Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Optimal Estimation for Networked Control Systems with Intermittent Inputs without Acknowledgement ⋆

Hong Lin * **Hongye Su** *† **Zhan Shu** ** **Zheng-Guang Wu** *
Yong Xu *

* *Institute of Cyber-Systems and Control, Zhejiang University, Hangzhou, Zhejiang, 310027, China*
(e-mail: linhongzju@163.com (H. Lin), hysu@iipc.zju.edu.cn (H. Su), nashwzhg@126.com (Z.-G. Wu), xuyong809@163.com (Y. Xu))
** *School of Engineering Sciences, University of Southampton, Southampton SO17 1BJ, UK. (e-mail: hustd8@gmail.com)*

Abstract: This paper investigates the optimal estimation problem for networked control systems, where the control packets are randomly dropped without acknowledgement to the estimator. Most existing results for this setup are concerned with the design of controller, while the optimal estimation and its performance evaluation have not been fully studied. This paper shows, unlike many other cases such as intermittent observations or TCP-like systems, the system state follows a Gaussian mixture distribution with exponentially increasing terms. The optimal estimation is obtained by Gaussian sum filtering, while the computation is time-consuming. By constructing an auxiliary estimator, a fast and stable filtering algorithm is proposed to improve computational efficiency.

Keywords: Networked control systems; Optimal estimation; Intermittent inputs; Packet loss

1. INTRODUCTION

In Networked Control Systems (NCSs), the information among sensors, controllers, and actuators is exchanged via networks. Due to various reasons, e.g., channel congestion, delay, signal degradation, etc., packet loss may occur in the information transmission. Usually, packet loss is modeled as an i.i.d. Bernoulli process or a Markov chain. Observation packets lost in sensor-to-estimator (S/E) channel are called intermittent observations, and packets of control signal dropped in controller-to-actuator (C/A) channel are named intermittent inputs. In addition, when intermittent inputs occur, two different fundamental protocols, that is, TCP-like and UDP-like, are introduced. TCP-like means that there is an acknowledgement (ACK) that informs the estimator whether the actuator has successfully received input packets or not, see Fig. 1, while in UDP-like case no acknowledgement (NACK) of reception is available to estimator. In this paper we focus on the Bernoulli packet loss case. To facilitate the problem formulation, we use "the S/E case" to denote the case where packet loss occurs in the S/E channel, and "the C/A(ACK) case" to denote the case where control packets are dropped with ACK, and "the S/E+C/A(ACK) case" to denote both. Similar notions apply to no acknowledgement (NACK), and related scenarios can be defined in a similar way.

⋆ The work is supported by National Natural Science Foundation of P.R. China (NSFC: 61320106009), (NSFC: 61134007), (NSFC: 61304072), and The National High Technology Research and Development Program of China (863 Program 2012AA040305).
†: Corresponding author

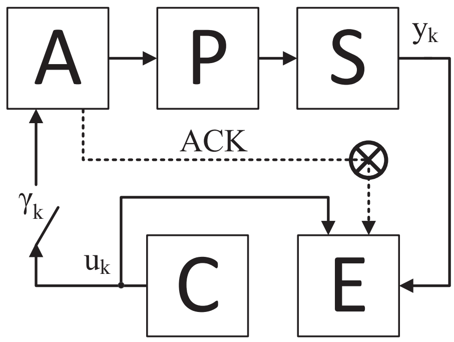

Fig. 1. NCSs with packet loss in C/A channel without acknowledgement. The blocks P, S, E, C and A denote the plant, sensor, estimator, controller and actuator, respectively. The symbol ⊗ indicates that there is no acknowledgement from actuator to estimator.

For the S/E case, in Sinopoli et al. (2004), it was pointed out that for an unstable system there exists a critical data arrival rate which determines the convergence of expectation of error covariance. Thereafter, the case of multiple packet dropouts was investigated in Sun et al. (2008) by an innovation analysis approach. Smith and Seiler (2003) developed a jump linear filter, which is more effective than standard time-varying Kalman filter (TVKF). The work on the Markov packet loss case can be found in Huang and Dey (2007); You et al. (2011) and the references therein. Although the estimation problems have been thoroughly investigated for the S/E case, these results cannot be applied directly to the C/A case.

For the C/A(ACK) case, when the ACK is available to estimator, the TVKF is still in effect for the optimal estimation (Imer et al. (2006)). For the C/A(NACK) case, in Schenato et al. (2007), the optimal control was investigated. It has been showed that without ACK the computation of the optimal control requires a nonlinear optimization, and the separation principle does not hold. While for the estimation issue we concerns, due to no acknowledgment from actuator to the estimator, the system can be viewed as Markov jump system with unknown jump mode. It was suggested in Costa et al. (2006) that the optimal filtering requires solving a bank of Kalman filters and its computation is time-consuming. To address the computational complexity, various suboptimal filtering algorithms and schemes have been proposed. These suboptimal filtering algorithms include the detection estimation (DE) algorithms (Tugnait (1982)), the generalized pseudo Bayes (GPB) algorithm (Jaffer and Gupta (1971)), and the interacting multiple model (IMM) algorithm (Blom and Bar-Shalom (1988)) and so on. As pointed out in Li and Bar-Shalom (1993) their performance and stability are usually uncertain and should be evaluated by Monte Carlo simulation method. In Epstein et al. (2007), the system setup is similar to that used in this paper. Epstein developed an estimation scheme consisting of a state estimator and a mode observer to recover the fate of control packet for the C/A(NACK) case. This method was extended by Blind in Blind and Allgower (2009) to solve the estimation problem for the S/E+C/A(NACK) case. While this scheme requires some additional assumptions for the system and control signals, the optimality of the estimation is not always guaranteed. So far for the C/A(NACK) case, the framework of the optimal filtering has been known, but due to the complex structure of the optimal estimation, most literatures focus on the suboptimal filtering algorithms. Consequently, the analytic characterization of the optimal estimation, to our knowledge, has rarely been presented, and the stability of these suboptimal filtering algorithms is usually uncertain.

In this paper, we aim at studying the optimal estimation in NCSs for the C/A(NACK) case. Firstly we show that without ACK the probability density functions (pdf) of system state are presented as Gaussian mixture. Then the optimal estimator is derived, but the computation is time-consuming. By constructing an auxiliary estimator, we develop a stable and efficient suboptimal filtering algorithm.

The rest of the paper is organized as follows: In Section 2, system and problems are formulated. In Section 3, the pdf and the optimal estimation of system state is derived. Then a fast filtering algorithm is proposed in Section 4. The stability of the proposed filtering algorithm is investigated in Section 5. A numerical example is used to illustrate the theoretical analysis in Section 6. The conclusions are presented in Section 7.

Notations:

- $\mathcal{N}_x(\mu, P)$ denotes the Gaussian pdf of the random variable x with mean μ and covariance P.
- $\mathbb{P}(\cdot)$ denotes probability measure.
- $p(\cdot)$ and $p(\cdot|\cdot)$ denote the pdf and the conditional pdf, respectively.
- $\underset{x}{\mathbb{E}}[\cdot]$, and $\underset{x}{\mathrm{cov}}(\cdot)$ denote probability expectation and covariance with respect to x, respectively.
- $\|\cdot\|$ denotes the 2-norm of a vector.
- $(\cdot)_2$ stands for the binary representation, e.g., $(101)_2 = 5$.

2. SYSTEM SETUP

Consider the system,

$$\begin{aligned} x_{k+1} &= Ax_k + \gamma_k Bu_k + \omega_k \\ y_k &= Cx_k + \upsilon_k \end{aligned} \tag{1}$$

where $x_k \in \mathbb{R}^n$ is the system state, $u_k \in \mathbb{R}^q$ the control input, and $y_k \in \mathbb{R}^p$ the observation. ω_k and υ_k are zero mean Gaussian noise with covariance $Q \geq 0$ and $R > 0$, respectively. γ_k is an i.i.d. Bernoulli random sequence with mean γ, which models the packet loss of the C/A channel. That is, $\gamma_k = 1$ indicates that the control packet u_k has been successfully delivered to actuator, otherwise $\gamma_k = 0$. The initial state x_0 is assumed to be Gaussian with mean $\bar{x}_0$ and covariance P_0.

The system described in (1) without ACK is denoted by $\mathcal{S}_{\mathcal{N}}$, and the one with ACK is denoted by $\mathcal{S}_{\mathcal{A}}$. The problem of estimation for the S/E case has been addressed in Sinopoli et al. (2004). Here we focus on the estimation for the C/A(NACK) case. Thus we assume the S/E channel is free of packet loss. Meanwhile, this paper does not involve the design of control. Thus in Fig. 1 the arrow from E to C, usually existing in the closed-loop system, is not presented here. For the system (1) we make an assumption as follows.
Assumption 1. ($\mathcal{A}1$) The pair (A, Q) is stabilizable, and the pair (A, C) detectable.

Define the Information set as $\mathcal{I}_k \triangleq \{y_k, \cdots, y_1\}$, and $\mathcal{I}_0 \triangleq \phi$ (empty set). In this paper, the *optimality* of the estimation is according to the minimum mean square error (MMSE) criterion. That is, the optimal estimation is the one, denoted by $\hat{x}_{k|k}$, minimizing $\mathbb{E}[\|x_k - \hat{x}_{k|k}\|^2]$. It is well known (Anderson and Moore (1979)) that the desired optimal estimation $\hat{x}_{k|k}$ is given by $\mathbb{E}[x_k|\mathcal{I}_k]$. Then denote $\hat{x}_{k+1|k} \triangleq \mathbb{E}[x_{k+1}|\mathcal{I}_k]$ as the state prediction. Denote $P_{k|k}$ and $P_{k+1|k}$ as the estimation and prediction error covariances, respectively. Let $p(x_k|\mathcal{I}_k)$ and $p(x_{k+1}|\mathcal{I}_k)$ stand for the pdf of x_k and x_{k+1} conditioned on $\mathcal{I}_k$, respectively.

For system $\mathcal{S}_{\mathcal{A}}$, the prediction error covariance and estimation error covariance, denoted by $S_{k+1|k}$ and $S_{k|k}$ respectively, can be calculated by standard Kalman filtering as follows.

$$S_{k+1|k} = AS_{k|k}A^T + Q \tag{2}$$

$$K_{k+1} = S_{k+1|k}C^T \left(CS_{k+1|k}C^T + R\right)^{-1} \tag{3}$$

$$\begin{aligned} S_{k+1|k+1} = {} & (I - K_{k+1}C)S_{k+1|k}(I - K_{k+1}C)^T \\ & + K_{k+1}RK_{k+1}^T \end{aligned} \tag{4}$$

with initial condition $S_{0|0} = P_0$. It is known (Anderson and Moore (1979)) that under the Assumption ($\mathcal{A}1$), $S_{k|k}$ converges.

The aims of this paper are to, for system $\mathcal{S}_{\mathcal{N}}$,

- Find out the pdf of system state and obtain the optimal estimation;
- Develop a stable and efficient filtering algorithm;
- Analyze the impact of packet loss on estimation.

Preliminaries:

Let X and b be Gaussian random variables with pdfs $\mathcal{N}_X(m, P)$ and $\mathcal{N}_b(0, W)$, respectively. Let $Y = CX + b$ where C is a constant matrix. Then

$$p(Y) = \mathcal{N}_Y(Cm, CPC^T + W) \tag{5}$$

$$p(X|Y) = \mathcal{N}_X(m + K(y - Cm), (I - KC)P) \tag{6}$$

where $K = PC^T(CPC^T + W)^{-1}$.

3. OPTIMAL ESTIMATION OF SYSTEM STATE

In this section, we derive the pdf of x_k, that is, $p(x_k|\mathcal{I}_{k-1})$ and $p(x_k|\mathcal{I}_k)$, then compute the optimal estimation.

We firstly introduce the presentation of the random events and its properties. For the packet loss random variable sequence $\{\gamma_k, \ldots, \gamma_0\}$, an event takes the following form $\{\gamma_k = \theta_k, \ldots, \gamma_0 = \theta_0\}$ where $\theta_j \in \{0,1\}$ for $0 \leq j \leq k$. The probability space denoted by Γ_k contains 2^{k+1} possible such events. For each binary-valued sequence $\boldsymbol{\theta}_k \triangleq (\theta_k \cdots \theta_0)$ there is associated with an unique integer i determined by $i = \rho(\boldsymbol{\theta}_k) \triangleq (\theta_k \cdots \theta_0)_2 + 1$. It is easy to check that the mapping ρ is a bijection. Hence the event can also be denoted by

$$\Theta_k^i \triangleq \{\gamma_k = \theta_k, \ldots, \gamma_0 = \theta_0 | i = \rho(\boldsymbol{\theta}_k)\}, 1 \leq i \leq 2^{k+1}.$$

An useful property of Θ_{k+1}^i is that for $1 \leq i \leq 2^{k+1}$,

$$\Theta_{k+1}^i = \{\gamma_{k+1} = 0, \Theta_k^i\} \text{ and } \Theta_{k+1}^{i+2^{k+1}} = \{\gamma_{k+1} = 1, \Theta_k^i\} \tag{7}$$

The equation (7) can be checked by using the knowledge of binary representation.

3.1 Probability density function of x_k

By using total probability law, we have

$$p(x_k|\mathcal{I}_{k-1}) = \sum_{i=1}^{2^k} p(x_k|\Theta_{k-1}^i, \mathcal{I}_{k-1}) p(\Theta_{k-1}^i|\mathcal{I}_{k-1}) \tag{8a}$$

$$p(x_k|\mathcal{I}_k) = \sum_{i=1}^{2^k} p(x_k|\Theta_{k-1}^i, \mathcal{I}_k) p(\Theta_{k-1}^i|\mathcal{I}_k). \tag{8b}$$

For each possible event Θ_k^i, all the four conditional pdfs in (8) are computed in the following two lemmas.

Lemma 1. The conditional pdfs of x_k in (8) are computed as follows: For $1 \leq i \leq 2^k$,

$$p(x_k|\Theta_{k-1}^i, \mathcal{I}_{k-1}) = \mathcal{N}_{x_k}(m_{k|k-1}^i, S_{k|k-1}) \tag{9a}$$

$$p(x_k|\Theta_{k-1}^i, \mathcal{I}_k) = \mathcal{N}_{x_k}(m_{k|k}^i, S_{k|k}) \tag{9b}$$

with initial condition $m_{0|0}^1 = \bar{x}_0$, where $m_{k|k-1}^i$ and $m_{k|k}^i$ evolve as (12). $S_{k|k-1}$ and $S_{k|k}$ are calculated by (2)-(4).

Proof. The proof is to be completed by mathematical induction. For $k=1$, it is easy to check that (9) and (12) hold. Then we suppose that (9) and (12) hold for $1, \ldots, k$, and check the case for $k+1$ with $1 \leq i \leq 2^{k+1}$. By this assumption we have $p(x_k|\Theta_{k-1}^i, \mathcal{I}_k) = \mathcal{N}_{x_k}(m_{k|k}^i, S_{k|k})$, First consider the case $1 \leq i \leq 2^k$, by (7), we have $\Theta_k^i = \{\gamma_k = 0, \Theta_{k-1}^i\}$. By using (5),

$$\begin{aligned} p(x_{k+1}|\Theta_k^i, \mathcal{I}_k) &= p(Ax_k + \gamma_k u_k + \omega_k | \gamma_k = 0, \Theta_{k-1}^i, \mathcal{I}_k) \\ &= \mathcal{N}_{x_{k+1}}(Am_{k|k}^i, AS_{k|k}A^T + Q) \\ &= \mathcal{N}_{x_{k+1}}(m_{k+1|k}^i, S_{k+1|k}). \end{aligned}$$

When y_{k+1} is observed, by using (6), we have

$$\begin{aligned} &p(x_{k+1}|\Theta_k^i, \mathcal{I}_{k+1}) \\ &= \mathcal{N}_{x_{k+1}}(m_{k+1|k}^i + K_{k+1}(y_{k+1} - Cm_{k+1|k}^i), S_{k+1|k+1}) \\ &= \mathcal{N}_{x_{k+1}}(m_{k+1|k+1}^i, S_{k+1|k+1}). \end{aligned}$$

Therefore (9) and (12) hold for $k+1$ when $1 \leq i \leq 2^k$. For $2^k + 1 \leq i \leq 2^{k+1}$, corresponding to the case without packet loss, by following a similar line, (9) and (12) can be proved to be true as well, and thus the derivation is omitted here for brevity. This completes the proof.

Lemma 2. Let $\alpha_{k|k-1}^i \triangleq p(\Theta_{k-1}^i|\mathcal{I}_{k-1})$ and $\alpha_{k|k}^i \triangleq p(\Theta_{k-1}^i|\mathcal{I}_k)$, for $1 \leq i \leq 2^k$. Then $\alpha_{k|k-1}^i$ and $\alpha_{k|k}^i$ evolve as (11) with initial condition $\alpha_{0|0}^1 = 1$.

Proof. We prove this lemma by mathematical induction. It is easy to check that (11) holds at $k = 1$. Suppose that (11) holds for $1, \ldots, k$. Then we check the condition for $k+1$ with $1 \leq i \leq 2^{k+1}$.

First, let $1 \leq i \leq 2^k$, by (7), $\alpha_{k+1|k}^i = p(\Theta_k^i|\mathcal{I}_k) = p(\gamma_k = 0, \Theta_{k-1}^i|\mathcal{I}_k) = \bar{\gamma}\alpha_{k|k}^i$. When $2^k + 1 \leq i \leq 2^{k+1}$, by the same derivation, $\alpha_{k+1|k}^i = \gamma\alpha_{k|k}^{i-2^k}$. Hence (11a) holds for $k+1$.

We check (11b) for $k+1$. By Lemma 1, $p(x_{k+1}|\Theta_k^i, \mathcal{I}_k) = \mathcal{N}_{x_{k+1}}(m_{k+1|k}^i, S_{k+1|k})$. By using (6), $p(y_{k+1}|\Theta_k^i, \mathcal{I}_k) = \mathcal{N}_{y_{k+1}}(Cm_{k+1|k}^i, S_{k+1}^Y) = \phi_k^i(y_{k+1})$. Then by using Bayesian formula again,

$$\begin{aligned} \alpha_{k+1|k+1}^i &\triangleq p(\Theta_k^i|\mathcal{I}_{k+1}) = p(\Theta_k^i|y_{k+1}, \mathcal{I}_k) \\ &= \frac{p(y_{k+1}|\Theta_k^i, \mathcal{I}_k) p(\Theta_k^i|\mathcal{I}_k)}{\sum_{j=1}^{2^{k+1}} p(y_{k+1}|\Theta_k^j, \mathcal{I}_k) p(\Theta_k^j|\mathcal{I}_k)} \\ &= \frac{\phi_k^i(y_{k+1})\alpha_{k+1|k}^i}{\sum_{j=1}^{2^{k+1}} \phi_k^j(y_{k+1})\alpha_{k+1|k}^j}. \end{aligned}$$

Therefore (11) holds for $k+1$. This completes the proof.

Theorem 3. For the system $\mathcal{S}_{\mathcal{N}}$ with i.i.d Bernoulli packet loss in the C/A channel, $p(x_k|\mathcal{I}_{k-1})$ and $p(x_k|\mathcal{I}_k)$ are Gaussian mixture, that is, for $k \geq 1$

$$p(x_k|\mathcal{I}_{k-1}) = \sum_{i=1}^{2^k} \alpha_{k|k-1}^i \mathcal{N}_{x_k}(m_{k|k-1}^i, S_{k|k-1}) \tag{10a}$$

$$p(x_k|\mathcal{I}_k) = \sum_{i=1}^{2^k} \alpha_{k|k}^i \mathcal{N}_{x_k}(m_{k|k}^i, S_{k|k}) \tag{10b}$$

where

$$\alpha_{k|k-1}^i = \begin{cases} \bar{\gamma}\alpha_{k-1|k-1}^i, & \text{for } 1 \leq i \leq 2^{k-1} \\ \gamma\alpha_{k-1|k-1}^{i-2^{k-1}}, & \text{for } 2^{k-1}+1 \leq i \leq 2^k \end{cases} \tag{11a}$$

$$\alpha_{k|k}^i = \frac{\phi_{k-1}^i(y_k)\alpha_{k|k-1}^i}{\sum_{j=1}^{2^k} \phi_{k-1}^j(y_k)\alpha_{k|k-1}^j} \tag{11b}$$

and

$$m_{k|k-1}^i = \begin{cases} Am_{k-1|k-1}^i, & \text{for } 1 \leq i \leq 2^{k-1} \\ Am_{k-1|k-1}^{i-2^{k-1}} + Bu_{k-1}, & \text{for } 2^{k-1}+1 \leq i \leq 2^k \end{cases} \tag{12a}$$

$$m_{k|k}^i = m_{k|k-1}^i + K_k(y_k - Cm_{k|k-1}^i) \tag{12b}$$

with initial condition, $\alpha_{0|0}^1 = 1$ and $m_{0|0}^1 = \bar{x}_0$, where $\phi_{k-1}^i(y_k) \triangleq p(y_k|\Theta_{k-1}^i, \mathcal{I}_{k-1}) = \mathcal{N}_{y_k}(Cm_{k|k-1}^i, S_k^Y)$, $S_k^Y \triangleq CS_{k|k-1}C^T + R$, $\bar{\gamma} \triangleq 1 - \gamma$. $S_{k|k-1}$, K_k and $S_{k|k}$ are calculated by (2)-(4).

Proof. Based on (8) and the conditional pdfs of x_k derived in Lemma 1 and 2, the proof is straightforward.

Remark 1. Unlike the cases of S/E or C/A(ACK) in Schenato et al. (2007) where the pdf of x_k is Gaussian, the pdf of x_k is Gaussian mixture with exponentially increasing terms for the C/A(NACK) case. Moveover for each term in the Gaussian mixture the covariance is the same and is equal to $S_{k|k}$, which is the estimation covariance of the system $\mathcal{S}_\mathcal{A}$ and can be determined off-line.

3.2 Optimal estimation

Since $p(x_k|\mathcal{I}_{k-1})$ and $p(x_k|\mathcal{I}_k)$ are Gaussian mixture pdfs, the optimal estimation of x_k can be directly calculated by Gaussian sum filter (Anderson and Moore (1979)), and is formulated as follows:

$$\hat{x}_{k|k} = \sum_{i=1}^{2^k} \alpha^i_{k|k} m^i_{k|k} \tag{13a}$$

$$P_{k|k} = S_{k|k} + \sum_{i=1}^{2^k} \alpha^i_{k|k} (m^i_{k|k} - \hat{x}_{k|k})^2_I \tag{13b}$$

$$\hat{x}_{k+1|k} = \sum_{i=1}^{2^{(k+1)}} \alpha^i_{k+1|k} m^i_{k+1|k} \tag{13c}$$

$$P_{k+1|k} = S_{k+1|k} + \sum_{i=1}^{2^{(k+1)}} \alpha^i_{k+1|k} (m^i_{k+1|k} - \hat{x}_{k+1|k})^2_I. \tag{13d}$$

By some algebraic computations, time update formulas of the estimator take the form as follow:

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k} + \gamma u_k \tag{14}$$

$$P_{k+1|k} = AP_{k|k}A^T + Q + \bar{\gamma}\gamma B u_k u_k^T B^T. \tag{15}$$

Remark 2. Due to the complexity of coupled coefficients in (11b), $\hat{x}_{k|k}$ and $P_{k|k}$ cannot be derived from $\hat{x}_{k|k-1}$ and $P_{k|k-1}$ recursively, and must be computed by (13a) and (13b), which results in a requirement of exponentially increasing memory and time. To address this problem, a fast filtering is proposed in the next section.

4. FAST FILTERING ALGORITHM

In this section, we develop a fast filtering algorithm by constructing an auxiliary estimator.

4.1 Construction of auxiliary estimator

Based on the (8), an auxiliary system state, denoted by $\tilde{x}_k$, is defined by assuming that it has the following pdfs,

$$p(\tilde{x}_k|\mathcal{I}_{k-1}) \triangleq \sum_{i=1}^{2^k} p(\tilde{x}_k|\Theta^i_{k-1}, \mathcal{I}_{k-1}) p(\Theta^i_{k-1}) \tag{16a}$$

$$p(\tilde{x}_k|\mathcal{I}_k) \triangleq \sum_{i=1}^{2^k} p(\tilde{x}_k|\Theta^i_{k-1}, \mathcal{I}_k) p(\Theta^i_{k-1}). \tag{16b}$$

where $p(\tilde{x}_k|\Theta^i_{k-1}, \mathcal{I}_{k-1})$ and $p(\tilde{x}_k|\Theta^i_{k-1}, \mathcal{I}_k)$ are the same functions as $p(x_k|\Theta^i_{k-1}, \mathcal{I}_{k-1})$ and $p(x_k|\Theta^i_{k-1}, \mathcal{I}_k)$ in (8) respectively, just by replacing the symbol x_k with $\tilde{x}_k$.

It is necessary to point out that $p(\tilde{x}_k|\mathcal{I}_k)$ and $p(\tilde{x}_{k+1}|\mathcal{I}_k)$ defined above are indeed pdfs. They satisfy two conditions: $p(x) \geq 0$ and $\int_{-\infty}^{\infty} p(x)\mathrm{d}x = 1$, which are easy to verify and are not presented here.

To compute the estimation of $\tilde{x}_k$, its pdfs are required and are formulated in following lemma.

Lemma 4. $p(\tilde{x}_k|\mathcal{I}_{k-1})$ and $p(\tilde{x}_k|\mathcal{I}_k)$ defined in (16) can be presented as follows:

$$p(\tilde{x}_k|\mathcal{I}_{k-1}) = \sum_{i=1}^{2^k} \tilde{\alpha}^i_{k|k-1} \mathcal{N}_{\tilde{x}_k}(m^i_{k|k-1}, S_{k|k-1}) \tag{17a}$$

$$p(\tilde{x}_k|\mathcal{I}_k) = \sum_{i=1}^{2^k} \tilde{\alpha}^i_{k|k} \mathcal{N}_{\tilde{x}_k}(m^i_{k|k}, S_{k|k}) \tag{17b}$$

where

$$\tilde{\alpha}^i_{k|k-1} = \begin{cases} \bar{\gamma}\tilde{\alpha}^i_{k-1|k-1}, & \text{for } 1 \leq i \leq 2^{k-1} \\ \gamma\tilde{\alpha}^{i-2^{k-1}}_{k-1|k-1}, & \text{for } 2^{k-1}+1 \leq i \leq 2^k \end{cases} \tag{18a}$$

$$\tilde{\alpha}^i_{k|k} = \tilde{\alpha}^i_{k|k-1}, \text{ for } 1 \leq i \leq 2^k \tag{18b}$$

with initial condition $\tilde{\alpha}^1_{0|0} = 1$. $m^i_{k|k-1}$ and $m^i_{k|k}$ evolve in the same way as (12).

Proof. Since $p(\tilde{x}_k|\Theta^i_{k-1}, \mathcal{I}_{k-1})$ and $p(\tilde{x}_k|\Theta^i_{k-1}, \mathcal{I}_k)$ are the same functions as $p(x_k|\Theta^i_{k-1}, \mathcal{I}_{k-1})$ and $p(x_k|\Theta^i_{k-1}, \mathcal{I}_k)$, respectively. By Lemma 1, we readily have

$$p(\tilde{x}_k|\Theta^i_{k-1}, \mathcal{I}_{k-1}) = \mathcal{N}_{\tilde{x}_k}(m^i_{k|k-1}, S_{k|k-1})$$
$$p(\tilde{x}_k|\Theta^i_{k-1}, \mathcal{I}_k) = \mathcal{N}_{\tilde{x}_k}(m^i_{k|k}, S_{k|k})$$

where $m^i_{k|k-1}$ and $m^i_{k|k}$ evolve in the same way as (12).

Next define $p(\Theta^i_{k-1})$ in (16a) as $\tilde{\alpha}^i_{k|k-1}$, and $p(\Theta^i_{k-1})$ in (16b) as $\tilde{\alpha}^i_{k|k}$. Since they are equal, $\tilde{\alpha}^i_{k|k} = \tilde{\alpha}^i_{k|k-1}$ is readily obtained. Thus (18b) holds.

Then we check (18a). Let $1 \leq i \leq 2^{k-1}$, by (7), $\Theta^i_{k-1} = \{\gamma_{k-1} = 0, \Theta^i_{k-2}\}$. $\tilde{\alpha}^i_{k|k-1} \triangleq p(\Theta^i_{k-1}) = p(\gamma_{k-1} = 0, \Theta^i_{k-2}) = p(\gamma_{k-1} = 0)p(\Theta^i_{k-2}) = \bar{\gamma}\tilde{\alpha}^i_{k-1|k-1}$. By the similar derivation procedures, we get $\tilde{\alpha}^i_{k|k-1} = \gamma\tilde{\alpha}^{i-2^{k-1}}_{k-1|k-1}$, for $2^{k-1}+1 \leq i \leq 2^k$. It shows that (18a) holds.

The proof is completed.

The pdfs of $\tilde{x}_k$ remain to be Gaussian mixture, then the estimator of $\tilde{x}_k$ can be computed by Gaussian sum filtering as well.

$$\tilde{x}_{k|k} = \sum_{i=1}^{2^k} \tilde{\alpha}^i_{k|k} m^i_{k|k} \tag{19a}$$

$$\tilde{P}_{k|k} = S_{k|k} + \tilde{\Psi}_{k|k} \tag{19b}$$

$$\tilde{x}_{k+1|k} = \sum_{i=1}^{2^{k+1}} \tilde{\alpha}^i_{k+1|k} m^i_{k+1|k} \tag{19c}$$

$$\tilde{P}_{k+1|k} = S_{k+1|k} + \tilde{\Psi}_{k+1|k}, \tag{19d}$$

where

$$\tilde{\Psi}_{k|k} \triangleq \sum_{i=1}^{2^k} \tilde{\alpha}^i_{k|k} (m^i_{k|k} - \tilde{x}_{k|k})^2_I \tag{20}$$

$$\tilde{\Psi}_{k+1|k} \triangleq \sum_{i=1}^{2^{k+1}} \tilde{\alpha}^i_{k+1|k} (m^i_{k+1|k} - \tilde{x}_{k+1|k})^2_I. \tag{21}$$

By some algebraic calculation it is easy to obtain the following equations from (20) and (21),

$$\tilde{\Psi}_{k+1|k} = A\tilde{\Psi}_{k|k}A^T + U_k \tag{22}$$

$$\tilde{\Psi}_{k+1|k+1} = (I - K_{k+1}C)\,\tilde{\Psi}_{k+1|k}\,(I - K_{k+1}C)^T, \tag{23}$$

where $U_k \triangleq \bar{\gamma}\gamma B u_k u_k^T B^T$ and $\tilde{\Psi}_{0|0} = 0$.

4.2 Fast filtering Algorithm

Compared to x_k, the estimator of $\tilde{x}_k$ still contains exponentially increasing terms, but there are recursive formulas for them as follows, which enables the estimator to be computed recursively and avoid the exponentially increasing computation.

For the system $\mathcal{S}_\mathcal{N}$, a *fast filtering algorithm* (FF) is formulated as a set of Kalman-filtering-like equations as follows:

$$\tilde{x}_{k+1|k} = A\tilde{x}_{k|k} + \gamma B u_k \tag{24a}$$

$$\tilde{P}_{k+1|k} = A\tilde{P}_{k|k}A^T + U_k + Q \tag{24b}$$

$$\tilde{x}_{k+1|k+1} = \tilde{x}_{k+1|k} + K_{k+1}(y_{k+1} - C\tilde{x}_{k+1|k}) \tag{24c}$$

$$\tilde{P}_{k+1|k+1} = (I-K_{k+1}C)\tilde{P}_{k+1|k}(I-K_{k+1}C)^T + K_{k+1}RK_{k+1}^T. \tag{24d}$$

The above equations can be obtained from (19) by some algebraic computations, and thus the detailed derivation is not included here due to page length consideration.

Remark 3. The pdfs of $\tilde{x}_k$ differs from x_k only in (18b). It is such a minor difference that brings the recursive form (24c) and (24d) for $\tilde{x}_k$ and $\tilde{P}_k$.

Remark 4. The equations in (24) appear to be the same as the estimator for the UDP-like case in Schenato et al. (2007). In fact they are quite different. In Schenato et al. (2007) it might assume the pdf of $p(x_{k+1}|\mathcal{I}_k)$ is Gaussian, then the estimator is obtained by Kalman filter. K_{k+1} is calculated by $P_{k+1|k}$, and $P_{k+1|k}$ evolves in Riccati equation. In this paper, (24) is derived from the auxiliary system state $\tilde{x}_k$ by Gaussian sum filter, and the pdfs of $p(\tilde{x}_k|\mathcal{I}_k)$ and $p(\tilde{x}_{k+1}|\mathcal{I}_k)$ are Gaussian mixture. Moreover, in (24), K_{k+1} is computed via $S_{k+1|k}$, not by $K_{k+1} = \tilde{P}_{k+1|k}C^T(C\tilde{P}_{k+1|k}C^T + R)^{-1}$. Thus by substituting (24d) into (24b), the obtained formula is not standard Riccati equation.

5. STABILITY OF THE FAST FILTERING ALGORITHM

As previously shown, the computation of optimal estimation is time-consuming. In general, the stability of those suboptimal filtering algorithms mentioned above is uncertain. Hence in this section, we show that under some condition the proposed FF algorithm is stable. Before presenting this result, we introduce three Lemmas.

Lemma 5. Let $L = CA$, and define Φ_i^k, for $0 \le i \le k$

$$\begin{aligned}\Phi_i^k &\triangleq \big(\textstyle\prod_{j=i+1}^{k}(A - K_{j+1}L)\big)(I - K_{i+1}C), \text{ for } i < k\\ \Phi_k^k &\triangleq (I - K_{k+1}C), \text{ for } i = k,\end{aligned} \tag{25}$$

where K_i is computed via $S_{i|i-1}$ in (3). Then $\tilde{\Psi}_{k|k}$ in (23) can be written as

$$\tilde{\Psi}_{k|k} = \textstyle\sum_{i=0}^{k-1} \Phi_i^{k-1}U_i(\Phi_i^{k-1})^T, \tag{26}$$

with $\tilde{\Psi}_{1|1} = \Phi_0^0 U_0(\Phi_0^0)^T$ as initial values.

Proof. This lemma can be readily proved by mathematical induction, and the detailed process is not presented here.

Lemma 6. Suppose that $\{u_k\} \in \mathbb{R}^q$ are bounded, that is, $\|u_k\| \le u$ for all k. Then there exits a positive semidefinite matrix, denoted by U such that $u_k u_k^T \le U$ for all k.

Proof. The proof of this lemma is straightforward, then is omitted.

Lemma 7. Let U be a positive semidefinite matrix. Then there exits a positive semidefinite matrix, denoted by S_τ such that $\sum_{i=0}^{k} \Phi_i^k U(\Phi_i^k)^T \le S_\tau$ for all k where Φ_i^k is defined in Lemma 5.

Proof. By substituting (2) into (4),

$$S_{k|k} = (I-K_kC)(AS_{k-1|k-1}A^T+Q)(I-K_kC)^T+K_kRK_k^T.$$

By the same algebraic computation, $S_{k|k}$ can formulated in a closed expression as follows.

$$\begin{aligned}S_{k|k} =&\Phi_0^{k-1}(AS_0A^T)(\Phi_0^{k-1})^T + \textstyle\sum_{i=0}^{k-1} \Phi_i^{k-1}Q(\Phi_i^{k-1})^T\\ &+ \textstyle\sum_{i=1}^{k-1} \Omega_i^k K_i R K_i^T(\Omega_i^k)^T + K_kRK_k^T, \text{for } k \ge 2\end{aligned}$$

where $\Omega_i^k \triangleq \prod_{j=i+1}^{k}(A - K_jC)$. Note that all the four terms in the preceding equation are positive semidefinite matrices. Since $S_{k|k}$ is convergent, then it is bounded, i.e., $S_{k|k} \le S$. Hence the second term $\sum_{i=0}^{k-1} \Phi_i^{k-1}Q(\Phi_i^{k-1})^T \le S$ for all k. For a positive semidefinite matrix U, there always exists a real number $\tau > 0$ such that $U \le \tau Q$. Then $\sum_{i=0}^{k-1} \Phi_i^{k-1}U(\Phi_i^{k-1})^T \le \sum_{i=0}^{k-1} \Phi_i^{k-1}\tau Q(\Phi_i^{k-1})^T \le S_\tau$. where $S_\tau \triangleq \tau S$. The proof is completed.

Theorem 8. Consider the system $\mathcal{S}_\mathcal{N}$ with i.i.d Bernoulli packet loss in C/A channel without acknowledgement and assume that the control inputs are bounded. Then $\tilde{P}_{k|k}$ is bounded if and only if $S_{k|k}$ converges.

Proof. If $S_{k|k}$ diverges, from (19b) $P_{k|k}$ diverge as well. So the necessary condition is obvious and we check for the sufficient case. Firstly, consider the term U_k in $\tilde{\Psi}_{k|k}$. As assumed in Theorem 8 that u_k is bounded, then by Lemma 6, there exits a positive defined matrix U, such that $u_k u_k^T \le U$. Since $\gamma + \bar{\gamma} = 1$, then $\gamma\bar{\gamma} \le 1/4$. So $U_k \le \frac{BUB^T}{4}$. As shown in (19d), $\tilde{P}_{k|k}$ is consist of $S_{k|k}$ and $\tilde{\Psi}_{k|k}$.

$$\begin{aligned}\tilde{P}_{k|k} &= S_{k|k} + \textstyle\sum_{i=1}^{k} \Phi_i^k U_i(\Phi_i^k)^T\\ &\le S_{k|k} + \textstyle\sum_{i=1}^{k} \Phi_i^k \frac{BUB^T}{4}(\Phi_i^k)^T\\ &\le S_{k|k} + S_\tau\end{aligned} \tag{27}$$

The equation (27) is obtained due to the Lemma 7. $S_{k|k}$ is convergent and thus is bounded. Therefore $\tilde{P}_{k|k}$ is bounded.

6. NUMERICAL EXAMPLE

In this section by a example, we compare the fast filtering algorithm with the optimal one, verify the boundedness of error covariance, and analyze the impact of packet loss. Consider the system in (1) with following parameters,

$$A = \begin{bmatrix} 0.9 & 0.1 \\ -0.1 & 0.8 \end{bmatrix}, B = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, C = [0\ 1], Q = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, R = 1.$$

Firstly, the performance of the fast filtering and optimal one are compared. We choose a bounded control inputs with packet loss rate 0.2 in the C/A channel, and use the trace of covariance to evaluate the performance. Subject to exponentially increasing computation of the optimal filtering, we run the simulation for 40 steps in a common desktop computer with CPU frequency 2.3 GHz. A comparison of running efficiency and performance between the optimal estimation and FF algorithm is listed in Table 1.

Table 1. Efficiency and performance of the optimal estimation and the fast filtering

Filtering	CPU time	Average of trace of covariance
Optimal estimation	104 minutes	6.26
Fast filtering	2.54 seconds	6.87

The simulation results are illustrated in Fig. 2 and 3. It is shown that $S_{k|k} \leq P_{k|k} \leq \tilde{P}_{k|k}$. The performance of fast algorithm is a little bit inferior to the optimal one, but is acceptable.

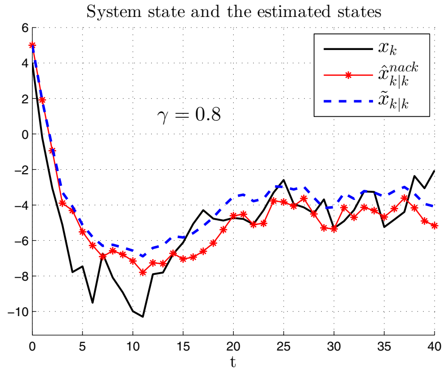

Fig. 2. System state and the estimated states.

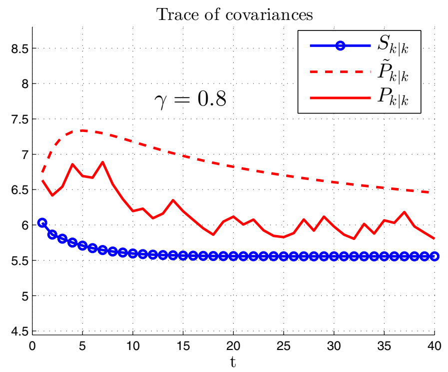

Fig. 3. Trace of covariances: $S_{k|k}$, $P_{k|k}$, and $\tilde{P}_{k|k}$

By fixing the input sequence and changing the packet loss rate from 0 to 1, the performance under different packet loss rate is illustrated in Fig. 4. The trace of covariance increases along with packet loss rate, attains a maximum when $\gamma = 0.5$, then descends when packet loss rate increases from 0.5 to 1.

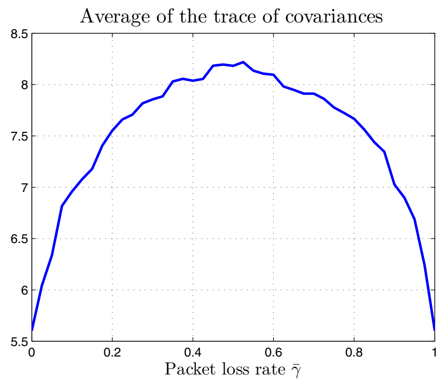

Fig. 4. Relationship between trace of covariance and packet loss rate

7. CONCLUSIONS

In this paper, we have studied the optimal estimation problem for the C/A(NACK) case and developed a fast filtering algorithm to improve computational efficiency. Hence for the case where minor degradation of estimation performance can be tolerated, the NACK setting together with FF algorithm is a fairly good alternate. Not only because it provides a stable and efficient filtering algorithm, but it brings some unique advantages such as simple implementation, less energy consumption, less restrictions on hardware, etc..

REFERENCES

Anderson, B. and Moore, J.B. (1979). *Optimal Filtering*. Englewood Cliffs: Prentice-Hall.

Blind, R. and Allgower, F. (2009). Estimating the fates of the control packets for networked control systems with loss of control and measurement packets. In *Proc. the 48th IEEE Conf.CDC/CCC 2009*, 2687–2692. IEEE.

Blom, H.A. and Bar-Shalom, Y. (1988). The interacting multiple model algorithm for systems with Markovian switching coefficients. *IEEE Trans. Automat. Control*, 33(8), 780–783.

Costa, O.L.V., Fragoso, M.D., and Marques, R.P. (2006). *Discrete-Time Markov Jump Linear Systems*. Springer.

Epstein, M., Shi, L., and Murray, R.M. (2007). Estimation schemes for networked control systems using UDP-like communication. In *Conf. Decision and Control, 2007, 46th IEEE*, 3945–3951. IEEE.

Huang, M. and Dey, S. (2007). Stability of Kalman filtering with Markovian packet losses. *Automatica*, 43(4), 598–607.

Imer, O.C., Yüksel, S., and Başar, T. (2006). Optimal control of LTI systems over unreliable communication links. *Automatica*, 42(9), 1429–1439.

Jaffer, A. and Gupta, S. (1971). On estimation of discrete processes under multiplicative and additive noise conditions. *Information Sciences*, 3(3), 267–276.

Li, X.R. and Bar-Shalom, Y. (1993). Performance prediction of the interacting multiple model algorithm. *IEEE Trans. Aerosp. Electron. Syst.*, 29(3), 755–771.

Schenato, L., Sinopoli, B., Franceschetti, M., Poolla, K., and Sastry, S.S. (2007). Foundations of control and estimation over lossy networks. *Proc. the IEEE*, 95(1), 163–187.

Sinopoli, B., Schenato, L., Franceschetti, M., Poolla, K., Jordan, M.I., and Sastry, S.S. (2004). Kalman filtering with intermittent observations. *IEEE Trans. Automat. Control*, 49(9), 1453–1464.

Smith, S.C. and Seiler, P. (2003). Estimation with lossy measurements: jump estimators for jump systems. *IEEE Trans. Automat. Control*, 48(12), 2163–2171.

Sun, S., Xie, L., Xiao, W., and Soh, Y.C. (2008). Optimal linear estimation for systems with multiple packet dropouts. *Automatica*, 44(5), 1333–1342.

Tugnait, J.K. (1982). Detection and estimation for abruptly changing systems. *Automatica*, 18(5), 607–615.

You, K., Fu, M., and Xie, L. (2011). Mean square stability for Kalman filtering with Markovian packet losses. *Automatica*, 47(12), 2647–2657.