Available online at www.uofk.edu/uofkej

UofKEJ Vol. 2 Issue 1 pp. 41-49 (2013)

UNIVERSITY OF KHARTOUM ENGINEERING JOURNAL (UOFKEJ)

Kalman Filtering with Correlated Noises over Unreliable Communication Network

Khalid A. Mageed Hag Elamin[1], Mirghani Fath Elrahman[2]

[1]*Department of Electrical and Electronic Engineering, College of Engineering, Alnelain University Khartoum, Sudan (Email: elaminkhalid32@yaho.com)*

[2]*Department of Electrical and Electronic Engineering, Faculty of Engineering, University of Khartoum Khartoum, Sudan (Tel.: +249122125301)*

Abstract: In this paper, we study the Kalman filtering problem with correlated noises one time apart, when part or all of the observation measurements are lost in a random way due to unreliable link between controller and estimator. By "one time apart" we mean that the process noise at time step k does not contribute to the system measurement at time step k. Rather, it is the process noise at time step $k-1$ that contributes to the system measurement at time step k. Observation losses can occur in a distributed control system where measurements are taken at different sensors in different physical locations, or one sensor needs to send its data in multiple packets. We formulate the Kalman filtering problem with correlated noises at one time apart and derived the Kalman filter update equations with random observation measurements. Then, we showed that the estimation error covariance (performance index) became a random quantity when the system measurements were collected at different intervals with different probability arrival rate. As a result, we investigated the statistical convergence properties of the performance index and found the stable and unstable regions so that this criterion could be bounded in a stable region and unbounded in an unstable one. The results are illustrated with some simple numerical examples.

Keywords: *Kalman Filtering; State estimation; Correlated Noises; dropping network*

1. INTRODUCTION

Networked control systems (NCSs) i.e., closed –loop feedback control systems that rely on an external separate communication network have received increasing attention over the last few years due to the severe needs for modern network applications [1], [2]. Generally, networked control systems offer many advantages such as reducing the system wiring, making the system easy to operate and maintaining and later diagnose in case of malfunctioning, and increasing system agility [3]. In spite of these advantages, when data are transmitted through these networks, classified as unreliable communication networks, the measurement information packet (from sensors to controllers) and command signals (from controllers to actuators) may be lost or delayed. The packets lost or delayed in this networked control system represent important factors influencing the behavior of the underlying control system [12]. So, inserting a network in between the plant and the controller can cause many potential problems. For instance, in our new system architecture possible finite bandwidth leads to zero delayed sensing and actuation.

In this work, we are specifically interested in the problem of estimation and control across an unreliable communication link that drops information packets. We consider a dynamical process evolving in time that is being observed by a sensor. The sensor needs to transmit the data over a network to a remote destination which can either be an estimator or a controller. The observed data provided by the sensor network are used to estimate the state of a controlled system, and this estimate is then used for control.

We studied the effect of data losses due to the unreliability of the network links. We used the most recursive estimation technique in control – the discrete – time Kalman filter [4], by modeling the arrival of an observation as a random process whose parameters are influenced by the characteristics of the communication link (Fig. 1).

In our setting, we are interested in the statistical convergence properties of the estimation error covariance which represents the expected difference between the actual state and the estimated state that yields as a result of a recursive Kalman filter algorithm.

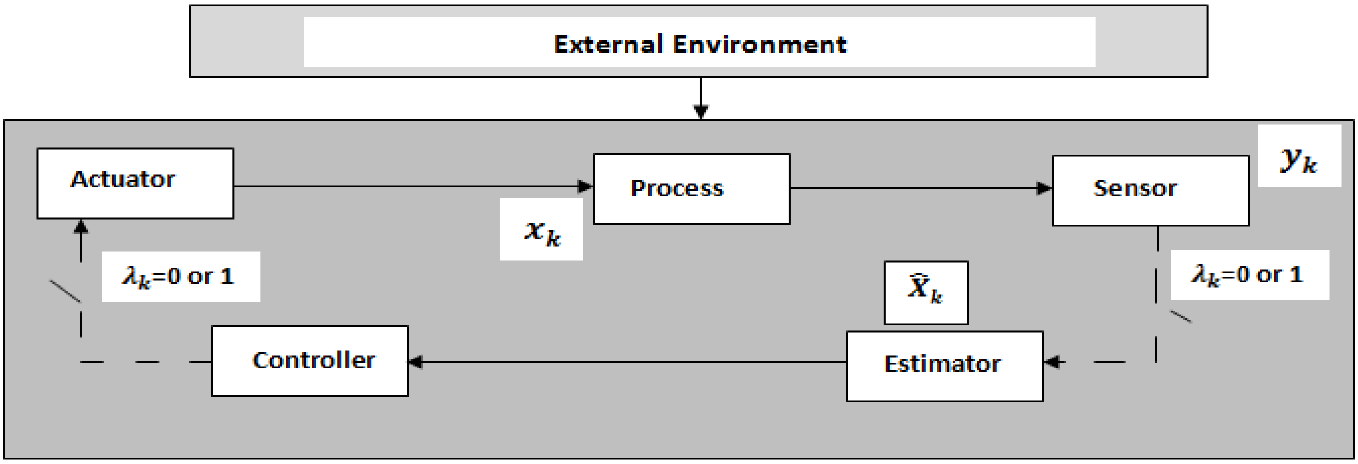

Fig. 1. System block diagram

The classical theory of Kalman filter is based on several assumptions that guarantee convergence of the Kalman filter. Consider the following discrete – time linear system:

$$X_{k+1} = AX_k + w_k \quad (1)$$

$$y_k = CX_k + v_k \quad (2)$$

where X_k is the state vector, y_k the output vector, w_k , v_k are Gaussian random vectors and are independent from each other with zero mean and covariance matrices $Q \geq 0$ and $R > 0$ respectively. w_k is independent from w_s for $s < k$. Assume that the initial state, X_0, is also a Gaussian vector of zero mean and covariance Σ_0. Under the hypothesis of stabilizability of the pair (A, Q) and detectability of the pair (A, C) , the estimation error covariance of the Kalman filter converges to a unique value from any initial condition [5]. These assumptions have been relaxed in various ways. For nonlinear systems, Extended Kalman filter is used [5]; particle filtering [6] is also used for nonlinear models and does not require the noise model to be Gaussian.

Recently, research areas including observation processes have been very important especially in a networked control system. The observation arrival rate in a drop out network is modeled as an identically independent distribution (i.i.d) random process [7]. These authors show the existence of a lower bound on the arrival rate of the observations below which the estimation error covariance is stable and bounded. However, the results are restricted to the case of the independence between system process noise and system measurement noise. We approach a similar problem within the framework of discrete time, and get results for a system with correlated noises at one time step apart.

Our main contribution is to show that, based on the eigenvalues of the system matrix A , the structure of the matrix C , and the correlation between process and measurement noises, there exists a critical observation arrival rate value in a drop out networked control system, μ_c , such that below this value the estimation error covariance is stable and bounded.

Optimal control system is possible for the dynamical system with uncertainty parameters if and only if the uncertainty does not exceed a given threshold. This concept is known as the uncertainty threshold principle [8], [9] with the uncertainty modeled as a white noise sequence acting on the system. In our case, we apply optimal estimation principles in which the uncertainty is due to the randomness arising from losses in the network.

More recently, this problem has been studied using jump linear systems (JLSs) [10]. JLSs are a stochastic system in which the stochastic process is described as a Markov chain so that the uncertainty in the dynamical system is modeled as a two state Markov chain. Following this approach, these authors obtained convergence criteria for the expected estimation error covariance. However, they restricted their formulation to the steady – state case, where the Kalman gain is constant, and they do not assume the correlation between process and measurement noises.

2. PROBLEM FORMULATION

Consider the discrete – time state estimation problem over a communication network as shown in Fig.1. The system dynamic and the measurement (sensor) equations are given as in equations (1) and (2). In this paper we consider two cases. In the first case, when the process noise and measurement noise are subject to Gaussian distribution and independent , and the information packet measurement y_k may be randomly dropped by a network or received by the Kalman filter estimator. An independent identically distributed (i.i.d) Bernoulli random variable λ_k with $E[\lambda_k] = \lambda$ is used to indicate whether y_k is dropped or not, i.e.,

$$\lambda_k = \begin{cases} 0, if\ y_k is\ dropped\ by\ the\ network. \\ 1, if\ y_k is\ received\ by\ the\ estimator \end{cases}$$

In the second case, the process and measurement noises are correlated to each other at one time step apart and the information packet measurement y_k may be randomly dropped by a network or received by the Kalman filter estimator. The Bernoulli random variable λ_k is also used to indicate the dropped and received packet process in this case. In the two cases considered, λ denotes the packet arrival rate at the Kalman filter estimator. In this study we will denote the first case as a packet dropping analysis with independent noises, and the second case as a packet dropping analysis with correlated noises. For both cases, we define the following state quantities at the remote state estimator:

$$\hat{X}_k \triangleq E[X_k | all\ data\ packets\ up\ to\ k],$$
$$P_k \triangleq E[(X_k - \hat{X}_k)(X_k - \hat{X}_k)^T | all\ data\ packets\ up\ to\ k].$$

Later in the next section we will see that P_k is a random quantity due to the randomness of the packet arrival at the estimator.

The question behind our great interest in this problem is how $\hat{X}_k$ approximates to X_k. So, from the above state quantities we are interested in properties of P_k rather than in its exact value. In particular, we are interested in the following problem (from the first case):

2.1 Problem 1 (Packet dropping analysis with independent noises):

With $E[w_k w_j^T] = \delta_{kj} Q \geq 0$, $E[v_k v_j^T] = \delta_{kj} R > 0$, $E[w_k v_j^T] = 0\ \forall j,k$ where δ_{kj} = 0 if $k \neq j$ and δ_{kj} =1 otherwise, find the minimum packet arrival rate λ such that

$$\lim_{k\to\infty} E[P_k | E[w_k v_j^T] = 0] \leq P_{desired} \qquad (3)$$

For a given $P_{desired} \geq 0$.
In the second scenario, we are interested in the following problem:

2.2 Problem 2 (packet dropping analysis with correlated noises):

With $E[w_k w_j^T] = \delta_{kj} Q \geq 0$, $E[v_k v_j^T] = \delta_{kj} R > 0$, $E[w_k v_j^T] = \delta_{kj-1} Z \geq 0$
Where δ_{kj} = 0 if $k \neq j$ and δ_{kj} =1 otherwise, find the minimum packet arrival rate λ such that

$$\lim_{k\to\infty} E[P_k | E[w_k v_j^T] = \delta_{kj-1} Z] \leq P_{desired}$$

For a given $P_{desired} \geq 0$.

3. STANDARD KALMAN FILTER FORMULATION

As in [4] to estimate the state X_k of a discrete –time system stated in equation (1) and (2), using standard Kalman filtering algorithm there are two cycles to be performed the prediction cycle and the filtering cycle. For the prediction cycle the following equations are performed

$$\hat{X}(k+1|k) = A_k \hat{X}(k|k) \qquad (4)$$

$$P(k+1|k) = A_k P(k|k) A_k^T + Q \qquad (5)$$

where P denotes the estimated error covariance. And for the filtering cycle the following equations are performed

$$\hat{X}(k|k) = \hat{X}(k|k-1) + K(k)[y(k) - C_k \hat{X}(k|k-1) \qquad (6)$$

$$K(k) = P(k|k-1) C_k^T [C_k P(k|k-1) C_k^T + R]^{-1} \qquad (7)$$

$$P(k|k) = [I - K(k) C_k]\, P(k|k-1) \qquad (8)$$

where $K(k)$ represents the Kalman Gain.

4. BASIC DIFINITIONS

Before we proceed to analysis we begin with some preliminaries:

4.1 Definitions

The following terms that are frequently used in subsequent sections are defined in this section. Assume that (A, C, Q, R) are the same as in section II. φ_+^n is the set of n by n positive semidefinite matrices. When $X \epsilon\ \varphi_+^n$, we simply write $X \geq 0$; when X is positive definite, we write $X > 0$. We define the function $F : \varphi_+^n \to \varphi_+^n$ as

$$F(X) \triangleq AXA^T + Q \qquad (9)$$

As we shall see shortly, applying F to the previous error covariance matrix corresponds to the update cycle of the standard Kalman filter.
For functions $f_1, f_2 : \varphi_+^n \to \varphi_+^n$, $f_1 \circ f_2$ is defined as

$$f_1 \circ f_2(X) \triangleq f_1(f_2(X)) \qquad (10)$$

Define the function $\tilde{g} : \varphi_+^n \to \varphi_+^n$ as
$$\tilde{g}(X) \triangleq X - XC^T[CXC^T + R]^{-1} CX\,, \qquad (11)$$

And function $g : \varphi_+^n \to \varphi_+^n$ as

$$g(X) \triangleq \tilde{g} \circ F(X) \qquad (12)$$

Denote $\hat{P}_\lambda \geq 0$ as the unique solution to the following equation

$$X = (1-\lambda) F(X) + \lambda g(X) \qquad (13)$$

Therefore $\hat{P}_\lambda$ satisfies $\hat{P}_\lambda = (1-\lambda) F(\hat{P}_\lambda) + \lambda g(\hat{P}_\lambda)$
when λ = 0 or 1 , we obtain $\hat{P}_0$ and $\hat{P}_1$ which satisfy

$$\hat{P}_0 = F(\hat{P}_0) \qquad (14)$$

$$\hat{P}_1 = g(\hat{P}_1) \qquad (15)$$

As stated before, λ_k is independent identically distributed (i.i.d) random variable with $E[\lambda_k] = \lambda$, which describes the arrival rate of measurement y_k. $\hat{P}_0$ in (9) represents the steady state error covariance matrix when $\lambda_k = 0$ for all $k \geq 1$, and $\hat{P}_1$ in (10) is the steady state error covariance matrix when $\lambda_k = 1$ for all $k \geq 1$.

4.2 Kalman filtering with uncorrelated noises

Consider the case when observation measurements are totally received by the estimator or when $\lambda_k = 1$. It is known that the optimal linear estimator for the system described by Eq. (1) and (2) is standard Kalman filter, denoted by KF. We can write equations (4), (5), (6), (7), and (8) in the following compact form: $(\hat{X}_k, P_k) = KF(\hat{X}_{k-1}, P_{k-1}, y_k)$. And with some manipulation, P_k can be shown to satisfy $P_k = g(P_{k-1})$. Therefore, $\hat{P}_1$ defined in equation (15) is the steady state error covariance matrix of the Kalman filter.

Now consider the case when observation measurements are totally dropped, or when λ_k may be equal to 0. Sinopoli *et al*. [11] showed that in this setting, the Kalman filter is still the optimal linear estimator. The two cycles of the standard Kalman filter algorithm take a different formulation, so that in this case only the prediction cycle is performed. But when $\lambda_k = 1$, both the prediction and filtering cycles are performed. The filtering equations are thus the same as the standard Kalman filter equations except for equations (6) and (8) become:

$$\hat{X}(k|k) = \hat{X}(k|k-1) + \lambda_k K(k)[y(k) - C_k \hat{X}(k|k-1) \quad (16)$$

$$P(k|k) = P(k|k-1) - \lambda_k K(k) C_k P(k|k-1) \quad (17)$$

It is obvious, that the randomness of λ_k makes P_k a random quantity.

4.3 Kalman filtering with correlated noises

Here, the goal is to reformulate the standard Kalman filter equations with respect to the fact that the measurement noise and process noise are correlated to each other at one time step apart. Mathematically, the necessary condition is

$E[w_k v_j^T] = \delta_{kj-1} Z \geq 0$ where $\delta_{kj} = 0$ if $k \neq j$ and $\delta_{kj} = 1$ otherwise. With reference to appendix (A) we can see equations (4), (5), (6), (7), and (8) for standard Kalman filtering with uncorrelated noise become:

$$\hat{X}(k+1|k) = A_k \hat{X}(k|k) \quad (18)$$

$$P(k+1|k) = A_k P(k|k) A_k^T + Q \quad (19)$$

Eqn. (18) and (19) represent the Kalman filter prediction cycle, and they are equivalent to equations, (4) and (5). And,

$$\hat{X}(k|k) = \hat{X}(k|k-1) + K(k)[y(k) - C_k \hat{X}(k|k-1) \quad (20)$$

$$K(k) = (P(k|k-1) C_k^T + Z)[C_k P(k|k-1) C_k^T + R + CZ + ZTCT-1 \quad (21)$$

$$P(k|k) = P(k|k-1) - K(k) C_k P(k|k-1) \quad (22)$$

5. PACKET DROPPING ANALYSIS WITH CORRELATED AND UNCORRELATED NOISES

5.1 Analysis with uncorrelated noises

We consider problem 1 for the first scenario. From section (3.1) and (3.2), we can write the error covariance matrix P_k as

$$P_k = \begin{cases} F(P_{k-1}), if\ \lambda_k = 0 \\ g(P_{k-1}), if\ \lambda_k = 1 \end{cases}$$

As seen before, P_k represents a random quantity, so it is difficult to find its exact expected value $E[P_k]$. Therefore, we are looking to its statistical properties to assess the performance of the Kalman filtering process. Before we proceed to analyze problem 1 , we introduce some lemmas that give some properties of the functions F , $\tilde{g}$ and g defined earlier in section IV(A).With reference to appendix (B) we can find the proofs of these lemmas, but here, the table below gives their summary:

From Table 1, X and Y represent positive semidefinite matrices. Now with reference to section (5.1), Since, $P_k = (1-\lambda) F(P_{k-1}) + \lambda g(P_{k-1})$, and as we mentioned before ,the exact value of $E[P_k]$ is difficult to find, so we can bound $E[P_k]$ instead of its exact value. By taking the expectation of the above equation yields:

$$E[P_k] = E[(1-\lambda) F(P_{k-1}) + \lambda g(P_{k-1})]$$
$$= E[(1-\lambda) F(P_{k-1})] + \lambda E[\tilde{g}(F(P_{k-1}))]$$

Hence, by applying lemma (2) from appendix (B), we get:

$$E[P_k] = E[(1-\lambda) F(P_{k-1})] + \lambda E[\tilde{g}(F(P_{k-1}))]$$
$$\leq E[(1-\lambda) F(P_{k-1})] + \lambda \tilde{g}(F(E[P_{k-1}]))\ ]$$
$$= (1-\lambda) F(E[P_{k-1}]) + \lambda g(E[P_{k-1}])$$
$$\therefore\ E[P_k] \leq (1-\lambda) F(E[P_{k-1}]) + \lambda g(E[P_{k-1}])$$

And by induction, it is easy to see that:

$$\lim_{k \to \infty} E[P_k] \leq \hat{P_\lambda}$$

At this point we can say that, If we have a desired error covariance $\bar{P} \geq \hat{P_\lambda}$ and both $\bar{P}$ and $\hat{P_\lambda}$ satisfy the equations:

Table 1: Important lemmas used in this section

No.	Condition	Statement
1	$0 \leq X \leq Y$	$F(X) \leq F(Y)$
2	$0 \leq X \leq Y$	$g(X) \leq g(Y)$
3	$0 \leq X \leq Y$	$g(X) \leq F(X)$
4	$0 \leq X \leq Y$ & $\propto [0,1]$	$\tilde{g}(X)$ $is\ a\ convex\ function$ i.e., $\tilde{g}(\propto X + (1-\propto) Y \leq \propto \tilde{g}(X) + (1-\propto) \tilde{g}(Y)$
5	P_k is a random variable, and $\tilde{g}$ is a convex function	By Jensen's Inequality: $\tilde{g}(E[P_k]) \leq E[\tilde{g}(P_k)]$
6	For any $\lambda_1, \lambda_2 \in [0,1]$	$\hat{P_{\lambda_1}} \geq \hat{P_{\lambda_2}}$, $iff\ \lambda_1 \leq \lambda_2$

$P_{\lambda}^{\wedge} = (1-\lambda)F(P_{\lambda}^{\wedge}) + \lambda g(P_{\lambda}^{\wedge})$, and $\bar{P} = (1-\lambda)F(\bar{P}) + \lambda g(\bar{P})$, we are looking for the minimum arrival rate λ of the state observations such that:

$P_{\lambda}^{\wedge} \leq \bar{P}$, that guarantee with $\lim_{k\to\infty} E[P_k] \leq \bar{P}$,and P_k will be bounded.

Before we proceed to find the value of the arrival rate λ, that satisfies the above statement , we will prove the following new important lemma, lemma 4:

With reference to equation (13) in section (4.1), we can state that: $X \geq (1-\lambda)AXA^T + Q$

Proof: From (9), and (12) in section (4.1) we have:
$F(X) \triangleq AXA^T + Q$ and,
$g(X) = F(X) - F(X)C^T(CF(X)C^T + R)^{-1}CF(X)$

Then (13) become:

$$= (1-\lambda)AXA^T + Q + \lambda[F(X) - F(X)C^T(CF(X)C^T + R)^{-1}CF(X)]$$
$$\geq (1-\lambda)(AXA^T + Q) + \lambda Q = (1-\lambda)(AXA^T) + (1-\lambda)(Q) + \lambda Q$$
$$= (1-\lambda)AXA^T + Q - \lambda Q + \lambda Q$$

Therefore: $X \geq (1-\lambda)AXA^T + Q$, and this is the complete the proof. Now, suppose that the exact solution of equation (9) in section (4.1) is:

$$\hat{X} = (1-\lambda)A\hat{X}A^T + Q$$

And if we take any value X^* that satisfies lemma (4) as:

$$X^* \geq (1-\lambda)AX^*A^T + Q$$

since $\hat{X}$ and X^* are positive semidefinite as mentioned previously, we find that:
$X^* - \hat{X} \geq (1-\lambda)A(X^* - \hat{X})A^T + Q \geq 0$

By putting $X^* = \bar{P}$ (desired error covariance), and $\hat{X} = P_{\lambda}^{\wedge}$, then the matrix $\hat{X}$ is solved by the lyapunov equation:

$$\hat{X} = \bar{\bar{A}}\hat{X}\bar{\bar{A}}^T + Q \text{ where: } \bar{\bar{A}} = \sqrt{(1-\lambda)}A$$

Therefore, $P_{\lambda}^{\wedge} = \bar{\bar{A}}P_{\lambda}^{\wedge}\bar{\bar{A}}^T + Q$. Here, we can say that for $P_{\lambda}^{\wedge}$ to be bounded or stable, that means $P_{\lambda}^{\wedge} \leq \bar{P}$ (desired error covariance), the eigenvalue of the matrix $\bar{\bar{A}}$, which is given by $\sqrt{(1-\lambda)}eig(A)$, should be less than 1, i.e.,

$$\sqrt{(1-\lambda)}eig(A) < 1$$

$$(1-\lambda)(eig(A))^2 < 1$$
$$\therefore 1-\lambda < \frac{1}{[\max eig(A)]^2}$$

Then, $\lambda > 1 - \frac{1}{[\max\ eig(A)]^2}$

We conclude that, the minimum value of the arrival rate λ, such that the estimated error covariance is bounded (stable) or $P_{\lambda}^{\wedge} \leq \bar{P}$ is given by:

$$\lambda = 1 - \frac{1}{[\max\ eig(A)]^2} \quad (23)$$

5.2 Analysis with correlated noises

Problem 2: In this section we are looking for the filtering performance when the system process noise and measurement noise are correlated to each other at one time step apart. This mean that, the process noise at time step k does not contribute to the measurement at time step k, but the process noise at time k-1 contributes to the measurement at time k through the following standard dynamic system equations

$$X_k = AX_{k-1} + w_{k-1}$$

$$y_k = CX_k + v_k$$

Then: $y_k = C(AX_{k-1} + w_{k-1}) + v_k$

In the standard Kalman filter the cross-correlation between the process and measurement noise is given by

$$E[w_k v_j^T] = 0$$

And as mentioned before, in the case of correlated noises we have

$E[w_k v_j^T] = \delta_{kj-1}Z \geq 0$, where $\delta_{kj} = 0$ if $k \neq j$ and $\delta_{kj} = 1$ otherwise.

And the prediction cycle in the case of correlated noises is equivalent to that of the standard Kalman filter, (see equations (18) and (19) in section (4.3). This cycle is only performed when $\lambda_k = 0$. When $\lambda_k = 1$, both the prediction cycle and filtering cycle are performed through equations (20), (21) and (22) in section (4.3). Therefore, the function $F(X)$ in section (4.1) which represents the estimation error covariance in the prediction cycle, has the same form as before, but the function $g(X)$ is given in a new form as

$$g(X) = F(X) - (F(X)C^T + Z)(CF(X)C^T + R + CZ + ZTCT)-1CFX$$

which represents the estimation error covariance in the filtering cycle. Using the same procedure adopted in problem 1, we can claim that, the minimum value of the observation arrival rate λ in which the estimated error covariance is bounded has the same form as that given by equation (23). Since the two noises are correlated in this case, the concept of correlation coefficient ρ is an important performance factor in the filtering process. From the fact that the correlation coefficient $\rho_{x,y}$ between two random variables x and y with zero mean, and standard deviation σ_x and σ_y is defined as

$$\rho_{x,y} = Corr(x,y) = \frac{cov(x,y)}{\sigma_x \sigma_y} = \frac{E[xy]}{\sigma_x \sigma_y}$$

where, E is the expected value operator, (cov means covariance, and corr the correlation coefficient). Then, for the process noise covariance Q and the measurement noise covariance R , the cross-correlation between process and measurement noise covariance Z as a function of correlation coefficient ρ is given by

$$Z = \rho\sqrt{Q}\sqrt{R}$$

6. ILLUSTRATIVE EXAMPLES

In this section, we provide two examples to illustrate the theories and algorithm developed in the previous sections. We start with an example for the packet dropping analysis with uncorrelated noises.

6.1 Problem 1: Packet dropping analysis with uncorrelated noises.

Consider system (1) and (2) with:

$$A = \begin{bmatrix} 1.25 & 1 & 0 \\ 0 & 0.9 & 7 \\ 0 & 0 & 0.6 \end{bmatrix}; C = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}; \; Q = 20 I_{3\times 3}, R = 2.5$$

Fig. 2 plots the estimation error covariance P_k as a function of packet arrival rate λ . The simulation is done using Matlab program. Figure shows that the critical value of the packet arrival rate λ is about 0.36, and this is equivalent to the value obtained by equation (23) in the previous section. When λ the estimation error covariance is below 0.36 becomes unbounded. The value of λ obtained depends on the eigenvalues of the system matrix A.

6.2 Problem 2: Packet dropping analysis with correlated noises

Fig.3 plots the effect of correlation coefficient ρ on the estimated error covariance around the critical value of the observation arrival rate $\lambda = 0.36$. From Fig. 3 when $\rho = 0$, the estimated error covariance is 1.9939. It is obvious that when the correlation coefficient increases in the positive direction the estimated error covariance has better performance characteristics

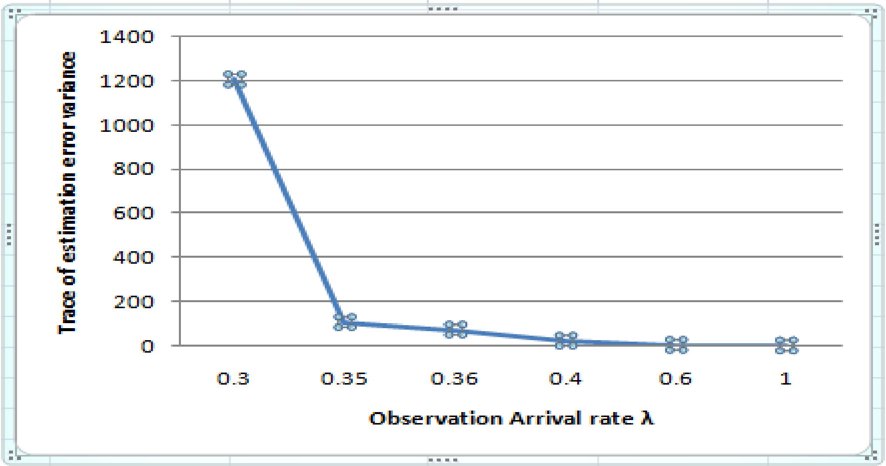

Fig. 2. Estimated error covariance versus λ

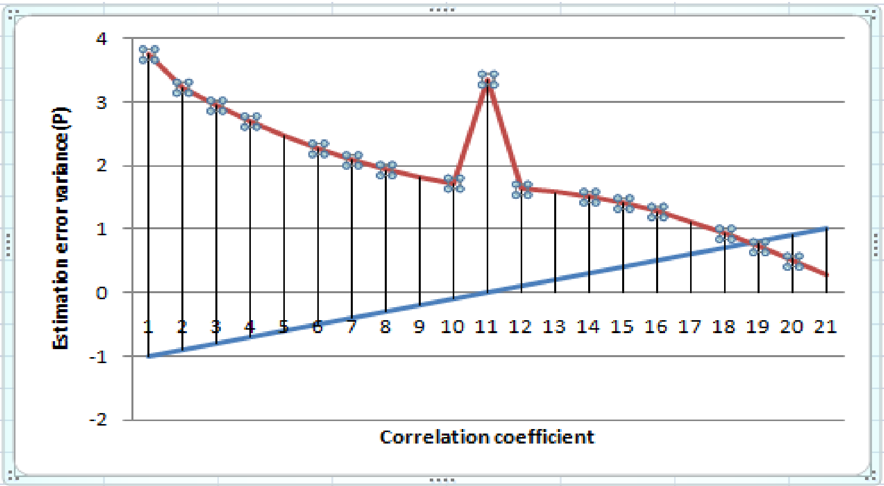

Fig. 3. Correlation coefficient ρ and estimation error covariance P_k

7. CONCLUSIONS

In this paper, we considered discrete-time state estimation over a network. Two scenarios were investigated. In "Packet dropping analysis with uncorrelated noises", we show that if the information packet is lost through the network, or if λ=0 , the Kalman filtering algorithm faces a problem in performing its task. In this situation, only the prediction cycle of the filter is executed. On the other hand, if the packet is totally received by the estimator, the Kalman filter formulation will be same as that of the standard Kalman filter. In this case, both the prediction and filtering cycles are executed. Furthermore, the packet arrival rate of the observation lies between 0 and 1 due to the unreliable communication link. We showed that the minimum packet arrival rate that gave a bounded estimated error covariance depends on the eigenvalues of the system matrix A and the measurement matrix C.

In a packet dropping analysis with correlated noises at one time step apart, we investigate the performance of the estimated error covariance around the critical value of the arrival rate. As a result, with the absolution of the correlation coefficient increasing, the error performance of the standard Kalman filter is worse than that of the correlated noises Kalman filtering.

There are many future directions along the line of this work. The dropping process may be assumed to be a markov chain with two transition states. One state is for packet drop or missing and the other for packet receiving. Also, we can extend these results to the case of nonlinear system analysis using extended Kalman filter as an optimal estimator.

APPENDIX A

Prediction and update cycles of Kalman filter with correlated noises at one time step apart

Suppose the dynamic equation and measurement equation are described by:

$$x_{k+1} = A_{k+1}x_k + w_k \quad (1)$$

$$y_k = C_k x_k + v_k \quad (2)$$

The covariance of the noises is given by:

$$E[w_k w_k^T] = Q_k \delta_{k,j},\ E[v_k v_k^T] = R_k \delta_{k,j}, E[w_k v_j^T] = Z_k \delta_{k,j-1} \quad (3)$$

From the above formulas, we see that the process noise in the system equation at time k-1 is correlated with the measurement noise at time k with covariance Z_k. The state prediction error is

$$\tilde{x}_{k,k-1} \triangleq x_k - \hat{x}_{k,k-1} = A_{k,k-1}\tilde{x}_{k-1} + w_{k-1} \quad (4)$$

The covariance between the state prediction error and the measurement noise is

$$E[\tilde{x}_{k,k-1}v_k^T] = E[(A_{k,k-1}\tilde{x}_{k-1} + w_{k-1})v_k^T] = Z_k \quad (5)$$

The covariance between the state and measurement is calculated by

$$P_{xy} = E[\tilde{x}_{k,k-1}\tilde{y}_{k,k-1}^k | Y^k] = E[\tilde{x}_{k,k-1}(C_k\tilde{x}_{k,k-1} + v_k)^T | Y^k] = P_{k,k-1}C_k^T + Z_k \quad (6)$$

The measurement prediction covariance is given as

$$P_{yy} = S_k = E[\tilde{y}_{k,k-1}\tilde{y}_{k,k-1}^k] = E[(C_k\tilde{x}_{k,k-1} + v_k)(C_k\tilde{x}_{k,k-1} + v_k)^T] = C_k P_{k,k-1}C_k^T + R_k + C_k Z_k + Z_k^T C_k \quad (7)$$

The corresponding filter gain is

$$K_k = (P_{k,k-1}C_k^T + Z_k)(C_k P_{k,k-1}C_k^T + R_k + C_k Z_k + Z_k^T C_k^T)^{-1} \quad (8)$$

According to the minimum mean square error criterion, the state estimation is

$$\hat{x}_k = \hat{x}_{k,k-1} + P_{xy}P_{yy}^{-1}(y_k - C_k\hat{x}_{k,k-1}) = \hat{x}_{k,k-1} + K_k(y_k - C_k\hat{x}_{k,k-1}) \quad (9)$$

And the associated covariance of the state vector is

$$P_k = P_{xx} - P_{xy}P_{yy}^{-1}P_{yx} = P_{k,k-1} - K_k(P_{k,k-1}C_k^T + Z_k)^T \quad (10)$$

So, from the deviation above, we can give the filtering procedure as follows:

Step 1: One time step prediction:

$$\hat{x}_{k,k-1} = A_{k,k-1}\hat{x}_{k-1} \qquad , P_{k,k-1} = A_{k,k-1}P_{k-1}A_{k,k-1}^T + Q_k \quad (11)$$

Step 2: Calculate Kalman gain by (8);

Step 3: Calculate the time update by using (9) and (10).
To be noticed that the above method is a generalization of the standard Kalman filter when the cross correlated covariance matrix $Z_k = 0$.

APPENDIX B

Proof of important lemmas used in section (5) Lemma (1)

For any $0 \leq X \leq Y$, and any $\propto \in [0,1]$, the following are holds:

(i) $F(X) \leq F(Y)$
(ii) $g(X) \leq g(Y)$
(iii) $g(X) \leq F(X)$

Proof: From the definition of F , it is easy to verify that $F(X) \leq F(Y)$.

(i) Since, $\tilde{g}(X) \triangleq X - (XC^T + Z)(CXC^T + R + CZ + ZTCT)-1CX$ represent affine function in X , and since $X \leq Y$, therefore, $\tilde{g}(X) \leq \tilde{g}(Y)$, also since $g(X) = \tilde{g}(F(X)) \leq \tilde{g}(F(Y)) = g(Y)$
Then, $g(X) \leq g(Y)$.

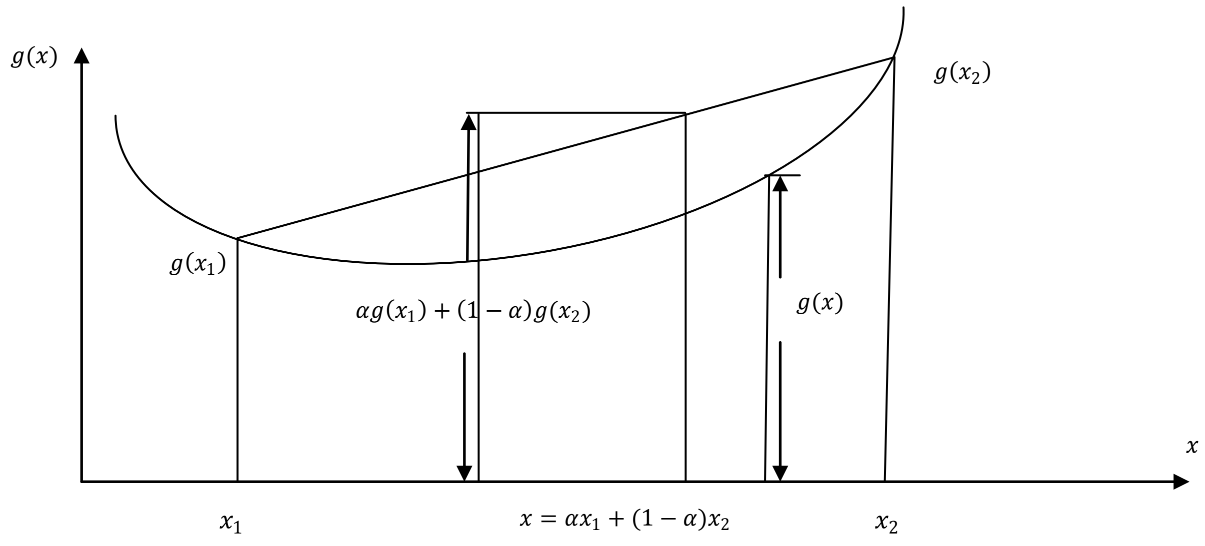

Fig. 4. Convexity function

(ii) $g(X) = \tilde{g}\big(F(X)\big)$
And $\tilde{g}(X) \triangleq X - (XC^T + Z)(CXC^T + R + CZ +$ ZTCT)−1CX,
Therefore, $\tilde{g}\big(F(X)\big)\ F(X)$
(iv) So, $g(X) \leq F(X)$, and this complete the proof.

Lemma (2)
If $\tilde{g}(X)$ is a concave function in X, then for a random variable W_k the following holds:

$$\tilde{g}(E[W_k]) \ \leq E[\tilde{g}(W_k)]$$

Proof: Firstly, for the convexity of $\tilde{g}(X)$, refers to appendix(C) in this paper.
Secondly, by applying Jensen's inequality to the function $\tilde{g}(X)$ and random variable W_k , we can verify the above inequality.

Lemma (3)
For any $\lambda_1 , \lambda_2 \in [0,1]$ and if $\lambda_1 \ \leq \lambda_2$, then $P^{\wedge}_{\lambda_1} \ \geq P^{\wedge}_{\lambda_2}$.
Proof (1):
Note that $(XC^T + Z)(CXC^T + R + CZ + Z^T C^T)^{-1} CX \geq 0$, then from (5.24),

$$P_{\lambda_1}(X) = AXA^T + Q - \lambda_1 A(XC^T + Z)(CXC^T + R + CZ + Z^T C^T)^{-1} CXA^T$$

$\geq AXA^T + Q - \lambda_2 A(XC^T + Z)(CXC^T + R + CZ +$ ZTCT)−1CXAT

$= P_{\lambda_2}(X)$, therefore, $P^{\wedge}_{\lambda_1} \ \geq P^{\wedge}_{\lambda_2}$, and this complete the

Another proof: Since , $P^{\wedge}_{\lambda}$ represent the steady – state error covariance, therefore:

$$P^{\wedge}_{\lambda} = \lim_{k\to\infty} P^{\wedge}_{\lambda,k}$$

With $P^{\wedge}_{\lambda,k+1} = (1 - \lambda) F\big(P^{\wedge}_{\lambda,k}\big) + \lambda g(P^{\wedge}_{\lambda,k})$
Then we can proof by induction:
let $P^{\wedge}_{\lambda_1,0} = P^{\wedge}_{\lambda_2,0}$, then :

$$P^{\wedge}_{\lambda_1,1} = (1 - \lambda_1) F\big(P^{\wedge}_{\lambda_1,0}\big) + \lambda_1 g(P^{\wedge}_{\lambda_1,0})$$
$$= (1 - \lambda_1) F\big(P^{\wedge}_{\lambda_2,0}\big) + \lambda_1 g(P^{\wedge}_{\lambda_2,0})$$
$$\geq (1 - \lambda_2) F\big(P^{\wedge}_{\lambda_2,0}\big) + \lambda_2 g(P^{\wedge}_{\lambda_2,0}) = P^{\wedge}_{\lambda_2,1}$$

Then, $P^{\wedge}_{\lambda_1,1} \geq P^{\wedge}_{\lambda_2,1}$, by lemma(1) (iii).
Now, we can assume that $P^{\wedge}_{\lambda_1,m} \geq P^{\wedge}_{\lambda_2,m}$, for some $m \ \geq 1$, then : $P^{\wedge}_{\lambda_1,m+1} = (1 - \lambda_1) F\big(P^{\wedge}_{\lambda_1,m}\big) + \lambda_1 g(P^{\wedge}_{\lambda_1,m})$

$$\geq (1 - \lambda_1) F\big(P^{\wedge}_{\lambda_2,m}\big) + \lambda_1 g(P^{\wedge}_{\lambda_2,m})$$
$$\geq (1 - \lambda_2) F\big(P^{\wedge}_{\lambda_2,m}\big) + \lambda_2 g(P^{\wedge}_{\lambda_2,m}) \ = P^{\wedge}_{\lambda_2,m+1}$$

Therefore, $P^{\wedge}_{\lambda_1,m+1} \geq P^{\wedge}_{\lambda_1,m+1}$, $iff \lambda_1 \ \leq \lambda_2$
Then, $iff\ \lambda_1 \ \leq \lambda_2$, implies that $P^{\wedge}_{\lambda_1} \ \geq P^{\wedge}_{\lambda_2}$, and this complete the proof.

From the above lemma (3), we can claim that $P^{\wedge}_{\lambda}$ is monotonically decreasing in $\lambda_1 , \lambda_2 \in [0,1]$.

APPNDEX C

Convexity

A set S is convex if $\forall\, x_1, x_2 \ \in S , \alpha \ \in [0,1] \Longrightarrow \alpha x_1 + (1 - \alpha) x_2 \in S$.
A function $g(x)$ is a convex function defined on a convex Set S if:

$$\forall\, x_1, x_2 \ \in S , \alpha \ \in [0,1] \Longrightarrow g(\alpha x_1 + (1 - \alpha) x_2) \leq \alpha g(x_1) + (1 - \alpha) g(x_2)$$

The Fig. 4 illustrates this concept:

A function $g(x)$ is strictly Convex if :
$g(\alpha x_1 + (1 - \alpha) x_2) < \alpha g(x_1) + (1 - \alpha) g(x_2) \quad \forall\, x_1, x_2 \ \in S , \alpha \ \in [0,1]$. And a function g is Concave if : - g is Convex.
Properties:
1- If $g(x^*)$ is a local minimum of a Convex function g on a Convex Set S , then it is also a global minimum.
2- If g is Convex and $g^{/}(x^*) = 0$, then x^* is a global minimum of g.

REFERENCES

[1] " *Special Issue on Networks and Control*," IEEE Control Systems Magazine , Vol.22, Feb.2002.

[2] " *Special Issue on Networks and Control*," IEEE Control Systems Magazine , Vol.22, Feb.2002.

[3] " *Special Issue on Networked Control Systems*," IEEE Transactions on Automatic Control, Vol.49,Issue 9, Sep.2004.

[4] A. Sahai, "Anytime information theory," Ph.D. dissertation, Massachusetts Institute of Technology, 2001.

[5] R. E. Kalman, "*A new approach to linear filtering and prediction problems*," Trans. ASME—J. Basic Eng. Automat. Control, vol. 82, no. D, pp. 35–45, Mar. 1960.

[6] P. S. Maybeck, Stochastic Models, Estimation, and Control, ser. Mathematics in Science and Engineering. New York: Academic, 1979, vol.141.

[7] N. Gordon, D. Salmond, and A. Smith, "*A novel approach to nonlinear nongaussian bayesian state estimation*," in Proc. IEEE Conf. Radar and Signal Processing, vol. 140, 1993, pp. 107–113.

[8] M.Micheli, "Random sampling of a continuous-time stochastic dynamical system: Analysis, state estimation, applications," M.S. thesis, Dept. Elect. Eng., Univ. California, Berkeley, 2001.

[9] M. Athans, R. Ku, and S. B. Gershwin, "The *uncertainty threshold principle, some fundamental limitations of optimal decision making under dynamic uncertainty*," IEEE Trans. Automat. Contr., vol. AC-22, pp.491–495, June 1977.

[10] R. Ku and M. Athans, "*Further results on the uncertainty threshold principle*," IEEE Trans. Automat. Contr., vol. AC-22, pp. 491–495, Oct. 1977.

[11] M. Mariton, Jump Linear Systems in Automatic Control. New York:Marcel Dekker, 1990.

[12] B. Sinopoli, L.Schenato, M. Franceschetti, K.Poolla,M.Jordan, and S.sastry," *Kalman Filtering with Intermittent Observations*," IEEE Trans. on Automatic Control,Vol.49,no.9,pp.1453-1464,September 2004.

[13] M. Moayedi, Y. C. Soh, and Y. K. Foo. "*Optimal Kalman filtering with random sensor delays, packet dropouts and missing measurements*". In Proceedings of 2009 American Control Conference, pages 3405–3410, St. Louis, MO, June 2009.