2011 American Control Conference
on O'Farrell Street, San Francisco, CA, USA
June 29 - July 01, 2011

An Innovative Packet-Splitting Approach for Kalman Filtering over Lossy Networks

Junfeng Wu*, Ling Shi*, Lihua Xie†

***Abstract*— We consider the problem of state estimation over lossy networks. Although a large number of approaches have been proposed to improve the estimator's performance, most of them demand either extra channel bandwidth or sensor energy budget. In this paper, we propose an innovative packet-splitting transmission approach and derive a corresponding packet-splitting Kalman Filter (PSKF). In this scheme, one bit of each packet is diverted from quantizing the current innovation to indicate the sign of the previous innovation. We show that if converges, the expected value of the *a posteriori* estimate error covariance ($\mathbb{E}[P_k]$) of the PSKF converges to a smaller value compared with that of modified Kalman filter in literature. Hence the proposed PSKF is able to tolerate a higher or at least equal data loss rate than the MKF. Examples are provided to illustrate the main ideas.**

Keywords: Packet-splitting, Kalman filter, Sign of innovations

I. INTRODUCTION

Networked Control Systems (NCSs) are control systems in which control loops are closed via networks. With the advantages resulting from using shared networks, such as low cost, system agility and self-configuration, NCSs have been applied in a wide range of areas. However, networks between distributed components also introduce new challenges, one of which is packet losses. Typically, packet losses are caused by transmission errors in physical network links, packet collisions or buffer overflows due to packet congestion [1]. When data dropouts occur, the performance of closed-loop NCSs may deteriorate and the closed-loop system may even become unstable. Thus, the effect of packet losses on the closed loop system can not be neglected.

The problem of estimation over packet-dropping networks has received significant attention recently. Sinopoli et al. [2] proposed a modified Kalman filter (MKF) to adopt observation losses. They modeled packet dropouts as a Bernoulli process and studied the statistical properties of the MKF. They proved the existence of a critical value for the packet loss rate, beyond which $\mathbb{E}[P_k]$ is unbounded. They also gave upper and lower bounds on $\mathbb{E}[P_k]$. Liu and Goldsmith [3] extended the results in [2] to allow partial observation losses.

The work by J. Wu and L. Shi was supported by HKUST DAG08/09.EG06.

∗ : Department of Eletronic and Computer Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. E-mails: {jfwu, eesling}@ust.hk

† : School of Electrical and Electrical Engineering, Nanyang Technological University, Singapore. E-mail: elhxie@ntu.edu.sg

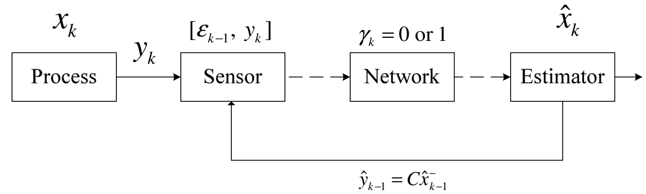

Fig. 1. Estimation over a Lossy network

Huang and Dey [4] described packet losses using a two state Markov chain. The authors introduced the notion of peak covariance and gave sufficient conditions for the stability of a peak covariance process in general vector case. In [5], Smith et al. modeled the packet loss process as a Markovian jump linear system (MJLS). Instead of using a time-varying Kalman filter they proposed a computationally simpler estimator to cope with packet losses. Other researchers also employed MJLSs to model NCSs over lossy networks, such as [6], [7], etc.

Xu and Hespanha [8] proposed an LTI estimation framework, in which local estimates are sent to the remote estimator. Their work assumed suffient local computational capability of the sensors. Similar result can be found in [9]. A. Ribeiro et al. [10] developed a distributed Kalman filter using the sign of innovations (SOI-KF) based on binary observations. In terms of performance and computational complexity it is comparable to the standard Kalman filter which is based on the original observations. By extending the SOI-KF, K. You et al. [11] developed a very general multi-level quantized innovation Kalman filter (MLQ-KF) for linear discrete-time stochastic systems. By optimizing the filter with respect to quantization levels, they obtained a close to optimal estimator MLQ-KF.

In this paper, we consider the problem of state estimation (Fig. 1) for linear discrete-time stochastic systems and propose an innovative packet-splitting Kalman filter (PSKF). Following the principle of SOI-KF in [10], we derive the PSKF. It is proved to be able to improve the estimation performance without requiring extra communication bandwidth and energy budget. What it needs is a buffer to store the sign of the previous innovation. The present work shows that if converges, $\mathbb{E}[P_k]$ of the PSKF converges to a smaller value than that of the MKF. Hence the PSKF is able to tolerate a higher or equal data loss rate than the MKF proposed in [2]. A smart sensor runs a Kalman filter and sends local estimates to the remote estimator. Although the estimation performance

978-1-4577-0079-8/11/$26.00 ©2011 AACC

of the smart sensor Kalman filter (SSKF) is better than that of the PSKF, it requires sufficient computation capacity.

The remainder of this paper is organized as follows. In Section II, we provide the mathematical models. In Section III, some frequently used notations are defined and a quick review of the MKF and SOI-KF is given. In Section IV, we propose an innovative packet-splitting transmission scheme and derive the corresponding optimal estimator. Performance analysis on the PSKF is carried out in Section V. In Section VI, we consider a simple scalar example to demonstrate the theory. Conclusion and results are summarized in Section VII.

Notation: $\mathbb{S}_+^n$ is the set of $n \times n$ positive semi-definite matrices. When $X \in \mathbb{S}_+^n$, we simply write $X \geq 0$; and when X is positive definite, we write $X > 0$. $\mathbb{R}_+$ is the set of positive real numbers. $\mathbb{R}^{n\times n}$ and $\mathbb{R}^n$ are the sets of $n \times n$ real matrices and $n \times 1$ real vectors respectively. We use $f(x|y)$ to denote the probability density function (pdf) of a random variable x given a random variable y. $\mathrm{N}(\mu, \Sigma)$ stands for Gaussian distribution with mean μ and covariance matrix Σ.

II. PROBLEM FORMULATION

Consider the following system in Fig. 1:

$$x_{k+1} = Ax_k + w_k, \tag{1}$$

$$y_k = h^T x_k + v_k, \tag{2}$$

where $x_k \in \mathbb{R}^n$ is the state vector, $y_k \in \mathbb{R}$ is the observation, $w_k \in \mathbb{R}^n$ and $v_k \in \mathbb{R}$ are mutually uncorrelated Gaussian white noises with zero mean and covariances $Q \in \mathbb{S}_+^n$, $r \in \mathbb{R}_+$ respectively. The initial state x_0 is assumed to be a Gaussian vector with zero mean and variance P_0. Furthermore, x_0 is uncorrelated with both w_k and v_k. We also assume that (A, h^T) is observable and $(A, \sqrt{Q})$ is controllable.

Suppose that the network inserted between the sensor and the estimator has an ideal communication path except packet dropouts. Let γ_k be the indicator of packet drop at time k, i.e., if the packet transmitted at time k is dropped, $\gamma_k = 0$, and otherwise $\gamma_k = 1$. We assume the packet arrival is an independent, identically distributed (i.i.d) Bernoulli random process, and $\mathbb{E}[\gamma_k] = \gamma$, where $\gamma \in [0, 1]$.

Assume the sensor can access the output prediction $\hat{y}_k = h^T A\hat{x}_{k-1}$ that is broadcasted by the estimator. For instance, the estimator is powered by external power source and has enough energy to broadcast the output prediction $\hat{y}_k$ to the sensor node once it is available. We also assume the sensor is able to compute the innovation $\varepsilon_k = y_k - \hat{y}_k$ and store its sign in a buffer such that it can be sent to the estimator at time $k + 1$.

Suppose that each packet sent by the sensor contains N bits, and N is sufficiently large so that the difference of quantized errors between an N-level and an (N-1)-level logarithmic quantizer is insignificant and neglectable. In [12], Minyue Fu et al. illustrated that the improvement achieved by increasing the number of quantization bits is marginal when $N \geq 4$. Thus, such an assumption is not as strong as one may imagine. Traditionally, all N bits are used to quantize the current innovation ε_k, whereas we will partition every packet into two portions: in the first portion, one bit is used to indicate the sign of previous innovation ε_{k-1}; the current observation y_k is quantized by the remaining $N-1$ bits in the second portion. In this paper, we will treat the "one bit" in the first portion as a nonlinear coarse quantizer and launch our study on it with the assistance of SOI-KF introduced in [10]. We will ignore quantization effect generated by the (N-1)-level quantizer in the second portion.

The main questions to be solved in this paper are:

1) how can we compute the best estimate $\hat{x}_k$ of x_k in (1) using such packet-splitting scheme?
2) does this scheme lead to better estimation performance compared with existing known results in existing literature?

The answers will be provided in subsequent sections.

III. PRELIMINARIES

A. Definitions

The following terms are frequently used in subsequent sections. Assume that A, h, Q, r are the same as they appear in section II. We define the functions t and g_λ: $\mathbb{S}_+^n \rightarrow \mathbb{S}_+^n$ as follows:

$$t(X) \triangleq AXA^T + Q, \tag{3}$$

$$g_\lambda(X) \triangleq X - \lambda Xh[hXh^T + r]^{-1}h^T X. \tag{4}$$

For functions f, f_1, f_2: $\mathbb{S}_+^n \rightarrow \mathbb{S}_+^n$, $f_1 \circ f_2$ and f^k are defined as

$$f_1 \circ f_2(X) \triangleq f_1(f_2(X)),$$

$$f^k(X) \triangleq f(\underbrace{f \circ \cdots \circ f}_{k-1}(X)).$$

We also define the function $\tilde{g}_\lambda$: $\mathbb{S}_+^n \rightarrow \mathbb{S}_+^n$ as

$$\tilde{g}_\lambda(X) = g_\lambda \circ t(X).$$

In the following text, if $\lambda = 1$, $g_{\lambda=1}$ and $\tilde{g}_{\lambda=1}$ will be written as g and $\tilde{g}$ for brevity.

Define $\mathbf{Y}_k$ as all received data by the estimator up to time step k, and define

$$\hat{x}_k^- \triangleq \mathbb{E}[x_k|\mathbf{Y}_{k-1}],\ e_k^- \triangleq \hat{x}_k^- - x_k,\ P_k^- \triangleq \mathbb{E}[e_k^- e_k^{-T}|\mathbf{Y_{k-1}}],$$

$$\hat{x}_k \triangleq \mathbb{E}[x_k|\mathbf{Y}_k], \quad e_k \triangleq \hat{x}_k - x_k,\ \ P_k \triangleq \mathbb{E}[e_k e_k^T|\mathbf{Y_k}],$$

The following Lemma is from [2].

Lemma 3.1: The following statements are true.

a) If $0 \leq X \leq Y$, then $\tilde{g}_\lambda(X) \leq \tilde{g}_\lambda(Y)$, $t(X) \leq t(Y)$.

b) If $0 \leq \lambda_1 \leq \lambda_2$, then $\tilde{g}_{\lambda_1}(X) \geq \tilde{g}_{\lambda_2}(X)$.

c) If X is a random variable, then $\mathbb{E}[g_\lambda(X)] \leq g_\lambda(\mathbb{E}[X])$.

B. Kalman Filtering with Intermittent Observations

Sinopoli et al. [2] proposed the MKF to compute the pair $(\hat{x}_k^-, P_k^-)$ and $(\hat{x}_k, P_k)$ over packet lossy networks. Both the time and measurement updates are implemented when a packet is perfectly received. If a packet is dropped, only the time update is performed at that time step. The MKF is given by the following set of equations:

$$\begin{cases} \hat{x}_k^- = A\hat{x}_{k-1}, \\ P_k^- = AP_{k-1}A^T + Q, \\ K_k = P_k^- h[h^T P_k^- h + R]^{-1}, \\ \hat{x}_k = \hat{x}_k^- + \gamma_k K_k (y_k - h^T \hat{x}_k^-), \\ P_k = (I - \gamma_k K_k h^T) P_k^-. \end{cases}$$

For ease of reference, we present Theorem 2 from [2] as follows:

Lemma 3.2: Consider the system described by (1) and (2). For the MKF, if A is unstable, then there exists a $\overline{\lambda} \in [0,1)$ such that

$$\lim_{k\to\infty} \mathbb{E}[P_k^-] = +\infty, \ \exists\, P_0 \geq 0, \ \forall\, 0 \leq \gamma \leq \overline{\lambda},$$

$$\lim_{k\to\infty} \mathbb{E}[P_k^-] < M_{P_0}, \ \forall\, P_0 \geq 0, \ \forall\, \overline{\lambda} < \gamma \leq 1,$$

where M_{P_0} depends on the initial condition P_0.

When $\gamma_k \equiv 1$ for $k = 1, 2, \ldots$, the MKF reduces to the standard Kalman filter. Since (A, h^T) is observable and $(A, \sqrt{Q})$ is controllable, then there exists $\overline{P} \geq 0$ such that $\overline{P} = t \circ g_{\lambda=1}(\overline{P})$. Define $\tilde{P} \geq 0$ as $\tilde{P} = g_{\lambda=1}(\overline{P})$. Then we have

$$\tilde{P} = g_{\lambda=1} \circ (t \circ g_{\lambda=1}(\overline{P})) = g_{\lambda=1} \circ t(\tilde{P}) = \tilde{g}_{\lambda=1}(\tilde{P}),$$

that is, $\lim_{k\to\infty} P_k = \tilde{P}$.

C. SOI-KF: Kalman Filtering Using the Sign of Innovations

In [10], A. Ribeiro et al. proposed a recursive algorithm for distributed state estimation based on the sign of innovations (SOI). Their goal was to achieve low cost communication by transmitting a single bit per observation.

The authors defined the message b_k as SOI:

$$b_k \triangleq sign[y_k - \hat{y}_k] = \begin{cases} +1, & \text{if } y_k \geq \hat{y}_k, \\ -1, & \text{if } y_k < \hat{y}_k. \end{cases} \quad (5)$$

The focus of [10] was to study the minimum mean squared error (MMSE) estimator of x_k based on $b_{0:k} \triangleq [b_0, \ldots, b_k]^T$. The author first proposed the exact MMSE estimator of which the expenditure is unaffordable for resource limited NCSs. Then a reduced-complexity approximation of the MMSE estimator was motivated to pursue. Approximating $f(x_k|b_{0:k-1})$ as a Gaussian distribution, that is, $f(x_k|b_{0:k-1}) \sim \mathrm{N}(\hat{x}_k^-, P_k^-)$, they obtained an approximate SOI-KF which is coarse but simple and efficient. The SOI-KF is described in the following lemma. For detailed proof, please refer to Proposition 2 in [10].

Lemma 3.3: Consider system (1), (2) and the message b_k defined in (5). Assume $f(x_k|b_{0:k-1}) \sim \mathrm{N}(\hat{x}_k^-, P_k^-)$, then the MMSE estimator can be obtained recursively as follows:

$$\hat{x}_k^- = A\hat{x}_{k-1}, \quad (6)$$

$$P_k^- = t(P_{k-1}), \quad (7)$$

$$\hat{x}_k = \hat{x}_k^- + \sqrt{2/\pi} P_k^- h (h^T P_k^- h + r)^{-0.5} b_k, \quad (8)$$

$$P_k = g_{\lambda=\frac{2}{\pi}}(P_k^-). \quad (9)$$

The main result in this paper is presented in the next two sections.

IV. PACKET-SPLITTING KALMAN FILTER WITHOUT EXTRA COST

To begin with, we will modify the definition of message b_k in [10] a little bit such that it is more appropriate from a theoretical angle.

$$b_k \triangleq sign[y_k - \hat{y}_k] = \begin{cases} +1, & \text{if } y_k > \hat{y}_k, \\ -1, & \text{if } y_k < \hat{y}_k. \end{cases} \quad (10)$$

When $\varepsilon_k = 0$, b_k is a random variable with the probability $\mathbf{Pr}(b_k = 1|\varepsilon_k = 0) = 0.5$, and $\mathbf{Pr}(b_k = -1|\varepsilon_k = 0) = 0.5$.

Let us denote the N-bit packet transmitted at time k by I_k. We partition I_k into two parts: one part consists of one bit which indicates the sign of the previous innovation ε_{k-1}, i.e., b_{k-1}; the other part contains $N-1$ bits which represent the recent measurement y_k. In other words, $I_k = [b_{k-1}, y_k]$. Intuitively, if I_{k-1} is dropped, the performance of the estimator could be improved by estimating x_{k-1} based on b_{k-1} as long as I_k is able to reach the estimator, as b_{k-1} contains some useful information of y_{k-1}. Thus the proposed scheme improves the performance of the estimator by reducing estimation error without requiring extra communication bandwidth or energy cost. What it needs is a one-bit buffer to store the sign of the pervious innovation.

We further define $\hat{x}_k^+$, e_k^+ and P_k^+ at time step $k+1$ as follows:

$$\hat{x}_k^+ \triangleq \mathbb{E}[x_k|\mathbf{Y}_k, \gamma_{k+1} b_k],$$

$$e_k^+ \triangleq \hat{x}_k - \hat{x}_k^+,$$

$$P_k^+ \triangleq \mathbb{E}[e_k^+ e_k^{+T}|\mathbf{Y}_k, \gamma_{k+1} b_k].$$

Following a traditional simplification in non-linear filtering, we approximate $f(x_k|\gamma_1 I_1, ..., \gamma_{k-1} I_{k-1}) \sim \mathrm{N}(\hat{x}_k^-, P_k^-)$, and next proposition gives an optimal linear MMSE estimation of $\hat{x}_k^+$. In section VI, we will verify the validity of this Gaussian approximation using a simple scalar example.

Proposition 4.1: Consider the system (1), (2) and assume $\gamma_{k-1} = 0$, $\gamma_k = 1$. If $f(x_k|\gamma_1 I_1, ..., \gamma_{k-1} I_{k-1}) \sim \mathrm{N}(\hat{x}_k^-, P_k^-)$, $\hat{x}_{k-1}^+$ and P_{k-1}^+ are computed based on $\hat{x}_{k-1}^-$ and P_{k-1}^- as follows:

$$\hat{x}_{k-1}^+ = \hat{x}_{k-1}^- + \sqrt{2/\pi} P_{k-1}^- h (h^T P_{k-1}^- h + r)^{-0.5} b_{k-1}, \quad (11)$$

$$P_{k-1}^+ = g_{\lambda=\frac{2}{\pi}}(P_{k-1}^-). \quad (12)$$

Proof: A direct result from the SOI-KF proposed in Lemma 3.3. ■

The following theorem illustrates our proposed packet-splitting Kalman filter (PSKF). At time step $k+1$, the estimator first re-estimates the state, that is, compute $\hat{x}_k^+$ from $\hat{x}_k$ given b_k and γ_{k+1}. Then the time update is carried out, i.e., $\hat{x}_{k+1}^-$ is computed. After that $\hat{x}_{k+1}$ is computed based on $\hat{x}_{k+1}^-$, y_{k+1} and γ_{k+1}. Such an iterative calculation of $(P_k^-, \hat{x}_k^-)$, $(P_k, \hat{x}_k)$ and $(P_k^+, \hat{x}_k^+)$ is given by the following theorem.

Theorem 4.2: In the PSKF, at time $k+1$, $(P_k^+, \hat{x}_k^+)$, $(P_{k+1}^-, \hat{x}_{k+1}^-)$ and $(P_{k+1}, \hat{x}_{k+1})$ can be computed recursively as follows:

$$\hat{x}_k^+ = \hat{x}_k + (1-\gamma_k)\gamma_{k+1}\sqrt{2/\pi}P_k h(h^T P_k h + r)^{-0.5} b_k,$$
$$P_k^+ = g_{\lambda=\frac{2}{\pi}(1-\gamma_k)\gamma_{k+1}}(P_k),$$
$$\hat{x}_{k+1}^- = A\hat{x}_k^+,$$
$$P_{k+1}^- = t(P_k^+),$$
$$K_{k+1} = P_{K+1}^- h[h^T P_{k+1}^- h + r]^{-1},$$
$$\hat{x}_{k+1} = \hat{x}_{k+1}^- + \gamma_{k+1}K_{k+1}(y_{k+1} - h^T\hat{x}_{k+1}^-),$$
$$P_{k+1} = g_{\gamma_{k+1}}(P_{k+1}^-).$$

Proof: Considering whether packets are dropped at time step k and $k+1$ or not, four cases are needed to consider:

Case 1: $\gamma_k = 0$, $\gamma_{k+1} = 0$. Since no new information reaches the estimator at time $k+1$, $\hat{x}_k^+$ is the same as $\hat{x}_k$ and the measurement update is skipped, which is the same if we substitute $\gamma_k = 0$ and $\gamma_{k+1} = 0$ into the equations of Theorem 4.2, i.e.,

$$\hat{x}_k^+ = \hat{x}_k,$$
$$P_k^+ = P_k,$$
$$\hat{x}_{k+1} = \hat{x}_{k+1}^- = A\hat{x}_k^+,$$
$$P_{k+1} = P_{k+1}^- = t(P_k^+).$$

Case 2: $\gamma_k = 0$, $\gamma_{k+1} = 1$. b_k is received as packet I_{k+1} arrives at the estimator. Thus, the SOI-KF based on b_k is implemented first, followed by the MKF:

$$\hat{x}_k^+ = \hat{x}_k + \sqrt{2/\pi}P_k h(h^T P_k h + r)^{-0.5} b_k,$$
$$P_k^+ = g_{\lambda=\frac{2}{\pi}}(P_k),$$
$$\hat{x}_{k+1}^- = A\hat{x}_k^+,$$
$$P_{k+1}^- = t(P_k^+),$$
$$K_{k+1} = P_{k+1}^- h[h^T P_{k+1}^- h + r]^{-1},$$
$$\hat{x}_{k+1} = \hat{x}_{k+1}^- + K_{k+1}(y_{k+1} - h^T\hat{x}_{k+1}^-),$$
$$P_{k+1} = g(P_{k+1}^-).$$

This agrees with Theorem 4.2.

Case 3: $\gamma_k = 1$, $\gamma_{k+1} = 1$. The optimal estimation reduces to the MKF.

$$\hat{x}_k^+ = \hat{x}_k,$$
$$P_k^+ = P_k,$$
$$\hat{x}_{k+1}^- = A\hat{x}_k^+,$$
$$P_{k+1}^- = t(P_k^+),$$
$$K_{k+1} = P_{k+1}^- h[h^T P_{k+1}^- h + r]^{-1},$$
$$\hat{x}_{k+1} = \hat{x}_{k+1}^- + K_{k+1}(y_{k+1} - h^T\hat{x}_{k+1}^-),$$
$$P_{k+1} = g(P_{k+1}^-).$$

Again this agrees with Theorem 4.2.

Case 4: $\gamma_k = 1$, $\gamma_{k+1} = 0$. No new information comes to the estimator at time $k+1$. The optimal estimation is the same as the one in Case 1. ∎

Remark 4.3: To remove confusion, we denote $\hat{x}_k$, e_k, P_k computed in our proposed PSKF as $\hat{x}_k^{PSKF}$, e_k^{PSKF}, P_k^{PSKF}. Similarly, in the MKF and SSKF, they are denoted as $\hat{x}_k^{MKF}$, e_k^{MKF}, P_k^{MKF} and $\hat{x}_k^{SSKF}$, e_k^{SSKF}, P_k^{SSKF} respectively.

Remark 4.4: In packet-dropping networks, unlike traditional estimator, due to randomness of γ_k, P_k is a random variable, therefore, we investigate the statistical properties of P_k. In this paper, we will focus on studying $\mathbb{E}[P_k^{PSKF}]$.

V. PERFORMANCE ANALYSIS

A. Convergence Analysis

Recall $\overline{\lambda}$ is the critical value in Lemma 3.2.

Lemma 5.1: If A is unstable, for $\forall\, \gamma \in (\overline{\lambda}, 1]$, we have

$$\lim_{k\to\infty} \mathbb{E}[P_k^{MKF}] \le \overline{M}_{P_0},\ \ \forall\ P_0^{MKF} \ge 0,$$

where $\overline{M}_{P_0} = g_\gamma(M_{P_0})$.

Proof: From Lemma 3.2, we obtain

$$\begin{aligned}\lim_{k\to\infty} \mathbb{E}[P_k^{MKF}] &= \lim_{k\to\infty} \mathbb{E}[g_{\gamma_k}(P_k^{MKF-})] \\ &= \lim_{k\to\infty} \mathbb{E}[g_{\gamma}(P_k^{MKF-})] \\ &\le \lim_{k\to\infty} g_\lambda(\mathbb{E}[P_k^{MKF-}]) \\ &\le g_\gamma(M_{P_0}) \triangleq \overline{M}_{P_0}\end{aligned}$$

where we used the property c) from Lemma 3.1. ∎

Lemma 5.2: Given an unstable matrix A, consider the sequence $V_k = (1-\lambda)t(V_{k-1}) + \lambda\tilde{g}^k(V_0)$, with any $V_0 \ge 0$ and $\lambda \in [0, 1)$. For $\underline{\lambda} \triangleq 1 - \frac{1}{\rho(A)^2} \in [0, 1)$, the following statement holds,

$$\lim_{k\to\infty} V_k = +\infty,\ \ \forall\, V_0 \ge 0,\ \ \forall\, 0 \le \lambda \le \underline{\lambda},$$
$$\lim_{k\to\infty} V_k = \overline{V},\ \ \ \ \forall\, V_0 \ge 0,\ \ \forall\, \underline{\lambda} < \lambda \le 1,$$

where $\rho(A)$ is the spectral radius of A and $\overline{V} \ge 0$ is the unique solution of the equation

$$\overline{V} = (1-\lambda)t(\overline{V}) + \lambda\tilde{P}. \qquad (13)$$

Proof: We shall first prove this lemma with V_0=0.

1) $\forall 0 \le \lambda \le \underline{\lambda} \implies \lim_{k\to\infty} V_k = +\infty$.
   Define a sequence $\{W_k\}_0^\infty$ satisfying $W_k = (1-\lambda)t(W_{k-1})$ with $W_0 = 0$. Let $\tilde{A} = \sqrt{1-\lambda}A$ and $\tilde{Q} = (1-\lambda)Q$. As $(\tilde{A}, \sqrt{\tilde{Q}})$ is controllable, the Lyapunov equation $X = \tilde{A}X\tilde{A}^T + \tilde{Q}$ has a unique positive semi-definite solution iff $\rho(\tilde{A}) < 1$, i.e., $\lambda > 1 - \frac{1}{\rho(A)^2} = \underline{\lambda}$. If $0 \le \lambda \le \underline{\lambda}$, it is impossible for W_k to converge to a finite matrix, otherwise $X = \tilde{A}X\tilde{A}^T + \tilde{Q}$ has a unique positive semi-definite solution. In other words, $\lim_{k\to\infty} W_k = +\infty$.
   Next $W_1 \le V_1$ and $W_k \le V_k$ imply that
   $$W_{k+1} = (1-\lambda)t(W_k) \le (1-\lambda)t(V_k) + \lambda\tilde{g}^k(V_0) = V_{k+1}$$
   By induction, $W_k \le V_k\ \ \forall k = 0, 1, 2, \ldots$. Thus, when $0 \le \lambda \le \underline{\lambda}$, $\lim_{k\to\infty} V_k = +\infty$.
2) $\forall \underline{\lambda} < \lambda \le 1 \implies \lim_{k\to\infty} V_k = \overline{V}$.
   It suffices to show that $\{V_k\}_0^\infty$ is bounded and monotonic increasing.
   If $\underline{\lambda} < \lambda \le 1$, $\tilde{A}$ is stable, then the Lyapunov equation $X = \tilde{A}X\tilde{A}^T + \tilde{Q} + \lambda\tilde{P}$ has a unique positive semi-definite solution $\overline{V} \ge 0$.

It is easy to obtain $\tilde{g}^k(0) \leq \tilde{P}$. Therefore,

$$V_k = (1-\lambda)t(V_{k-1}) + \lambda\tilde{g}^k(0) \leq (1-\lambda)t(V_{k-1}) + \lambda\tilde{P} \leq \overline{V},$$

i.e., V_k is bounded.

Clearly, $V_1 = (1-\lambda)t(V_0) + \lambda\tilde{g}(V_0) = g_\lambda(Q) \geq V_0$, and $V_k \geq V_{k-1}$ implies

$$V_{k+1} = (1-\lambda)t(V_k) + \lambda\tilde{g}(V_k) \geq (1-\lambda)t(V_{k-1}) + \lambda\tilde{g}(V_{k-1}) = V_k.$$

By induction, $\{V_k\}_0^\infty$ is nondecreasing. Therefore, V_k converges and

$$\lim_{k\to\infty} V_k = (1-\lambda)t(\lim_{k\to\infty} V_k) + \lambda \lim_{k\to\infty} \tilde{g}^k(V_0) = (1-\lambda)t(\lim_{k\to\infty} V_k) + \lambda\tilde{P}.$$

Thus, $\lim_{k\to\infty} V_k = \overline{V}$.

Finally, it is straightforward to extend the above analysis to $V \geq 0$ as $\lim_{k\to\infty} \tilde{g}^k(V_0) = \tilde{P}$. ■

Theorem 5.3: If A is unstable, then there exists a $\gamma_c \in [0,1)$ such that

$$\lim_{k\to\infty} \mathbb{E}[P_k^{PSKF}] = +\infty,\ \exists\, P_0^{PSKF} \geq 0,\ \forall\, 0 \leq \gamma \leq \gamma_c,$$
$$\lim_{k\to\infty} \mathbb{E}[P_k^{PSKF}] < +\infty,\ \forall\, P_0^{PSKF} \geq 0,\ \forall\, \gamma_c < \gamma \leq 1,$$

and γ_c is bounded by $\underline{\lambda}$ and $\overline{\lambda}$, i.e., $\underline{\lambda} \leq \gamma_c \leq \overline{\lambda}$.

Proof: First, consider the sequence $V_k = (1-\gamma)t(V_{k-1}) + \gamma\tilde{g}^k(V_0)$, $V_0 = P_0^{PSKF} \geq 0$. If $0 \leq \gamma \leq \underline{\lambda}$, from Lemma 5.2 $\lim_{k\to\infty} V_k = +\infty$.

It is clear that

$$\mathbb{E}[P_1^{PSKF}] = (1-\gamma)t(P_0^{PSKF}) + \gamma\tilde{g}(P_0^{PSKF}) = V_1, \qquad \forall\, P_0^{PSKF} = V_0 \geq 0.$$

Moreover, $\mathbb{E}[P_k^{PSKF}] \geq V_k$ implies

$$\begin{aligned}\mathbb{E}[P_{k+1}^{PSKF}] &= (1-\gamma)\mathbb{E}[t(P_k^{PSKF})] + \gamma\mathbb{E}[\tilde{g}(P_k^{PSKF})]\\ &\geq (1-\gamma)t(\mathbb{E}[P_k^{PSKF}]) + \gamma\tilde{g}(\mathbb{E}[P_k^{PSKF}])\\ &\geq (1-\gamma)t(V_k) + \gamma\tilde{g}(V_k) = V_{k+1}.\end{aligned}$$

By applying induction, we obtain $\mathbb{E}[P_k^{PSKF}] \geq V_k\ \forall\ k = 1,2,\ldots$. It implies that $\mathbb{E}[P_k^{PSKF}]$ is unbound for any $0 \leq \gamma \leq \underline{\lambda}$. Therefore $\gamma_c \geq \underline{\lambda}$.

Now consider $\gamma > \overline{\lambda}$. P_k^{PSKF} and P_k^{MKF} can be shown to satisfy a recursive algorithm as:

$$P_{k+1}^{PSKF} = g_{\gamma_{k+1}} \circ t \circ g_{\theta_k}(P_k^{PSKF}) \tag{14}$$

and

$$P_{k+1}^{MKF} = g_{\gamma_{k+1}} \circ t(P_k^{MKF}) \tag{15}$$

where $0 \leq \theta_k = \frac{2(1-\gamma_k)\gamma_{k+1}}{\pi} \leq 1$.

From (14) and (15), we have

$$P_k^{MKF} = (1-\gamma_k)t \circ \tilde{g}_{\gamma_{k-1}} \circ \cdots \circ \tilde{g}_{\gamma_1}(P_0^{MKF}) + \gamma_k\tilde{g}_{\gamma_k} \circ \cdots \circ \tilde{g}_{\gamma_1}(P_0^{MKF})$$

and

$$P_k^{PSKF} = (1-\gamma_k)t \circ \tilde{g}_{\gamma_{k-1}+\theta_{k-1}} \circ \cdots \circ \tilde{g}_{\gamma_1+\theta_1}(P_0^{PSKF}) + \gamma_k\tilde{g}_{\gamma_k} \circ \tilde{g}_{\gamma_{k-1}+\theta_{k-1}} \circ \cdots \circ \tilde{g}_{\gamma_1+\theta_1}(P_0^{PSKF}),$$

where $\gamma_k \leq \gamma_k + \theta_k \leq 1$. By properties a) and b) in Lemma 3.1, it is easy to see $P_k^{MKF} \geq P_k^{PSKF}$. Moreover, it is shown in Theorem 5.1 that $\mathbb{E}[P_k^{MKF}] \leq \overline{M}_{P_0},\ \forall P_0 \geq 0$. Thus, $\mathbb{E}[P_k^{PSKF}] \leq \mathbb{E}[P_k^{MKF}] \leq \overline{M}_{P_0}$. This implies that $\gamma_c \leq \overline{\lambda}$. The proof is complete. ■

B. Performance Comparison with the MKF and SSKF

Now consider the SSKF in which the sensor has sufficient computation capability to compute its state estimate $\hat{x}_k^s$ locally and sends $\hat{x}_k^s$ to the remote estimator. At the estimator side, $(\hat{x}_k,\ P_k^{SSKF})$ is computed as follows:

$$(\hat{x}_k,\ P_k^{SSKF}) = \begin{cases}(A\hat{x}_{k-1},\ t(P_{k-1}^{SSKF})), & \text{if } \gamma_k = 0\\ (\hat{x}_k^s,\ P_k^s), & \text{if } \gamma_k = 1\end{cases} \tag{16}$$

In particular, P_k^{SSKF} can be written as

$$P_k^{SSKF} = (1-\gamma_k)t(P_{k-1}^{SSKF}) + \gamma_k\tilde{g}^k(P_0^{SSKF}). \tag{17}$$

Remark 5.4: Taking expectation at both side of (17), we obtain $\mathbb{E}[P_k^{SSKF}] = (1-\gamma)t(\mathbb{E}[P_{k-1}^{SSKF}]) + \gamma\tilde{g}^k(P_0^{SSKF})$. Hence,$\{\mathbb{E}[P_k^{SSKF}]\}_0^\infty$=$\{V_k\}_0^\infty$, if $P_0^{SSKF} = V_0$. Applying Lemma 5.2,

$$\lim_{k\to\infty} \mathbb{E}[P_k^{SSKF}] = +\infty,\ \forall\, \mathbb{E}[P_0^{SSKF}] \geq 0,\ \forall\, 0 \leq \lambda \leq \underline{\lambda},$$
$$\lim_{k\to\infty} \mathbb{E}[P_k^{SSKF}] = \overline{V}, \quad \forall\, \mathbb{E}[P_0^{SSKF}] \geq 0,\ \forall\, \underline{\lambda} < \lambda \leq 1.$$

The following theorem shows that if the MKF, PSKF and SSKF all converge, then the *a posteriori* estimate error covariance of them descend orderly.

Theorem 5.5: Assume $\exists\ \gamma > 0$, such that $\mathbb{E}[P_k^{MKF}]$, $\mathbb{E}[P_k^{PSKF}]$ and $\mathbb{E}[P_k^{SSKF}]$ are all convergent, then

$$\lim_{k\to\infty} \mathbb{E}[P_k^{MKF}] \geq \lim_{k\to\infty} \mathbb{E}[P_k^{PSKF}] \geq \lim_{k\to\infty} \mathbb{E}[P_k^{SSKF}]$$

for any $P_0^{MKF} = P_0^{PSKF} = P_0^{SSKF} \geq 0$. In particular, if $\forall\ k = 1,2,\ldots,\ \gamma_k \equiv 0$ (or $\gamma_k \equiv 1$),

$$\lim_{k\to\infty} \mathbb{E}[P_k^{MKF}] = \lim_{k\to\infty} \mathbb{E}[P_k^{PSKF}] = \lim_{k\to\infty} \mathbb{E}[P_k^{SSKF}].$$

Proof: In Theorem 5.3, the following is proved,

$$\lim_{k\to\infty} \mathbb{E}[P_k^{MKF}] \geq \lim_{k\to\infty} \mathbb{E}[P_k^{PSKF}].$$

From (17), we obtain

$$P_k^{SSKF} = (1-\gamma_k)t(P_{k-1}^{SSKF}) + \gamma_k\tilde{g}^k(P_0^{SSKF}).$$

Clearly, $\mathbb{E}[P_1^{SSKF}] = \mathbb{E}[P_1^{PSKF}]$ and

$$\begin{aligned}\mathbb{E}[P_2^{SSKF}] &= (1-\gamma)t(\mathbb{E}[P_1^{SSKF}]) + \gamma\tilde{g}^2(P_0^{SSKF})\\ &\leq (1-\gamma)t(\mathbb{E}[P_1^{PSKF}])\\ &\quad + \gamma\mathbb{E}[\tilde{g}_{\gamma_2} \circ \tilde{g}_{\gamma_1+\theta_1}(P_0^{PSKF})]\\ &= \mathbb{E}[P_2^{PSKF}].\end{aligned}$$

For $k \geq 2$, $\mathbb{E}[P_k^{SSKF}] \leq \mathbb{E}[P_k^{PSKF}]$ implies

$$\begin{aligned}&\mathbb{E}[P_{k+1}^{SSKF}] = (1-\gamma)t(\mathbb{E}[P_k^{SSKF}]) + \gamma\tilde{g}^{k+1}(P_0^{SSKF})\\ &\leq (1-\gamma)t(\mathbb{E}[P_k^{PSKF}]) + \gamma\mathbb{E}[\tilde{g}_{\gamma_{k+1}} \circ \tilde{g}_{\gamma_k+\theta_k} \circ \cdots \circ\\ &\quad \tilde{g}_{\gamma_1+\theta_1}(P_0^{PSKF})] = \mathbb{E}[P_{k+1}^{PSKF}].\end{aligned}$$

By induction, it is easy to show that

$$\lim_{k\to\infty} \mathbb{E}[P_k^{PSKF}] \geq \lim_{k\to\infty} \mathbb{E}[P_k^{SSKF}].$$

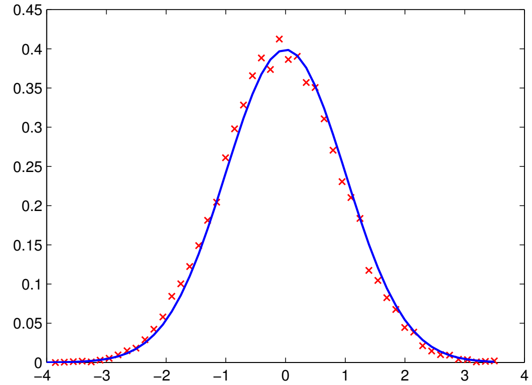

Fig. 2. the red cross shows the pdf of $(x_k - \hat{x}_k^-)(P_k^-)^{-0.5}$ and the blue solid line shows the pdf of standard Gaussian distribution.

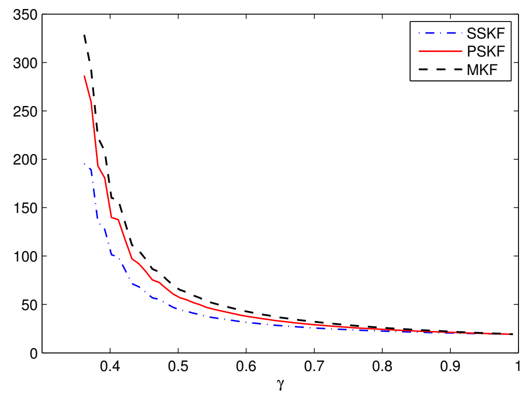

Fig. 3. $\mathbb{E}[P_k]$ of the MKF, PSKF and SSKF

In particular, if $\forall\ k = 0, 1, 2 \ldots\ \gamma_k \equiv 0\,(\text{or}\, \gamma_k \equiv 1)$, the MKF, PSKF and SSKF reduce to open-loop predictors (or standard KF), i.e.,

$$\lim_{k\to\infty} \mathbb{E}[P_k^{MKF}] = \lim_{k\to\infty} \mathbb{E}[P_k^{PSKF}] = \lim_{k\to\infty} \mathbb{E}[P_k^{SSKF}].$$

■

Remark 5.6: As shown in Theorem 5.5, the PSKF provides a tradeoff between the MKF and SSKF in terms of communication, computation cost and estimation quality.

VI. EXAMPLES

Consider a simple scalar system

$$x_{k+1} = 1.25x_k + w_k, \tag{18}$$

$$y_k = x_k + v_k. \tag{19}$$

where w_k and v_k have zero mean and variances $Q = 1$ and $r = 50$ respectively.

First we will verify $f(x_k|\gamma_1 I_1, ..., \gamma_{k-1} I_{k-1}) \sim \mathrm{N}(\hat{x}_k^-, P_k^-)$ approximately holds, which is the prerequisite of Proposition 4.1. Fig. 2 shows the pdf of the normalized empirical prediction error $(x_k - \hat{x}_k^-)(P_k^-)^{-0.5}$ obtained by a Monte Carlo simulation, along with the pdf of standard Gaussian distribution. We can observe these two curves fit each other very well, i.e., $f(x_k|\gamma_1 I_1, ..., \gamma_{k-1} I_{k-1}) \sim \mathrm{N}(\hat{x}_k^-, P_k^-)$ holds approximately.

As Fig. 3 shown, $\mathbb{E}[P_k^{PSKF}]$ is bounded by $\mathbb{E}[P_k^{MKF}]$ and $\mathbb{E}[P_k^{SSKG}]$ under the same γ. Note that in this example, $\mathbb{E}[P_k^{SSKF}]$, $\mathbb{E}[P_k^{PSKF}]$ and $\mathbb{E}[P_k^{MKF}]$ tend to infinity as λ approach $\gamma_c \approx 0.36$ simultaneously. This is because h^T is invertible, $\overline{\lambda} = 1 - \frac{1}{\rho(A)^2}$ as shown in [2]. For this system, $\gamma_c = \underline{\lambda} = \overline{\lambda} = 1 - \frac{1}{\rho(A)^2} = 0.36$. Fig. 3 clearly shows the transition at $\lambda = 0.36$.

VII. CONCLUSION

In this paper, we consider the problem of state estimation over lossy networks and propose an innovative packet-splitting approach for Kalman filtering. The PSKF provides a tradeoff between the MKF and SSKF in terms of resource usage and estimation quality. Without extra bandwidth or energy cost, the PSKF has a better performance and tolerates a higher or at least equal data loss rate than the MKF. Unlike the SSKF which requires sufficient computation capability of the sensor, the PSKF only needs a one-bit buffer. Thus the PSKF can be applied in a wider range, especially in a communication channel with a low packet arrival rate.

There are still some interesting works in the future. For example, although we have assumed i.i.d Bernoulli random process for the packet arrival rate, in the sensor network there generally exist correlations among the continuous packet dropouts. We also ignored the quantization error, and it will be of interest to see whether the quantization error has influence on the PSKF.

REFERENCES

[1] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 138–162, January 2007.

[2] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poola, M. Jordan, and S. Sastry, "Kalman filtering with intermittent observations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1453–1464, 2004.

[3] X. Liu and A. Goldsmith, "Kalman filtering with partial observation losses," in *Proceedings of the 43rd IEEE Coference on Decision and Control*, vol. 4, December 2004, pp. 4180 – 4186.

[4] M. Huang and S. Dey, "Stability of kalman filtering with markovian packet losses," *Automatica*, vol. 23, no. 4, pp. 598–607, April 2007.

[5] S. Smith and P. Seiler, "Estimation with lossy measurements: Jump estimators for jump system," *IEEE Transactions on Automatic Control*, vol. 48, no. 12, pp. 1453–1464, December 2003.

[6] P. Seiler and R. Sengupta, "Analysis of communication losses in vehicle control problems," in *Proceedings of American Control Conference*, vol. 2, June 2001, pp. 1491 – 1496.

[7] O. Costa, "Stationary filter for linear minimum mean square error estimator of discrete-time markovian jump systems," *IEEE Transactions on Automatic Control*, vol. 47, no. 8, pp. 1351–1456, August 2002.

[8] Y. Xu and J. P. Hespanha, "Estimation under uncontrolled and controlled communications in networked control systems," in *Proceedings of the 44th IEEE Conference on Decision and Control*, June 2005, pp. 842–847.

[9] V. Gupta, D. Spanos, B. Hassibi, and R. M. Murray, "On lqg control across a stochastic packet-dropping link," in *Proceedings of American Control Conference*, June 2005, pp. 360–365.

[10] G. B. G. A. Ribeiro and S. Roumeliotis, "Soi-kf: Distributed kalman filtering with low-cost communications using the sign of innovations," *IEEE Transactions on Signal Processing*, vol. 54, no. 12, pp. 4782–4795, December 2006.

[11] K. You, L. Xie, S. Sun, and W. Xiao, "Multiple-level quantized innovation kalman filter," in *Proceedings of the 17th International Federation of Automatic Control*, 2008.

[12] M. Fu and C. E. de Souza, "State estimation for linear discrete-time systems using quantized measurements," *Automatica*, vol. 45, no. 12, pp. 2937–2945, December 2009.