Remote State Estimation with Stochastic Event-triggered Sensor Schedule in the Presence of Packet Drops

Liang Xu, Yilin Mo and Lihua Xie

Abstract—This paper studies the remote state estimation problem of linear time-invariant systems with stochastic event-triggered sensor schedules in the presence of packet drops between the sensor and the estimator. It is shown that the system state conditioned on the available information at the estimator side is Gaussian mixture distributed. Minimum mean square error (MMSE) estimators are subsequently derived for both open-loop and closed-loop schedules. Since the optimal estimators require exponentially increasing computation and memory, sub-optimal estimators to reduce the computational complexities are further provided. In the end, simulations are conducted to illustrate the performance of the optimal and sub-optimal estimators.

I. INTRODUCTION

Sensor networks have wide applications in environment and habitat monitoring, industrial automation, smart buildings, etc. In many applications, sensors are battery powered and are required to reduce the energy consumption to prolong their service life. Sensor scheduling algorithms are therefore proposed as an efficient method by scheduled transmissions to reduce the communication frequency to prolong the service time of sensor devices. Sensor scheduling algorithms can be roughly categorized as off-line schedules and event-triggered schedules. The off-line schedules are designed based on the communication frequency requirement and the statistics of the systems [1], [2], [3]. Compared with off-line schedules, event-triggered schedules depend on both the statistics and the realization of the system, which are expected to achieve better performance than off-line ones. Many triggering rules have been proposed in the literature based on the condition that, the estimation error [4], error in predicated output [5], functions of the estimation error [6], [7], or the error covariance [8], exceeds a given threshold.

Wireless communications are mostly utilized in sensor networks, and packet drops are inevitable in wireless communications. Therefore, it is necessary to study how packet drops affect sensor scheduling algorithms [9], [3]. It should be noted that, for off-line schedulers and estimation error covariance based event-triggered schedulers, there is no need to distinguish between the channel loss event and the hold of transmission event when designing estimators. As long as the estimator receives the packet, it can conduct the measurement update to improve the estimate and vice versa. However, the case is different for the event-triggered sensor scheduling algorithms in [6], [7] where the sensor measurement is used as the trigger criterion and the hold of transmission event contains information about the sensor measurement. In the presence of possible channel losses, the estimator cannot decide whether the non-reception of the packet can be attributed to the sensor measurement or the channel loss. If it is due to that the sensor measurement lies below the given threshold, then this information can be leveraged to improve the estimate. However, if it is caused by the channel loss, the estimator will have no information about the sensor measurement and no update will be carried out. This fact complicates the optimal estimator design. Furthermore, it is proved that, in the presence of channel losses, the Gaussian properties with the stochastic event-triggered sensor scheduling algorithms in [7] no longer hold [10].

This paper considers the same problem setting as in [7] with the additional consideration of the presence of packet drops between the sensor and the estimator. We try to derive the MMSE estimator in the case that the estimator has no knowledge about the channel loss events and only knows the channel loss rate. We show that the conditional distributions of the system state at the estimator side are mixture Gaussian, based on which MMSE estimators are derived. Moreover, sub-optimal estimation algorithms to reduce computational complexities are provided. This paper is organized as follows. The problem formulation is given in Section II. The optimal estimators for the open loop scheduler case and the closed-loop scheduler case are studied in Section III and Section IV, respectively. Strategies to reduce the computational complexities are discussed in Section V. Simulations evaluations are given in Section VI. This paper ends with some concluding remarks in Section VII.

Notation: \( N_x(\bar{x}, \Sigma) \) denotes the Gaussian pdf of the random variable \( x \) with the mean \( \bar{x} \) and the covariance matrix \( \Sigma \). \( f(x) \) (Pr(\( x \))) denotes the probability density function (probability) of the random variable \( X \). \( f(x|y) \) (Pr(\( x|y \))) denotes the probability density function (probability) of the random variable \( X \) conditioned on the event that \( Y = y \). \( E\{\cdot\} \) denotes the expectation operator. \( A^T, A^{-1} \) and \( |A| \) are the transpose, the inverse and the determinant of matrix \( A \), respectively. The term \( x^TAx \) for the symmetric matrix \( A \) and vector \( x \) is abbreviated as \( x^*A(*) \).

II. PROBLEM FORMULATION

In this paper, we are interested in the following linear dynamic system

\[
x_{k+1} = Ax_k + w_k,
\]
\( y_k = C x_k + v_k, \)

where \( x_k \in \mathbb{R}^n, y_k \in \mathbb{R}^m \) are the state and output; \( w_k \) and \( v_k \) are the process and measurement noises. We assume that \( \{w_k\}_{k \geq 0} \) and \( \{v_k\}_{k \geq 0} \) are white Gaussian processes with zero mean and covariance matrices \( Q \) and \( R \), respectively. Moreover, the initial system state satisfies \( x_0 \sim \mathcal{N}(0, \Sigma_0) \) and is independent with \( w_k \) and \( v_k \).

We consider the remote estimation problem where the sensor output is transmitted to the estimator through a wireless network. To reduce the communication frequency, after measuring \( y_k \), the sensor follows the stochastic sensor scheduling algorithm [7] to decide whether to transmit \( y_k \) to the estimator or not. Let \( s_k \) denote the decision variable by the sensor. When \( s_k = 1 \), the sensor transmits \( y_k \) to the estimator and \( s_k = 0 \), otherwise. We assume that the communication channel between the sensor and the estimator is a memoryless erasure channel. Let us define independent Bernoulli random variables \( \gamma_k \)'s such that \( \gamma_k = 1 \) if the channel is in the good state at time \( k \) and \( \gamma_k = 0 \) if otherwise. Hence, the estimator can only receive \( y_k \) when both \( s_k = 1 \) and \( \gamma_k = 1 \). We further assume that \( \Pr(\gamma_k = 0) = p \) and the estimator can distinguish whether a packet arrives or not. However, in case of no packet arrival, the estimator cannot decide whether it is due to channel loss or the inactivity of the event-trigger. The following information is available to the estimator at time \( k \)

\[ I_k = \{s_0 \gamma_0, \ldots, s_k \gamma_k, s_0 \gamma_0 y_0, \ldots, s_k \gamma_k y_k\} \quad (1) \]

with \( I_{k-1} = \emptyset \). The following notions are defined first and will be used in subsequent analysis.

\[ \hat{x}_{k|k} = \mathbb{E}\{x_k|I_k\}, \quad \hat{x}_{k|k-1} = \mathbb{E}\{x_k|I_{k-1}\}, \]
\[ \hat{y}_{k|k} = \mathbb{E}\{y_k|I_k\}, \quad \hat{y}_{k|k-1} = \mathbb{E}\{y_k|I_{k-1}\}, \]
\[ e_{k|k} = x_k - \hat{x}_{k|k}, \quad e_{k|k-1} = x_k - \hat{x}_{k|k-1}, \]
\[ P_{k|k} = \mathbb{E}\{e_{k|k} e_{k|k}^T\}, \quad P_{k|k-1} = \mathbb{E}\{e_{k|k-1} e_{k|k-1}^T\}. \]

The stochastic sensor scheduling algorithm [7] operates as below. At the time \( k \), the sensor randomly generates a variable \( \zeta_k \) from a uniform distribution on \([0, 1]\). Then \( \zeta_k \) is compared with a function \( \phi(y_k, \hat{y}_{k|k-1}) \), where \( \phi(y_k, \hat{y}_{k|k-1}) : \mathbb{R}^m \times \mathbb{R}^m \rightarrow [0, 1] \). The sensor schedules transmissions based on the following rule

\[ s_k = \begin{cases} 0, & \text{if } \zeta_k \leq \phi(y_k, \hat{y}_{k|k-1}), \\ 1, & \text{if } \zeta_k > \phi(y_k, \hat{y}_{k|k-1}). \end{cases} \quad (2) \]

Two stochastic sensor scheduling algorithms are proposed in [7]. The first one is the open-loop scheduler, in which \( \phi(y_k, \hat{y}_{k|k-1}) = e^{y_k^T Y y_k} \) with \( Y > 0 \). The open-loop scheduler uses the current measurement \( y_k \) only to schedule the transmission. Another scheduler is the closed-loop scheduler, in which \( \phi(y_k, \hat{y}_{k-1}) = e^{-Z_k^T Z_k} \) with \( Z > 0 \) and \( Z_k = y_k - \hat{y}_{k|k-1} \). The closed-loop scheduler relies on the feedback information \( \hat{y}_{k|k-1} \) to schedule the transmission. The diagram of the system is shown in Fig. 1.

In subsequent sections, we will show that in the presence of channel losses, the distribution of \( x_k \) conditioned on \( I_k \) is mixture Gaussian with an exponentially increasing number of components. Moreover, MMSE estimators are derived from the expectation of the mixture Gaussian distribution.

**Remark 1:** There are several reasons for not implementing the Kalman filter at the sensor side. The first reason is that the sensor might be primitive [7], so it does not have a sufficient computation capability to run a local Kalman filter. Secondly, the system parameters might not be available to the sensor. Thirdly, in decentralized settings where there are multiple sensors measuring the same process, only the fusion center which has access to all the sensor measurements can run the Kalman filter. In the end, the state dimension might be larger than the output dimension. Therefore, it reduces the communication cost to transmit the sensor output and perform the Kalman filter at the estimator side.

**Remark 2:** The closed-loop scheduler assumes perfect feedback channels and this is possible if the estimator has a much larger transmitting power than the sensor. A good example is the communication with a satellite. The power in the ground-to-satellite direction can be much larger than that in the reverse direction, so the first link can be considered as a (essentially) noiseless link.

### III. Optimal Open-Loop Estimator

In this section, we assume that the open-loop stochastic sensor scheduler is applied and try to derive the MMSE estimator. First of all, the following notions are defined. For any given \( i \in \mathbb{N} \) and \( k \in \mathbb{N} \), let

\[ i_k^- = \begin{cases} i, & \text{if } i < 2^k, \\ i - 2^k, & \text{if } i \geq 2^k. \end{cases} \]

For any given \( k \in \mathbb{N} \) and \( i \in \mathbb{N} \) with \( 0 \leq i \leq 2^{k+1} - 1 \), define the event

\[ \Gamma_i^k = \{\gamma_0, \ldots, \gamma_k = (b_0, \ldots, b_k)\}, \]

where \( b_k \) is the \((k+1)\)-th element of the binary expansion of \( i \), i.e., \( i = b_0 2^k + b_{k-1} 2^{k-1} + \ldots + b_0 2^0 \). Therefore, \( \Gamma_i^k \) denotes a packet drop sequence \( \{\gamma_0, \ldots, \gamma_k\} \) specified by the index \( i \). \( i_k^- \) is the index of the sub-sequence \( \{\gamma_0, \ldots, \gamma_{k-1}\} \) extracted from the sequence \( \{\gamma_0, \ldots, \gamma_k\} \) specified by the index \( i \).

We shall prove that the posterior distribution of \( x_k \) can be written as follows:

\[ f(x_k|I_k) = \sum_{i=0}^{2^{k+1}-1} f(x_k|\Gamma_i^k, I_k) \Pr(\Gamma_i^k|I_k), \quad (3) \]
One can interpret $\Gamma_k$ as a possible realization of the channel loss sequence up to time $k$, and $f(x_k|\Gamma_k, I_k)$ is the pdf of $x_k$ assuming that we know the channel loss sequence is indeed $\Gamma_k$, which we shall prove later is indeed Gaussian. Moreover, Pr($\Gamma_k|I_k$) represents the estimator’s estimate of the likelihood of channel loss sequence $\Gamma_k$ with the available information $I_k$. Therefore, $f(x_k|I_k)$ is the pdf of a Gaussian mixture. In the sequel, we will derive the expressions for $f(x_k|\Gamma_k, I_k)$ and Pr($\Gamma_k|I_k$). Then in view of (3) and (4), the optimal estimator can be obtained.

The following results are required and are presented first.

\[ f(x_k|\Gamma_k, I_k) = f(x_k|\Gamma_k, I_k) = f(x_k|\Gamma_k, I_k) \]  

since only knowing $\gamma_k$ without knowing $y_k$ cannot help to improve the knowledge about $x_k$.

**Lemma 3:**

\[ f(x_k|\Gamma_k, I_k) = N_x(m_i^k|k, P_i^k|k), \quad 0 \leq i \leq 2^{k+1} - 1 \]

where $m_i^k|k, P_i^k|k$ satisfy the following recursive equations.

**Time Update:**

\[ m_{i+1}|k = Am_i^k|k, P_{i+1}|k + AP_i^k|k A^T + Q \]

**Measurement Update:**

- For $i < 2^k$,
  \[ m_i^k|k = m_i^k|k, P_i^k|k = P_i^k|k \]
- For $i \geq 2^k$,
  \[ m_i^k|k = (I - K_i^k C)m_i^k|k - 1 + s_k \gamma_k K_i^k y_k \]
  \[ P_i^k|k = P_i^k|k - 1 - K_i^k P_i^k|k - 1 C_i^k \]

with initial conditions $m_0^0 = 0, P_0^0 = 1 = \Sigma_0$.

**Proof:** The proof of the initialization and the time-update is straightforward. The measurement update (6) follows from the fact that for $i < 2^k$, we have $\gamma_k = 0$. Therefore, no new information is available and the measurement update is not needed. The measurement updates (7), (8), (9) follow from the fact that for $i \geq 2^k$, we have $\gamma_k = 1$. Therefore, the measurement update is the same as [7]. It should be noted that in the case $i \geq 2^k, s_k = \gamma_k s_k$.

Next we calculate the probabilities of $\alpha_{k-1}^i = \Pr(\Gamma_k|I_k), \alpha_{k-1}^i = \Pr(\Gamma_i^1|I_k)$.

**Lemma 4:** $\alpha_{k-1}^i$ and $\alpha_{k-1}^i$ with $0 \leq i \leq 2^{k+1} - 1$ satisfy the following recursive equation.

**Time Update:**

\[ \alpha_{k-1}^i = p\alpha_{k-1}^{i-1} \]

- For $i < 2^k$,
  \[ \alpha_{k-1}^i = (1 - p)\alpha_{k-1}^{i-1} \]

**Measurement Update:**

\[ \alpha_{k-1}^i = \frac{\Pr(s_k \gamma_k|\Gamma_i^1, I_k)\alpha_{k-1}^i}{\sum_{j=0}^{2^{k+1} - 1} \Pr(s_k \gamma_k|\Gamma_j^1, I_k)\alpha_{k-1}^j} \]

where

- For $j < 2^k$,
  \[ \Pr(s_k \gamma_k|\Gamma_j^1, I_k) = 1 - s_k \gamma_k \]

- For $j \geq 2^k$,
  \[ \Pr(s_k \gamma_k|\Gamma_j^1, I_k) = 0 \]

Proof: Time Update: follows from the fact that for $i < 2^k$, we have $\gamma_k = 0$. Therefore

\[ \alpha_{k-1}^i = \Pr(\Gamma_i^1, \gamma_k = 0|I_k - 1) \]

\[ = \Pr(\gamma_k = 0) \Pr(\Gamma_i^1|I_k - 1) = 0 \]

Measuring Update: Since

\[ \alpha_{k-1}^i = \Pr(\Gamma_i^1|I_k) = \Pr(\Gamma_i^1|s_k \gamma_k, s_k \gamma_k y_k, I_k - 1) \]

\[ = \Pr(\Gamma_i^1|s_k \gamma_k, I_k - 1) \]

\[ = \frac{\Pr(s_k \gamma_k|\Gamma_i^1, I_k - 1)\Pr(\Gamma_i^1|I_k - 1)}{\Pr(s_k \gamma_k|I_k - 1)} \]

\[ = \frac{\sum_{j=0}^{2^{k+1} - 1} \Pr(s_k \gamma_k|\Gamma_j^1, I_k - 1)\alpha_{k-1}^j}{\sum_{j=0}^{2^{k+1} - 1} \Pr(s_k \gamma_k|\Gamma_i^1, I_k - 1)\alpha_{k-1}^j} \]

where the third equality follows from the fact that when $s_k \gamma_k = 0$, $s_k \gamma_k y_k = 0$, it is useless to know $s_k \gamma_k y_k$; when $s_k \gamma_k = 1$, knowing $s_k \gamma_k y_k$ is equivalent to know $y_k$, which is also useless in inferring $\Gamma_i$. Next we will show how to calculate $\Pr(s_k \gamma_k|\Gamma_i^1, I_k - 1)$. When $i < 2^k$, since $\gamma_k = 0$, we have that $s_k \gamma_k = 0$. Therefore

\[ \Pr(s_k \gamma_k|\Gamma_i^1, I_k - 1) = 1 - s_k \gamma_k \]
When $i \geq 2^k$, we have $\gamma_k = 1$. Let $M_k^{\hat{x}} = CP_{k+1}^{\hat{x}} C' + R$, we then have

$$\Pr(s_k|\gamma_k, \Gamma^i_k, \mathcal{I}_{k-1}) = \Pr(s_k|\Gamma^i_k, \mathcal{I}_{k-1})$$

$$= \int_{\mathbb{R}^m} \Pr(s_k|y_k, \Gamma^i_k, \mathcal{I}_{k-1}) f(y_k|\Gamma^i_k, \mathcal{I}_{k-1}) dy_k$$

$$= \int_{\mathbb{R}^m} \Pr(s_k|y_k) f(Cx_k + v_k|\Gamma^i_k, \mathcal{I}_{k-1}) dy_k$$

(a)\hspace{1cm} = \int_{\mathbb{R}^m} \left( s_k(1 - 2e^{-\frac{1}{2}y_k^T y_k}) + e^{-\frac{1}{2}y_k^T y_k} \right) \times f(Cx_k + v_k|\Gamma^i_k, \mathcal{I}_{k-1}) dy_k$$

$$= \int_{\mathbb{R}^m} \left( s_k(1 - 2e^{-\frac{1}{2}y_k^T y_k}) + e^{-\frac{1}{2}y_k^T y_k} \right) \times N_{y_k}(Cm_k^{\hat{x}}, |M_k^{\hat{x}}|) dy_k$$

$$= s_k + (1 - 2s_k) \int_{\mathbb{R}^m} e^{-\frac{1}{2}y_k^T y_k} N_{y_k}(Cm_k^{\hat{x}}, |M_k^{\hat{x}}|) dy_k$$

$$= s_k + \frac{1 - 2s_k}{\sqrt{(2\pi)^m|M_k^{\hat{x}}|}} \times e^{-\frac{1}{2}(Cm_k^{\hat{x}})^T (M_k^{\hat{x}})^{-1}Cm_k^{\hat{x}}} \times \int_{\mathbb{R}^m} e^{-\frac{1}{2}y_k^T \left( Y + (M_k^{\hat{x}})^{-1}\right) y_k + (Cm_k^{\hat{x}})^T (M_k^{\hat{x}})^{-1} y_k} dy_k$$

(b)\hspace{1cm} = s_k + \frac{1 - 2s_k}{\sqrt{(2\pi)^m|M_k^{\hat{x}}|}} \times \sqrt{\left| (M_k^{\hat{x}})^{-1} \right|} \times e^{\frac{1}{2}(Cm_k^{\hat{x}})^T (M_k^{\hat{x}})^{-1} \left( Y + (M_k^{\hat{x}})^{-1}\right) (M_k^{\hat{x}})^{-1} Cm_k^{\hat{x}}} \times \int_{\mathbb{R}^m} e^{-\frac{1}{2}(Cm_k^{\hat{x}})^T (M_k^{\hat{x}})^{-1} Y + (M_k^{\hat{x}})^{-1} y_k} dy_k$$

(13)

where (a) follows from (5); (b) follows from the Gaussian integral and (c) follows from the matrix inversion lemma.

In view of Lemma 3 and Lemma 5, the MMSE estimate can be calculated by the Gaussian sum filter [11] and is given as follows.

**Theorem 5:** With the open-loop scheduler and in the presence of packet drops, the MMSE estimator is given by

$$\hat{x}_{k|k} = \sum_{i=0}^{2^{k+1} - 1} \alpha_{k|k}^{i} m_{k|k}^{i},$$

$$P_{k|k} = \sum_{i=0}^{2^{k+1} - 1} \alpha_{k|k}^{i} \left( P_{k|k}^{i} + (m_{k|k}^{i} - \hat{x}_{k|k}) (m_{k|k}^{i} - \hat{x}_{k|k})^T \right),$$

$$\hat{x}_{k|k-1} = \sum_{i=0}^{2^{k-1} - 1} \alpha_{k-1|k-1}^{i} m_{k|k-1}^{i},$$

$$P_{k|k-1} = \sum_{i=0}^{2^{k-1} - 1} \alpha_{k-1|k-1}^{i} (P_{k|k-1}^{i} + (m_{k|k-1}^{i} - \hat{x}_{k|k-1}) (m_{k|k-1}^{i} - \hat{x}_{k|k-1})^T).$$

We can verify that when there is no packet drop, the optimal estimator degenerates to the one given in [7]. Besides, it is straightforward from the above expressions that the time update of the MMSE estimator can be written as

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k}, \quad P_{k+1|k} = AP_{k|k}A' + Q.$$  

However, there are no such simple relations for the measurement update of the optimal estimator.

**IV. OPTIMAL CLOSED-LOOP ESTIMATOR**

In this section, we derive the optimal estimators when the closed-loop scheduler is applied. Similar to the open-loop scheduler case, for closed-loop schedulers, we have the following result describing the pdf of $x_k$ conditioned on $\Gamma^i_k$ and $\mathcal{I}_k$. The proof is similar to that of Lemma 3 and is omitted for brevity.

**Lemma 6:**

$$f(x_k|\Gamma^i_k, \mathcal{I}_k) = \mathcal{N}_{x_k}(m_{k|k}^{i}, P_{k|k}^{i}), \quad 0 \leq i \leq 2^{k+1} - 1$$

$$f(x_k|\Gamma^i_k, \mathcal{I}_{k-1}) = \mathcal{N}_{x_k}(m_{k|k-1}^{i}, P_{k|k-1}^{i}), \quad 0 \leq i \leq 2^{k} - 1$$

where $m_{k|k}^{i}, P_{k|k}^{i}, m_{k|k-1}^{i}, P_{k|k-1}^{i}$ satisfy the following recursive equations.

**Time Update:**

$$m_{k|k}^{i} = Am_{k|k-1}^{i}, P_{k+1|k}^{i} = AP_{k|k}^{i}A' + Q.$$  

**Measurement Update:**

- For $i < 2^k$, $m_{k|k}^{i} = m_{k|k-1}^{i}, P_{k|k}^{i} = P_{k|k-1}^{i}.$
- For $i \geq 2^k$, $m_{k|k}^{i} = m_{k|k-1}^{i} + s_k \gamma_k K^i_k z_k,$ $P_{k|k}^{i} = P_{k|k-1}^{i} - K^i_k C P_{k|k-1}^{i} K^i_k.$

Next we will show how to calculate $\alpha_{k-1|k}^{i} = \Pr(\Gamma^i_k|\mathcal{I}_k), \alpha_{k|k}^{i} = \Pr(\Gamma^i_k|\mathcal{I}_k)$, for the closed-loop scheduler case.

**Lemma 7:** $\alpha_{k-1|k}^{i}$ and $\alpha_{k|k}^{i}$ with $0 \leq i \leq 2^{k+1} - 1$ can be calculated recursively as

**Time Update:**

- For $i < 2^k$, $\alpha_{k-1|k-1}^{i} = p\alpha_{k-1|k-1}^{i}$
- For $i \geq 2^k$, $\alpha_{k-1|k-1}^{i} = (1 - p)\alpha_{k-1|k-1}^{i}$

**Measurement Update:**

$$\alpha_{k|k}^{i} = \frac{\Pr(s_k \gamma_k|\Gamma^i_k, \mathcal{I}_{k-1}) \alpha_{k|k-1}^{i}}{\sum_{j=0}^{2^{k+1}-1} \Pr(s_k \gamma_k|\Gamma^j_k, \mathcal{I}_{k-1}) \alpha_{k|k-1}^{j}},$$

where
• For \( j < 2^k \),
  \[
  \Pr(s_k\gamma_{|k}\Gamma_k\hat I_{k-1}) = 1 - s_k\gamma_k
  \]

• For \( j \geq 2^k \),
  \[
  \Pr(s_k\gamma_{|k}\Gamma_k\hat I_{k-1}) = s_k\gamma_k + \frac{1 - 2s_k\gamma_k}{\sqrt{(CP_k^{\hat I_{k-1}}C' + R)Z + I}}
  \]
\[
\times e^{-\frac{1}{2}(C(m_k^{\hat I_{k-1}} - \hat x_k)|\hat I_{k-1}))'(Z^{-1} + (CP_k^{\hat I_{k-1}}C' + R))^{-1}(s)}
\]

with initial conditions \( \alpha_0^0 = p, \alpha_1^0 = 1 - p \).

**Proof:** The proof of the time update is the same as that of Lemma [4]. The measurement update only differs in the calculation of \( \Pr(s_k\gamma_{|k}\Gamma_k\hat I_{k-1}) \) for \( j \geq 2^k \) which is demonstrated as follows. For \( j \geq 2^k \), let \( M_k^{\hat I_{k-1}} = CP_k^{\hat I_{k-1}}C' + R \), we have

\[
\Pr(s_k\gamma_{|k}\Gamma_k\hat I_{k-1}) = \Pr(s_k\gamma_{|k}\Gamma_k\hat I_{k-1})
\]

\[
= \int_{\R_m} \Pr(s_kz_k\Gamma_k\hat I_{k-1})f(x_k|\Gamma_k\hat I_{k-1})dz_k
\]

\[
= \int_{\R_m} \Pr(s_kz_k)f(C(x_k - \hat x_k)|\Gamma_k\hat I_{k-1})dz_k
\]

\[
= \int_{\R_m} (s_k(1 - 2e^{-\frac{1}{2}z_kZz_k}) + e^{-\frac{1}{2}z_kZz_k})
\]
\[
\times f(C(x_k - \hat x_k)|\Gamma_k\hat I_{k-1})dz_k
\]

\[
= \int_{\R_m} (s_k(1 - 2e^{-\frac{1}{2}z_kZz_k}) + e^{-\frac{1}{2}z_kZz_k})
\]
\[
\times N_z(C(m_k^{\hat I_{k-1}} - \hat x_k)|\Gamma_k\hat I_{k-1})C' + R)dz_k
\]

\[
= s_k + \frac{1 - 2s_k}{\sqrt{(2\pi)^{m|CP_{\hat I_{k-1}}C' + R|Z + I}}}
\]
\[
\times e^{-\frac{1}{2}(C(m_k^{\hat I_{k-1}} - \hat x_k)|\hat I_{k-1})'(Z^{-1} + (CP_k^{\hat I_{k-1}}C' + R))^{-1}(s)}
\]

where the last equality can be proved similarly as [13].

In view of Lemma [6] and Lemma [7] we have the following theorem.

**Theorem 8:** With the closed-loop scheduler and in the presence of packet drops, the MMSE estimator is given by

\[
\hat x_k = \sum_{i=0}^{2^{k+1}-1} \alpha_{i|k}^i m_{i|k}^i,
\]
\[
P_k = \sum_{i=0}^{2^{k+1}-1} \alpha_{i|k}^i \left( P_{i|k}^i + (m_{i|k}^i - \hat x_{i|k})(m_{i|k}^i - \hat x_{i|k})' \right),
\]
\[
\hat x_{k-1} = \sum_{i=0}^{2^k-1} \alpha_{i|k-1}^i m_{i|k-1}^i,
\]
\[
P_{k|k-1} = \sum_{i=0}^{2^k-1} \alpha_{i|k-1}^i P_{i|k-1}^i + \sum_{i=0}^{2^k-1} \alpha_{i|k-1}^i (m_{i|k-1}^i - \hat x_{i|k-1})(m_{i|k-1}^i - \hat x_{i|k-1})',
\]

**V. Reduced Computational Complexities**

The problem considered in this paper is similar to the state estimation problem of Markov jump systems with unknown jump modes, where it is shown that the optimal nonlinear filter is obtained from a bank of Kalman filters, which requires exponentially increasing memory and computation with time [12]. The generalized pseudo Bayes (GPB) algorithm [13] and the interacting multiple model (IMM) algorithm [14] are two commonly used sub-optimal algorithms to overcome the computational complexities. The approximations in the GPB algorithm consist of restricting the probability density \( f(x_k|I_{k-1}) \) to depend on at most the last \( N \) random variables \( \gamma_k, \ldots, \gamma_{k-N+1} \) and approximate each hypothesis \( f(x_k|\gamma_k, \ldots, \gamma_{k-N+1}, I_{k-1}) \) with a Gaussian distribution. Moreover, a hypothesis merging operation is introduced at every step to prevent the increase of hypothesis numbers with time. The suboptimum procedure approaches the optimum one with increasing \( N \). The number \( N \) is to be chosen on the basis of the desired estimation performance subject to the constraint of the allowable storage capacity. The IMM estimator further exploits the timing of hypothesis merging to reduce the computational complexity. Owing to its excellent estimation performance and low computational cost, the IMM algorithm has been widely applied in various fields [15]. The principles of the GPB and the IMM algorithms can be utilized to derive sub-optimal estimators for the problem considered in this paper. The derivations are straightforward following [13] and [16] and are omitted here.

Another simple strategy is to directly apply the Gaussian mixture reduction algorithms [17] at each step of the measurement update to reduce the numbers of components in the Gaussian mixture model. This strategy has been applied to the Bayes filtering problem to reduce the computational complexities [18]. However, it should be noted that the performance of all the above mentioned algorithms can only be evaluated via Monte Carlo simulations and there are no systematic methods to analyze the performance.

**VI. Simulations**

In simulations, we adopt the same system parameters as in [7], which are

\[
A = \begin{bmatrix} 0.8 & 0.95 \\ 0.95 & 1 \end{bmatrix}, C = [1, 1], \Sigma_0 = Q = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, R = 1.
\]

We only conduct simulations under the open loop scheduler setting with the optimal estimator, in conjunction with the oracle estimator, the GPB estimator and the OLSET-KF estimator in [7], where no packet drops are considered. The OLSET-KF estimator does not consider packet drops. When the estimator fails to receive a packet, it always assumes that this is caused by the hold of transmission from the scheduler. The oracle estimator is the optimal estimator under the assumption that the estimator knows the value of \( \gamma_k \).
at each step. Therefore, if the estimator fails to receive a packet, it knows whether this is caused by the channel or the scheduler. If this is caused by the channel, i.e., $\gamma_k = 0$, no measurement update is conducted and vice versa. Clearly, the oracle estimator has the smallest mean square error (MSE) and can be used as a benchmark to evaluate the performance of other estimators.

In simulations, the schedule parameter $Y$ is selected as $Y = 1$ and $N = 2$ is selected for the GPB estimator. We compare the performance of different estimators and we adopt Monte Carlo methods with 1000 independent experiments to evaluate the sum of MSE $\sum_{k=0}^9 E\{\|x_k - \hat{x}_{k|k}\|^2\}$ under different packet drop rates. The simulation results are illustrated in Fig. 2 where the relative sum of MSE is plotted. The relative sum of MSE is defined as the sum of MSE of an estimator divided by the sum of MSE of the oracle estimator. It is clear from Fig. 2 that the sum of MSE of the GPB estimator is close to that of the optimal estimator and is much smaller than the OLSET-KF, which shows the superior performance of the GPB estimator and also indicates the advantage of considering packet drops in the remote state estimation problem. Moreover, in the case of $p = 0$ and $p = 1$, the sum of MSE of the OLSET-KF, the optimal estimator, the GPB estimator and the oracle estimator are equal. This is because when $p = 0$ ($p = 1$), the optimal estimator (GPB estimator) assigns zero probability to all the hypotheses with $\gamma_k = 0$ ($\gamma_k = 1$). Therefore, only the hypothesis with $\gamma_k = 1$ ($\gamma_k = 0$) for all $k$ is preserved. As a result, the estimate of the optimal estimator (GPB estimator) is the same with the oracle estimator. Therefore, for the case that $p = 0$ ($p = 1$), the optimal estimator (GPB estimator) and the oracle estimator have the same sum of MSE. The recursions of the OLSET-KF and the oracle estimator are the same for the case $p = 0$, where there are no packet drops. For the case that $p = 1$, even though the recursions of the OLSET-KF and the oracle estimator are different, since they both start with $\hat{x}_{0|0} = 0$, their estimates would always be $\hat{x}_{k|k} = 0$. Therefore, for the case that $p = 0$ and $p = 1$, the OLSET-KF and the oracle estimator have the same sum of MSE.

VII. CONCLUSIONS

This paper studies the remote state estimation problem of linear systems with stochastic event-triggered sensor schedulers in the presence of packet drops. The conditional PDFs are computed, the optimal estimators are derived and the corresponding communication rates are analyzed. Strategies to reduce the computational complexities are discussed. However, the performance of sub-optimal estimators can only be evaluated via simulations. Sub-optimal estimators with performance guarantees are to be proposed.

REFERENCES

[1] C. Yang and L. Shi, “Deterministic sensor data scheduling under limited communication resource,” IEEE Transactions on Signal Processing, vol. 59, no. 10, pp. 5050–5056, 2011.
[2] L. Shi, P. Cheng, and J. Chen, “Sensor data scheduling for optimal state estimation with communication energy constraint,” Automatica, vol. 47, no. 8, pp. 1693–1698, 2011.
[3] Y. Mo, E. Garone, and B. Sinopoli, “On infinite-horizon sensor scheduling,” Systems & control letters, vol. 67, pp. 65–70, 2014.
[4] M. Xia, V. Gupta, and P. J. Antsaklis, “Networked state estimation over a shared communication medium,” IEEE Transactions on Automatic Control, vol. 62, no. 4, pp. 1729–1741, 2017.
[5] S. Trumpe, “Stability analysis of distributed event-based state estimation,” in Proceedings of the 53rd Annual Conference on Decision and Control, (Los Angeles, California, USA), pp. 2013–2019, 2014.
[6] J. Wu, Q. Jia, K. H. Johansson, and L. Shi, “Event-based sensor data scheduling: Trade-off between communication rate and estimation quality,” IEEE Transactions on Automatic Control, vol. 58, no. 4, pp. 1041–1046, 2013.
[7] D. Han, Y. Mo, J. Wu, S. Weerakkody, B. Sinopoli, and L. Shi, “Stochastic event-triggered sensor schedule for remote state estimation,” IEEE Transactions on Automatic Control, vol. 60, no. 10, pp. 2661–2675, 2015.
[8] S. Trumpe and R. D’Andrea, “Event-based state estimation with variance-based triggering,” IEEE Transactions on Automatic Control, vol. 59, no. 12, pp. 3266–3281, 2014.
[9] A. S. Leong, S. Dey, and D. E. Quevedo, “Sensor scheduling in variance based event triggered estimation with packet drops,” IEEE Transactions on Automatic Control, vol. 62, no. 4, pp. 1880–1895, 2017.
[10] E. Kang, J. Wu, D. Shi, and L. Shi, “On the existence of event-based triggers that preserve gaussian state in presence of package-drop,” in Proceedings of the 2017 American Control Conference, (Seattle, USA), pp. 1233–1237, 2017.
[11] B. D. Anderson and J. B. Moore, Optimal filtering. New Jersey: Prentice-Hall, Inc., 1979.
[12] O. L. d. V. Costa, M. D. Fragoso, and R. P. Marques, Discrete-time Markov jump linear systems. Probability and its applications, London: Springer, 2005.
[13] A. G. Jaffer and S. C. Gupta, “On estimation of discrete processes under multiplicative and additive noise conditions,” Information Sciences, vol. 3, no. 3, pp. 267–276, 1971.
[14] H. A. Blom and Y. Bar-Shalom, “The interacting multiple model algorithm for systems with markovian switching coefficients,” IEEE Transactions on Automatic Control, vol. 33, no. 8, pp. 780–783, 1988.
[15] E. Mazor, A. Averbuch, Y. Bar-Shalom, and J. Dayan, “Interacting multiple model methods in target tracking: a survey,” IEEE Transactions on aerospace and electronic systems, vol. 34, no. 1, pp. 103–123, 1998.
[16] C. E. Seah and I. Hwang, “Algorithm for performance analysis of the imm algorithm,” IEEE Transactions on Aerospace and Electronic Systems, vol. 47, no. 2, pp. 1114–1124, 2011.
[17] D. F. Crouse, P. Willett, K. Pattipati, and L. Svensson, “A look at gaussian mixture reduction algorithms,” in 14th International Conference on Information Fusion, (Chicago, IL, USA), pp. 1–8, 2011.
[18] A. G. Wills, J. Hendriks, C. Renton, and B. Ninness, “A bayesian filtering algorithm for gaussian mixture models,” arXiv preprint arXiv:1705.05495, 2017.