{-1} (H_{k/k-N+1}^s) - \varepsilon (\Xi_{k-1/k-N})^{-1} \quad (67) \\
& \geq \frac{\varepsilon a}{1+\varepsilon} I - \frac{\varepsilon^2}{1+\varepsilon} (\Xi_{k-1/k-N})^{-1} \\
& \geq \frac{\varepsilon(a-\varepsilon/\xi)}{1+\varepsilon} I
\end{aligned}$$

Thus, if we select ε as

$$\varepsilon = \frac{a\xi}{2}, \quad (68)$$

the following inequality results.

$$(H_{k/k-N+1}^\perp)^T (R_{k/k-N+1}^s)^{-1} (H_{k/k-N+1}^\perp) \geq \frac{(a\xi)^2}{2(2+a\xi)} I > 0 \quad (69)$$

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

Remark: The OCM introduced in Section 3 becomes full column-rank if $(H_{k/k-N+1}^\perp)^T (H_{k/k-N+1}^\perp)$ is non-singular. Lemma 4-6 means that a full column-rank OCM exists if Eq. (64) is satisfied. The physical meaning of Eq. (64) 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. (41) and that the system matrices are bounded satisfying Eq. (42). Then, the orthogonalized observability grammian $\Theta_{k/k-N+1}^\perp$ is upper-bounded by

$$\begin{aligned}
& (H_{k/k-N+1}^\perp)^T (R_{k/k-N+1}^\perp)^{-1} (H_{k/k-N+1}^\perp) \leq \frac{1}{\underline{\gamma}} I \\
\underline{\gamma} & = \frac{1}{2r} \left[\bar{\theta} + \bar{m} \left(\frac{1}{\underline{m}} \right)^2 (r + qh_o^2) \right]^{-1} > 0. \quad (70)
\end{aligned}$$

Proof: By Eq. (28), Lemma 4-2, and Lemma 4-3, the following inequality holds.

$$R_{k/k-N+1}^s \leq (r + qh_o^2) I \quad (71)$$

Due to Eq. (71) and Lemma 4-4, the following inequality also holds.

$$\begin{aligned}
& \frac{1}{r + qh_o^2} (S_{k/k-N+1}^s)^T (S_{k/k-N+1}^s) \leq \quad (72) \\
& (S_{k/k-N+1}^s)^T (R_{k/k-N+1}^s)^{-1} (S_{k/k-N+1}^s) \leq M_k
\end{aligned}$$

By Eqs. (50) and Lemma 4-1, we obtain

$$(S_{k/k-N+1}^s)^T (S_{k/k-N+1}^s) \leq \bar{m} (r + qh_o^2) I. \quad (73)$$

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

$$\begin{aligned}
& (H_{k/k-N+1}^\perp)^T (R_{k/k-N+1}^\perp)^{-1} (H_{k/k-N+1}^\perp) \\
& \leq (H_{k/k-N+1}^\perp)^T (R_{k/k-N+1}^o)^{-1} (H_{k/k-N+1}^\perp) \\
& = r (H_{k/k-N+1}^s + S_{k/k-N+1}^s M_k^{-1})^T (H_{k/k-N+1}^s + S_{k/k-N+1}^s M_k^{-1}) \\
& = r \left[(H_{k/k-N+1}^s)^T (H_{k/k-N+1}^s) + M_k^{-1} (S_{k/k-N+1}^s)^T H_{k/k-N+1}^s \right. \\
& \quad \left. + (H_{k/k-N+1}^s)^T S_{k/k-N+1}^s M_k^{-1} + M_k^{-1} (S_{k/k-N+1}^s)^T S_{k/k-N+1}^s M_k^{-1} \right] \\
& \leq \frac{r(\varepsilon+1)}{\varepsilon} (H_{k/k-N+1}^s)^T (H_{k/k-N+1}^s) \\
& \quad + r(\varepsilon+1) M_k^{-1} (S_{k/k-N+1}^s)^T S_{k/k-N+1}^s M_k^{-1}. \quad (74)
\end{aligned}$$

Letting $\varepsilon = 1$, we have

$$\begin{aligned}
& (H_{k/k-N+1}^\perp)^T (R_{k/k-N+1}^\perp)^{-1} (H_{k/k-N+1}^\perp) \\
& \leq 2r \left[\bar{\theta} + \bar{m} \left(\frac{1}{\underline{m}} \right)^2 (r + qh_o^2) \right]. \quad (75)
\end{aligned}$$

Lemma 4-8: Bounds for orthogonalized observability grammian

Assume that the given system is uniformly completely controllable and uniformly completely observable satisfying Eq. (41). If there exists an integer N and a positive constant a that satisfies

$$(H_{k/k-N+1}^s)^T (R_{k/k-N+1}^s)^{-1} (H_{k/k-N+1}^s) \geq (\Xi_{k-1/k-N})^{-1} + aI, \quad (76)$$

then the error covariance matrix of the OCM is bounded by

$$\underline{\gamma} I \leq R_{k/k-N+1}^n \leq \bar{\gamma} I \quad (77)$$

where the positive bounding constants $\bar{\gamma}$ 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 the 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 , $\nu > 0$, and $0 < \vartheta < 1$, such that

$$E\{\|\zeta_n\|^2\} \leq \eta \|\zeta_0\|^2 \vartheta^{n+\nu} \quad (78)$$

holds for every $n \geq 0$.

Definition 4-2: Boundedness with probability one [6~8]

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

$$\sup_{n \geq 0} \|\zeta_n\| < \infty \quad (79)$$

holds with probability one.

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

Lemma 4-8 [8]

Assume a stochastic process $V_n(e_n)$ and real numbers $\underline{\nu}, \bar{\nu}, \mu > 0$ and $0 < \alpha \leq 1$ such that

$$\underline{\nu} \|e_n\|^2 \leq V_n(e_n) \leq \bar{\nu} \|e_n\|^2 \quad (80)$$

and

$$E\{V_{n+1}(e_{n+1})|e_n\} - V_n(e_n) \leq \mu - \alpha V_n(e_n) \quad (81)$$

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

$$E\{\|e_n\|^2\} \leq \frac{\bar{\nu}}{\underline{\nu}} E\{\|e_0\|^2\} (1-\alpha)^n + \frac{\mu}{\underline{\nu}} \sum_{i=1}^{n-1} (1-\alpha)^i \quad (82)$$

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

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

Theorem: Stochastic radius of attraction of the LTV KF

Consider the discrete LTV system of Eqs. (1)~(3) which is uniformly completely controllable and uniformly completely observable satisfying Eq. (41) and in which the system matrices are bounded satisfying Eq. (42). Given the conditions that

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

$$0 < \underline{\gamma} I \leq R_{k/k-N+1}^n \leq \bar{\gamma} I \text{ holds for all } k \geq N \quad (83)$$

and that

- ii) there exists an integer N and a positive constant a that satisfies Eq. (64),

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

$$\pi = \sqrt{\bar{\gamma} \bar{p} \underline{p} \left(\frac{\bar{\xi}}{\underline{m}} + 2 \frac{\bar{p}}{\underline{\gamma}} \right)} \quad (84)$$

where the positive constants \bar{p} , \underline{p} , \underline{m} , $\bar{\xi}$, $\bar{\gamma}$, and $\underline{\gamma}$ are defined in Eqs. (41), (45), (65), and (70).

Proof: The matrices

$$\begin{aligned} M_k &:= P_{k/k-N} \\ K_{k/k-N+1}^n &:= M_k (M_k + R_{k/k-N+1}^n)^{-1} = P_k (R_{k/k-N+1}^n)^{-1} \\ U_k &:= -K_{k/k-N+1}^n e_{k/k-N} - K_{k/k-N+1}^n v_{k/k-N+1}^n \\ &= -[P_k (R_{k/k-N+1}^n)^{-1}] e_{k/k-N} - [P_k (R_{k/k-N+1}^n)^{-1}] v_{k/k-N+1}^n \\ \Gamma_{k/k-N} &:= F_{k/k-N}^{-1} \\ N_k &:= (F_{k/k-N}^T M_k^{-1} F_{k/k-N})^{-1} \end{aligned} \quad (85)$$

satisfy the relationships

$$\begin{aligned} P_k^{-1} &= M_k^{-1} + (R_{k/k-N+1}^n)^{-1} \\ P_k M_k^{-1} &= I - P_k (R_{k/k-N+1}^n)^{-1} = I - K_{k/k-N+1}^n \\ e_k &= P_k M_k^{-1} e_{k/k-N} - K_{k/k-N+1}^n v_{k/k-N+1}^n \\ &= e_{k/k-N} + (P_k M_k^{-1} - I) e_{k/k-N} - K_{k/k-N+1}^n v_{k/k-N+1}^n \\ e_{k/k-N} &= M_k P_k^{-1} e_k + M_k (R_{k/k-N+1}^n)^{-1} v_{k/k-N+1}^n \\ N_k^{-1} &= F_{k/k-N}^T M_k^{-1} F_{k/k-N} \end{aligned} \quad (86)$$

We also have

$$e_k = e_{k/k-N} + U_k = (I - K_{k/k-N+1}^n) e_{k/k-N} - K_{k/k-N+1}^n v_{k/k-N+1}^n \quad (87)$$

Defining a Lyapunov candidate $V_k(e_k)$ as

$$V_k(e_k) := e_k^T P_k^{-1} e_k, \quad (88)$$

then, by the properties of Eqs. (85)~(87), $V_k(e_k)$ satisfies

$$\begin{aligned} V_k(e_k) &= e_{k/k-N}^T M_k^{-1} e_{k/k-N} - e_k^T (R_{k/k-N+1}^n)^{-1} e_k + 2e_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_{k/k-N+1}^n)^{-1} e_k + 2e_k^T M_k^{-1} e_{k/k-N} \\ &\quad - 2e_k^T P_k^{-1} K_{k/k-N+1}^n v_{k/k-N+1}^n - 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_{k/k-N+1}^n)^{-1} e_k \\ &\quad - (e_k - e_{k/k-N})^T M_k^{-1} (e_k - e_{k/k-N}) - 2e_k^T (R_{k/k-N+1}^n)^{-1} v_{k/k-N+1}^n \\ &= e_{k/k-N}^T M_k^{-1} e_{k/k-N} - e_k^T (R_{k/k-N+1}^n)^{-1} e_k \\ &\quad - (e_k - e_{k/k-N})^T M_k^{-1} (e_k - e_{k/k-N}) - 2e_k^T (R_{k/k-N+1}^n)^{-1} v_{k/k-N+1}^n \\ &= e_{k/k-N}^T M_k^{-1} e_{k/k-N} - e_k^T (R_{k/k-N+1}^n)^{-1} e_k - U_k^T M_k^{-1} U_k \\ &\quad - 2e_{k/k-N}^T (I - K_{k/k-N+1}^n)^T (R_{k/k-N+1}^n)^{-1} v_{k/k-N+1}^n \\ &\quad + 2(v_{k/k-N+1}^n)^T (R_{k/k-N+1}^n)^{-1} P_k (R_{k/k-N+1}^n)^{-1} v_{k/k-N+1}^n \\ &= (F_{k/k-N} e_{k-N} + W_{k-1/k-N})^T M_k^{-1} (F_{k/k-N} e_{k-N} + W_{k-1/k-N}) \\ &\quad - e_k^T (R_{k/k-N+1}^n)^{-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 (I - K_{k/k-N+1}^n)^T (R_{k/k-N+1}^n)^{-1} v_{k/k-N+1}^n \\ &\quad + 2(v_{k/k-N+1}^n)^T (R_{k/k-N+1}^n)^{-1} P_k (R_{k/k-N+1}^n)^{-1} v_{k/k-N+1}^n \\ &\leq e_{k-N}^T P_{k-N}^{-1} e_{k-N} \\ &\quad - e_k^T (R_{k/k-N+1}^n)^{-1} e_k - U_k^T M_k^{-1} U_k \\ &\quad + 2e_{k-N}^T F_{k/k-N}^T M_k^{-1} W_{k-1/k-N} \\ &\quad - 2e_{k-N}^T F_{k/k-N}^T (I - K_{k/k-N+1}^n)^T (R_{k/k-N+1}^n)^{-1} v_{k/k-N+1}^n \\ &\quad - 2W_{k-1/k-N}^T (I - K_{k/k-N+1}^n)^T (R_{k/k-N+1}^n)^{-1} v_{k/k-N+1}^n \\ &\quad + 2(v_{k/k-N+1}^n)^T (R_{k/k-N+1}^n)^{-1} P_k (R_{k/k-N+1}^n)^{-1} v_{k/k-N+1}^n \\ &\quad + W_{k-1/k-N}^T M_k^{-1} W_{k-1/k-N} \\ &\leq V_{k-N}(e_{k-N}) \\ &\quad - e_k^T (R_{k/k-N+1}^n)^{-1} e_k \\ &\quad - U_k^T M_k^{-1} U_k \\ &\quad + W_{k-1/k-N}^T M_k^{-1} W_{k-1/k-N} \\ &\quad + 2(v_{k/k-N+1}^n)^T (R_{k/k-N+1}^n)^{-1} P_k (R_{k/k-N+1}^n)^{-1} v_{k/k-N+1}^n \\ &\quad + 2e_{k-N}^T F_{k/k-N}^T M_k^{-1} W_{k-1/k-N} \\ &\quad - 2e_{k-N}^T F_{k/k-N}^T (I - K_{k/k-N+1}^n)^T (R_{k/k-N+1}^n)^{-1} v_{k/k-N+1}^n \\ &\quad - 2W_{k-1/k-N}^T (I - K_{k/k-N+1}^n)^T (R_{k/k-N+1}^n)^{-1} v_{k/k-N+1}^n \\ &\leq V_{k-N}(e_{k-N}) \\ &\quad - e_k^T (R_{k/k-N+1}^n)^{-1} e_k \\ &\quad + W_{k-1/k-N}^T M_k^{-1} W_{k-1/k-N} \\ &\quad + 2(v_{k/k-N+1}^n)^T (R_{k/k-N+1}^n)^{-1} P_k (R_{k/k-N+1}^n)^{-1} v_{k/k-N+1}^n \\ &\quad + C(k) \end{aligned} \quad (89)$$

where $C(k)$ in Eq. (89) is defined as

$$\begin{aligned} C(k) &:= 2e_{k-N}^T F_{k/k-N}^T M_k^{-1} W_{k-1/k-N} \\ &\quad - 2e_{k-N}^T F_{k/k-N}^T (I - K_{k/k-N+1}^n)^T (R_{k/k-N+1}^n)^{-1} v_{k/k-N+1}^n \\ &\quad - 2W_{k-1/k-N}^T (I - K_{k/k-N+1}^n)^T (R_{k/k-N+1}^n)^{-1} v_{k/k-N+1}^n \end{aligned} \quad (90)$$

By Lemma 4-1 and Lemma 4-8,

$$e_k^T (R_{k/k-N+1}^n)^{-1} e_k \geq \frac{1}{\bar{\gamma}} \|e_k\|^2 \geq \frac{1}{\bar{\gamma} \underline{p}} e_k^T P_k^{-1} e_k = \frac{1}{\bar{\gamma} \underline{p}} V_k(e_k) \quad (91)$$

Substituting Eq. (91) for Eq. (89), we have

$$\begin{aligned} V_k(e_k) &\leq V_{k-N}(e_{k-N}) \\ &\quad - \frac{1}{\bar{\gamma} \underline{p}} V_k(e_k) \\ &\quad + W_{k-1/k-N}^T M_k^{-1} W_{k-1/k-N} \\ &\quad + 2(v_{k/k-N+1}^n)^T (R_{k/k-N+1}^n)^{-1} P_k (R_{k/k-N+1}^n)^{-1} v_{k/k-N+1}^n \\ &\quad + C(k). \end{aligned} \quad (92)$$

Rearranging and scaling the above equation, we obtain

$$\begin{aligned} V_k(e_k) &\leq \frac{\bar{\gamma} \underline{p}}{1 + \bar{\gamma} \underline{p}} V_{k-N}(e_{k-N}) \\ &\quad + \frac{\bar{\gamma} \underline{p}}{1 + \bar{\gamma} \underline{p}} W_{k-1/k-N}^T M_k^{-1} W_{k-1/k-N} \\ &\quad + 2 \frac{\bar{\gamma} \underline{p}}{1 + \bar{\gamma} \underline{p}} (v_{k/k-N+1}^n)^T (R_{k/k-N+1}^n)^{-1} P_k (R_{k/k-N+1}^n)^{-1} v_{k/k-N+1}^n \\ &\quad + \frac{\bar{\gamma} \underline{p}}{1 + \bar{\gamma} \underline{p}} C(k). \end{aligned} \quad (93)$$

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

$$\begin{aligned} E[V_k(e_k) | e_{k-N}] - V_{k-N}(e_{k-N}) &\leq -\frac{1}{1 + \bar{\gamma} \underline{p}} V_{k-N}(e_{k-N}) \\ &\quad + \frac{\bar{\gamma} \underline{p}}{1 + \bar{\gamma} \underline{p}} \cdot \frac{\bar{\xi}}{\underline{m}} \\ &\quad + 2 \frac{\bar{\gamma} \underline{p}}{1 + \bar{\gamma} \underline{p}} \cdot \frac{\bar{p}}{\underline{\gamma}}. \end{aligned} \quad (94)$$

To summarize, we have the inequality

$$E[V_k(e_k) | e_{k-N}] - V_{k-N}(e_{k-N}) \leq -\alpha V_{k-N}(e_{k-N}) + \beta \quad (95)$$

where

$$\alpha := \frac{1}{1 + \bar{\gamma} \underline{p}} > 0, \quad \beta := \frac{\bar{\gamma} \underline{p}}{1 + \bar{\gamma} \underline{p}} \left(\frac{\bar{\xi}}{\underline{m}} + 2 \frac{\bar{p}}{\underline{\gamma}} \right). \quad (96)$$

In addition, from

$$\beta < \frac{\alpha}{\bar{p}} \pi^2 \leq \frac{\alpha}{\bar{p}} \|e_{k-N}\|_2^2 \leq \alpha V_{k-N}(e_{k-N}) \quad (97)$$

we find that the radius of attraction is

$$\pi = \sqrt{\frac{\bar{p}\beta}{\alpha}} = \sqrt{\bar{\gamma} \bar{p} \underline{p} \left(\frac{\bar{\xi}}{\underline{m}} + 2 \frac{\bar{p}}{\underline{\gamma}} \right)}. \quad (98)$$

5. Conclusion

To simplify the stability analysis of the discrete linear time-varying Kalman filter, we presented a stochastic

Lyapunov 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 the five Lemmas, three concepts of equivalent measurements were introduced. All the measurements in a specified time interval were stacked to a vector to form a stacked measurement. A de-correlation process was applied to the stacked measurement to generate a de-correlated measurement. Finally, the large dimension of the de-correlated measurement was 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 was applied to derive the radius of attraction. During the derivation, the complex multiple-step propagations of a stochastic Lyapunov function candidate driven by measurement and process noises was shown to be simplified to a one-step Lyapunov propagation by the orthonormalized compressed measurement. Since the unique analysis procedure in this study considered the effects of the noise terms correctly, this study will also help to better explain the physical meanings of grammians and the stability of linear time-varying Kalman filters.

Acknowledgements

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

References

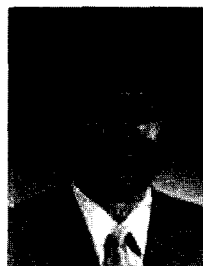
- [1] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems", *ASME J. Basic Eng.*, vol. 82, pp. 35~45, March, 1960.
- [2] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, New York: Academic Press, 1970.
- [3] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Eaglewood 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. Control*, vol. 9, pp. 372~384, 1971.
- [7] T. J. Tam and Y. Rasis, "Observers for nonlinear stochastic systems", *IEEE Transactions on Automatic Control*, vol. AC-21, pp. 441~448, 1976.
- [8] K. Reif, S. Ginthe, E. Yaz, R. Unbehauen, "Stochastic

Stability of the Discrete-Time Extended Kalman Filter", *IEEE Transactions on Automatic Control*, vol. 44, no. 4, pp. 714~728, April, 1999.



Hyung Keun Lee received the B.S. and M.S. degrees in Control and Instrumentation Engineering and Ph.D degree in Electrical Engineering and Computer Science from Seoul National University in 1990, 1994, and 2002, respectively. From 1994 to 1998, he was with the Hyundai Space and Aircraft Corporation. He is currently

holding a Research Fellow Position at the New South Wales University in Australia. His research interests include fusion theory, fault detection, distributed filter architecture, and various hybrid navigation systems.



Jang Gyu Lee received the B.S. degree from Seoul National University in 1971 and the M.S. and Ph.D. degrees from University of Pittsburgh in 1974 and 1977, respectively, all in electrical engineering. From 1977 to 1982, he was with The Analytic Sciences Corporation (TASC), Reading, MA. and with the Charles Stark

Draper Laboratory, Cambridge, MA. In 1982, he joined the faculty of Seoul National University where he has been a full professor at the School of Electrical Engineering and Computer Science since 1992. In addition to his academic duties, he is also the director of the Automatic Control Research Center established by the Ministry of Defense in the Seoul National University. His research interests are the theory and applications of navigation, guidance, and control.