

Event-based State Estimation in Cyber-Physical Systems

by

Dawei Shi

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Control Systems

Department of Electrical and Computer Engineering
University of Alberta

©Dawei Shi, 2014

Abstract

This thesis focuses on event-based state estimation problems in the context of cyber-physical systems (CPSs), targeting at low-complexity event-based state estimators that are optimal in a certain sense. The motivation stems from the resource limitations in the applications of CPSs (e.g., wireless sensor networks) as well as the increased computation burden in calculating the optimal state estimates caused by the event-triggering conditions.

Several event-based estimation problems are formulated and solved using different approaches, including the maximum likelihood estimation approach, the approximate Gaussian filtering approach, the set-valued Kalman filtering approach and the change of probability measure approach. For all these investigations, optimal state estimates with simple structures that can be recursively calculated are obtained, which form the major contributions of this thesis. Also, the performance improvements in the sense of average estimation errors by exploiting the information contained in the event-triggering conditions are addressed either by theoretical proofs or extensive numerical simulations. Several results on communication rate analysis are proposed, which are relevant and necessary for event-based estimation, considering the potential communication resource limitations in CPSs.

Based on the developed results, the outcome of the research attempts on event-based estimation is encouraging, and a distinct and systematic approach to event-based estimation seems on the horizon. The results are not only of theoretical value, but are potentially implementable to a variety of applications in industrial processes,

due to the practical considerations in both the problem formulations and the design procedures.

Preface

Chapters 2-4 of this thesis form parts of the international research collaborations with Dr. Ling Shi at the Hong Kong University of Science and Technology, and Chapter 5 is part of the international research collaboration with Prof. Robert J. Elliott at the University of Adelaide, Australia. The ideas, technical derivations and simulation examples are my original work, as well as the introduction and literature review in Chapter 1.

Parts of Chapter 2 of this thesis have been separately published as

- D. Shi, T. Chen and L. Shi, “Event-triggered maximum likelihood state estimation,” *Automatica*, 50(1), pp. 247-254, 2014.
- D. Shi, T. Chen and L. Shi, “Event-based state estimation of linear dynamical systems: Communication rate analysis,” in *Proceedings of the 2014 American Control Conference (ACC 2014)*, Portland, OR, 2014.

The material in Chapter 3 has been published as

- D. Shi, T. Chen and L. Shi, “An event-triggered approach to state estimation with multiple point- and set-valued measurements,” *Automatica*, 50(6), pp. 1641-1648, 2014.

Parts of the results in Chapter 4 is conditionally accepted for publication as

- D. Shi, T. Chen and L. Shi, “On set-valued Kalman filtering and its application

to event-based state estimation,” conditionally accepted by *IEEE Transactions on Automatic Control* as a full paper, 2014.

In all these papers, the ideas and technical results were developed by myself. Dr. Chen was the supervisory author and provided a lot of valuable suggestions in the development. Dr. Shi provided comments on the technical derivations and contributed to manuscript edits.

Parts of the results in Chapter 5 are under review as

- D. Shi, R. J. Elliott and T. Chen, “Event-based state estimation of discrete-state hidden Markov models,” submitted to *IEEE Transactions on Automatic Control*, 2014.

The problem was formulated and solved mostly on my own. Prof. Elliott provided a number of insightful comments and suggestions on alternative methods to validate the results. Dr. Chen was the supervisory author and provided suggestions on the results and the manuscript composition.

Acknowledgements

Firstly, I would like to express my most sincere appreciation to my supervisor, Dr. Tongwen Chen, for all his time, effort and contributions to my PhD research. I still remember when I came here four years ago, I was only a student with a control engineering background, knowing almost nothing about control theory. It was Dr. Chen who patiently taught me every single piece in theoretical research, from formulating a problem to presenting a theorem and to writing a scientific paper, from telling the distinctions between a lemma and a proposition to writing a response to review letter and to the style of giving a conference presentation. Without his help, there would be no chance for me to become a theoretical researcher.

Secondly, I would like to thank my previous supervisor, Prof. Junzheng Wang at the School of Automation, Beijing Institute of Technology for his persistent encouragement and support during my PhD studies. Around five years ago, Prof. Wang asked me to visit his office, and suggested the opportunity of pursuing a PhD degree in an overseas university, which was the starting point of my journey to Canada.

My special gratitude also goes to my academic collaborators, Dr. Ling Shi from the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, and Prof. Robert J. Elliott from the School of Mathematics, University of Adelaide, for their altruistic help and suggestions on the results we developed in our research collaboration. It was a great honor and an invaluable experience to work with these excellent and enthusiastic researchers during my PhD program.

I would also like to thank all the members in the Advanced Control Group and the nice staffs in the Department of Electrical and Computer Engineering, University of Alberta, for their kind help in the past four years.

Finally, I would like to thank my parents, for their understanding, encouragement

and support during all these years. As the only child in the family, I owe them really a lot. To them I dedicate my dissertation work. Better late than never.

Financial support from the China Scholarship Council and the Natural Sciences and Engineering Research Council of Canada is also gratefully acknowledged.

Contents

1	Introduction	1
1.1	Cyber-physical systems and event-based control and estimation	1
1.2	Literature review of event-based estimation	2
1.3	Summary of the contributions	4
2	Event-triggered maximum likelihood state estimation	6
2.1	Introduction	6
2.2	Problem formulation	7
2.3	Solution to the event-based ML estimation problem	10
2.4	One-step event-based ML state estimation	14
2.4.1	Solution to the problem	14
2.4.2	Analysis of the communication rate	16
2.5	Examples	24
2.5.1	Example 1	25
2.5.2	Example 2: Sensorless event-based estimation of a DC motor system	26
2.6	Summary	30
3	An approximate Gaussian approach to event-triggered state estimation	34
3.1	Introduction	34
3.2	System description and problem setup	35
3.3	Optimal fusion of sequential event-triggered measurement information	40
3.4	Experimental verification of the proposed results based on Monte Carlo simulations	46
3.5	Summary	48

4	On set-valued Kalman filtering and its application to event-based state estimation	51
4.1	Introduction	51
4.2	Problem setup	54
4.3	Sensor fusion	58
4.4	Asymptotic properties of the set of means of the estimates	63
4.5	Performance improvement	68
4.6	Application to event-based state estimation	73
4.6.1	Analysis and parameter design	73
4.6.2	Examples	76
4.7	Summary	80
5	Event-based state estimation of discrete-state hidden Markov models	83
5.1	Introduction	83
5.2	Problem Description	84
5.3	Results on reliable communication channel	88
5.3.1	Results on recursive estimation	90
5.3.2	Communication rate analysis	93
5.3.3	Estimation with a lossy communication channel: a special case	94
5.4	Results on unreliable communication channel	95
5.4.1	Results on recursive estimation	98
5.4.2	Communication rate analysis	101
5.5	Numerical implementation and verification	102
5.5.1	Numerical implementation	102
5.5.2	A numerical example	104
5.6	Summary	107
6	Conclusions and Future Work	109
6.1	Concluding remarks	109
6.2	Future work	110
6.2.1	Application of the discrete state HMM results to general dynamical systems	110
6.2.2	Event-based estimation problems for HMMs with continuous state spaces	110
6.2.3	The effect of time delays in communication protocols	110
6.2.4	Periodic local estimation vs. event-based remote estimation	111

List of Tables

2.1	Motor parameters	29
-----	----------------------------	----

List of Figures

2.1	Block diagram of the overall system.	8
2.2	Relationship of inner and outer ellipsoidal approximations of Ω for the case of $m = 2$	18
2.3	Relationship of inner and outer ellipsoidal approximations of Ω for the case of $m = 2$	22
2.4	Performance of the state estimation strategy (x_k^i denotes the trajectory of the i th state at time instant k , and \hat{x}_k^i denotes the estimate of the i th state at time instant k).	26
2.5	Plot of γ_k	27
2.6	Bounds on $\mathbf{E}(\gamma_k)$ (UB1 and LB1 respectively denote the upper and lower bounds derived in Theorem 2.3, UB2 and LB2 respectively denote the upper and lower bounds derived in Theorem 2.4, and UB3 and LB3 respectively denote the upper and lower bounds derived in Corollary 2.1).	28
2.7	Comparison of the relative tightness of different bounds (UB1-LB1 denotes the distance between the upper and lower bounds derived in Theorem 2.3, and UB2-LB2 denotes the distance between the upper and lower bounds derived in Theorem 2.4).	29
2.8	Tradeoff between estimation performance and communication rate.	30
2.9	Plot of the input signals.	31
2.10	Performance of the state estimation strategy.	32
2.11	Plot of $\mathbf{E}(\gamma_k)$	32
2.12	Tradeoff between estimation performance and communication rate.	33
3.1	Plot of the conditional distributions.	39
3.2	Performance validation of the proposed event-based estimator.	49
3.3	Performance comparison between different fusion sequences.	50

4.1	Comparison of the sizes of the sets of estimation means for different choices of Y_k^2	69
4.2	Multiple-sensor event-based remote estimation architecture.	74
4.3	Performance comparison of the different estimation strategy for $\delta = 0.5$ (the sets of estimation means are calculated by projecting the two-dimensional ellipsoids on one dimension).	78
4.4	Performance comparison of the different estimation strategy for $\delta = 1.5$ (the sets of estimation means are calculated by projecting the two-dimensional ellipsoids on one dimension).	79
4.5	Performance of the set-valued state estimation strategy (the envelopes are calculated by projecting the three-dimensional ellipsoids on one dimension).	81
4.6	Plot of sensor transmissions.	82
5.1	Block diagram of the event-based remote estimation system with a reliable communication channel.	86
5.2	Block diagram of the periodic estimation problem with a lossy communication channel.	94
5.3	Block diagram of the event-based remote estimation system with an unreliable communication channel.	95
5.4	Estimation performance comparison for the reliable communication channel case.	105
5.5	Tradeoff between average communication rate and estimation performance for the reliable communication channel case.	106
5.6	Estimation performance comparison for the unreliable communication channel case.	107
5.7	Tradeoff between average communication rate and estimation performance for the unreliable communication channel case.	108

List of Abbreviations

CPS	Cyber-Physical System
HMM	Hidden Markov Model
ML	Maximum Likelihood
MMSE	Minimum Mean Square Error
QP	Quadratic Programming
WSN	Wireless Sensor Network

Chapter 1

Introduction

1.1 Cyber-physical systems and event-based control and estimation

With the overwhelming developments and advances in wireless communication, micro-electromechanical systems and digital electronics, cyber-physical systems (CPSs) now play a central role to the technological advancements in our life, from unmanned automobiles to humanoid robots to smart grids, from GPS navigation to face recognition to speech recognition, and from remote control to building automation and to wireless systems [18].

According to US National Science Foundation [19], “CPSs are engineered systems that are built from, and depend upon, the seamless integration of computational algorithms and physical components”. One of the major problems that faces the applications of this type of systems, however, is the limitation of communication and power resources. For instance, in wireless monitoring, the wireless sensors/actuators are usually powered by batteries, some of which are not even replaceable [1]. Thus it is important to maintain a reasonable battery life in wireless applications [31], e.g., power saving mechanisms are included in ZigBee RF4CE standard [75]. Moreover, according to IEEE 802.15.4 [31], the number of communication channels are normally limited, which means only one or a subset of the sensors/actuators can transmit measurement information or perform their control actions at each time instant.

In this context, the emerging event-based control and estimation strategies become attractive to applications in CPSs, especially in wireless sensor networks (WSNs) as well as systems with limited communication or computational resources, due to their capability to maintain the system performance at reduced communication or computation rates. The basic rationale of the event-based action strategies is that the

sensors and actuators do not update their actions until certain events happen (e.g., the difference between the current measurement and the measurement at the previous event time goes beyond a pre-specified level). The event-based control and estimation problems are potentially related with a number of other topics in the literature, e.g., nonuniformly sampled-data systems [28, 49], set-membership identification and estimation [3, 14, 36], quantized control systems [22, 40], just to name a few. However, the presence of the event-triggers has introduced new and distinct challenges to controller/estimator design, and performance guarantees (e.g., stability, optimality) are more difficult to be determined theoretically compared with their periodic counterparts, which forms the basic motivation of our investigation in this thesis.

1.2 Literature review of event-based estimation

The scope of this thesis is concerned with the topic of event-based estimation for discrete-time systems. This topic has received considerable attention during the last few years, since the pioneer work of Åström and Bernhardsson on Lebesgue sampling [9].

Earlier results mainly focused on the design of event-triggering strategies. The optimal event-based finite-horizon sensor transmission scheduling problems were studied in [32, 55] for continuous-time and discrete-time scalar linear systems, respectively. The results were extended to vector linear systems in [39] by relaxing the zero mean initial conditions and considering measurement noises. The tradeoff between performance and the average sampling period was analyzed in [38], and a sub-optimal event-triggering scheme with a guaranteed least average sampling period was proposed. Adaptive sampling for state estimation of continuous-time linear systems was considered in [54]. The authors in [61] proposed a hybrid sensor data scheduling method by combining time and event-based methods with reduced computational complexity. In [72], a distributed event-triggered estimation problem was considered and a global event-triggered communication policy for state estimation was proposed by minimizing a weighted function of network energy consumption and communication cost while considering estimation performance constraints. The joint design of event-trigger and estimator for first-order stochastic systems with arbitrary distributions was considered in [46], where a game-theoretic framework was utilized to analyze the optimal trade-off between the mean squared estimation error and the expected transmission rate. These results form the first line of research in event-based estimation.

In addition to the scheduling issues, another important problem is to find the optimal estimate for a specified event-triggering scheme. Although the sensors measure the state of the system at every sampling instant, they only send their measurements to the estimator when certain pre-specified conditions are satisfied (e.g., the current measurement value deviates the previous transmitted value by a specified level [44]). As a result, the estimator has to handle combined point- and set-valued hybrid measurement information at each time instant: when the measurement is received, the estimator updates itself with the “point-valued” information; when the measurement is not received, the estimator still knows the information that the event-triggering conditions are satisfied by the current measurement (which we refer to as “set-valued” information). The information pattern becomes even more complicated if the effect of packet dropout is further taken into account, which is necessary when the measurements are transmitted through wired or wireless communication networks. The difficulty caused by the hybrid measurement information mainly affects the estimation problems from the computational aspect, as in many cases the conditional distributions of the states no longer have simple closed-form expressions, and need to be calculated by numerically evaluating multi-dimensional integrals. This not only adds to the computational burden but also inevitably leads to accumulative numerical errors.

To solve this problem, a number of interesting attempts have been made in the literature, mostly for linear Gaussian systems, which form the second line of research in event-based estimation. Based on Gaussian assumptions on the conditional distributions of the states on the available hybrid measurement information, the minimum mean square error (MMSE) estimators were derived for an event-triggering scheme quantifying the magnitude of the innovation of the estimator in [73]. By approximating the uniform distribution with the sum of a finite number of Gaussian distributions, an event-based estimator with a hybrid update was proposed in [62], based on a general description of event-based sampling. In [26], a class of stochastic event-triggering conditions were proposed, and closed-form MMSE estimates were obtained without introducing additional approximations. In [63], an event-based estimate was obtained by minimizing the worst-case mean squared error, treating the noises and event-triggering conditions as stochastic and non-stochastic uncertainties, respectively. However, the developments obtained so far are still at a relatively early stage, and systematic approaches to handling the hybrid set- and point-valued measurement information are still missing in the literature.

1.3 Summary of the contributions

The results presented in this thesis target at finding the optimal estimators for given event-triggering schemes, which belong to the second line of research introduced in the previous section. The event-based estimation problem is approached from a few distinct perspectives, resulting in different event-based estimates that are optimal in a certain sense, which form the major contents and contributions of this thesis.

In Chapter 2, the event-triggered state estimation problem for linear time-invariant systems is considered in the framework of maximum likelihood (ML) estimation. We show that the optimal estimate is parameterized by a special time-varying Riccati equation, and the computational complexity increases exponentially with respect to the time horizon. For ease in implementation, a one-step event-based ML estimation problem is further formulated and solved, and the solution behaves like a Kalman filter with intermittent observations. For the one-step problem, time-varying upper and lower bounds on the expectation of the communication rate are provided. For stable systems, time-invariant upper and lower bounds are derived as well. For scalar sensors (namely, sensors with one measurement channel), the exact expression for the expectation of the communication rate is obtained. Numerical examples are presented and the benefits of the proposed one-step event-based estimator are demonstrated by comparative simulations.

In Chapter 3, we consider state estimation based on information from multiple sensors that provide their measurement updates according to separate event-triggering conditions. An optimal sensor fusion problem based on hybrid measurement information is formulated and explored. We show that under a commonly-accepted Gaussian assumption, the optimal estimator depends on the conditional mean and covariance of the measurement innovations, which applies to general event-triggering schemes. For the case that each channel of the sensors has its own event-triggering condition, closed-form representations are derived for the optimal estimate and the corresponding error covariance matrix, and it is proved that the exploration of the set-valued information provided by the event-triggering sets guarantees improvement of estimation performance. The effectiveness of the proposed event-based estimator is demonstrated by extensive Monte Carlo simulation experiments for different categories of systems and comparative simulation with the classical Kalman filter.

In Chapter 4, the properties of the exact and approximate set-valued Kalman filters with multiple sensor measurements for linear time-invariant systems are investigated. First, we show that the exact and the proposed approximate set-valued filters are in-

dependent of the fusion sequence at each time instant. Second, the boundedness of the size of the set of estimation means is proved for the exact set-valued filter. For the approximate set-valued filter, if the closed-loop matrix is contractive, then the set of estimation means has a bounded size asymptotically; otherwise a nonsingular linear transform is constructed such that the size of the set of estimation means for the transformed states is asymptotically bounded. Third, the effect of set-valued measurements on the size of the set of estimation means is analyzed and conditions for performance improvement in terms of smaller size of the set of estimation means are proposed. Finally, the results are applied to event-based estimation, which allow the event-triggering conditions to be designed by considering requirements on performance and communication rates. The efficiency of the proposed results is illustrated by examples and comparative simulation.

In Chapter 5, the state estimation problem for Hidden Markov Models (HMMs) subject to event-based sensor measurement updates is considered, using the change of probability approach. We assume the measurement updates are transmitted through wired or wireless communication networks. For the scenarios of reliable and unreliable communication channels, analytical expressions for the probability distributions of the states conditioned on all the past measurement information are obtained. Also, we show that the scenario of a lossy channel, but without the event-trigger, can be treated as a special case of the reliable channel results. Based on these results, closed-form expressions for the estimated communication rates are presented, which are shown to be the ratio between a weighted 1-norm and the 1-norm of the conditional probability distributions of the states. Implementation issues are discussed, and the effectiveness of the results is illustrated by a numerical example and comparative simulations.

Chapter 2

Event-triggered maximum likelihood state estimation*

2.1 Introduction

In this chapter, the event-based state estimation problem is considered under the maximum-likelihood estimation framework. We study the remote state estimation of a process based on the measurements taken by a battery-powered smart-sensor on the process side, the output of which is transmitted to the remote estimator through a wireless channel. Comparison between standard ZigBee chips designed according to [31] (e.g., CC2530 by [71]) and analog to digital converters (e.g., AD7988, 16-digit ADC from [6]) indicates that the energy consumption of wireless transmission is at least one magnitude greater than that of data acquisition and basic calculation. Consequently, we assume wireless transmission consumes more energy than basic calculation, and thus an event-based data-scheduler is proposed on the process side to prolong the battery life (utilizing the limited calculation capacity of the smart sensor). The main contributions of the work in this chapter are three-fold:

1. The structure of the event-based ML state estimator is provided. We show that the optimal estimate is parameterized by a special time-varying Riccati equation, and the computational complexity increases exponentially with the time horizon. Note that the solution to the Riccati equation is not necessarily the covariance matrix of the estimation error for event-based ML state estimation problems, due to event-based data updating.

*Parts of the results in this chapter appeared in *Automatica*, vol. 50, no. 1, pp. 247-254, 2014.

2. For ease in implementation of the event-based ML estimator, a one-step event-based ML estimation problem is formulated, and its solution is shown to behave like the Kalman filter with intermittent observations [64] and only requires feedback communication when an event occurs at the smart sensor. This is different from the results in [73], where feedback communication is always needed.
3. For the one-step event-based ML estimator, communication rate analysis is performed from the process side. Two sets of time-varying upper and lower bounds on the expectation of the communication rate are proposed in terms of incomplete Gamma functions, which can be iteratively calculated. For stable systems, time-invariant upper and lower bounds are provided as well. For the case of single-channel sensors, the exact expression of the expectation of communication rate is obtained.

Notation: \mathbb{N} and \mathbb{N}^+ denote the sets of nonnegative and positive integers, respectively. For $a, b \in \mathbb{N}$ and $a \leq b$, $u_{a:b}$ denotes $\{u(a), u(a+1), \dots, u(b)\}$. \mathbb{R} denotes the set of real numbers. For $m, n \in \mathbb{N}^+$, $\mathbb{R}^{m \times n}$ denotes the set of m by n real-valued matrices, whereas \mathbb{R}^m is short for $\mathbb{R}^{m \times 1}$. For $Z \in \mathbb{R}^{m \times n}$, Z^\top denotes the transpose of Z , whereas $Z^{-\top}$ denotes $(Z^\top)^{-1}$ if Z is square and nonsingular. For a random variable x , $\mathbf{E}(x)$ denotes its expectation, and x denotes its realization.

2.2 Problem formulation

Consider the system in Fig. 2.1. The process is Linear Time-Invariant (LTI) and evolves in discrete time driven by white noise:

$$\mathbf{x}_{k+1} = A\mathbf{x}_k + \mathbf{w}_k, \quad (2.1)$$

where $\mathbf{x}_k \in \mathbb{R}^n$ is the state, and $\mathbf{w}_k \in \mathbb{R}^n$ is the noise input, which is zero-mean Gaussian with covariance $Q > 0$.

The initial state \mathbf{x}_0 is Gaussian with $\mathbf{E}(\mathbf{x}_0) = \mu_0$ and covariance $P_0 > 0$. Assume A is nonsingular. Note that this assumption is not restrictive as (2.1) is typically a model that comes from discretizing a stochastic differential equation $dx = A_1xdt + B_1dw$, in which case $A = e^{A_1h}$, for a sampling period h , is clearly invertible. The state information is measured by a battery-powered smart sensor, which communicates with a remote state estimator through a wireless channel, and the measurement equation is

$$y_k = C\mathbf{x}_k + \mathbf{v}_k, \quad (2.2)$$

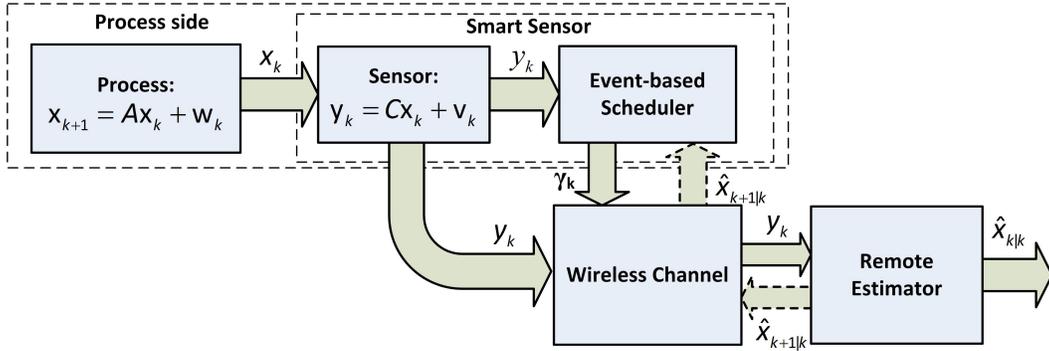


Figure 2.1: Block diagram of the overall system.

where $v_k \in \mathbb{R}^m$ is zero-mean Gaussian with covariance $R > 0$. In addition, x_0 , w_k and v_k are uncorrelated with each other. We assume (C, A) is detectable. For consideration of the limited sensor battery capacity and the communication cost, an event-based data scheduler is integrated in the sensor. At each time instant k , the measurement information y_k is sent directly to the event-based scheduler; the estimator provides a prediction $\hat{x}_{k|k-1}$ of the current state x_k , and sends the prediction $\hat{x}_{k|k-1}$ to the scheduler via the wireless channel. Based on y_k and $\hat{x}_{k|k-1}$, the scheduler computes γ_k according to the following event-triggered condition:

$$\gamma_k = \begin{cases} 0, & \text{if } \|y_k - C\hat{x}_{k|k-1}\|_\infty \leq \delta \\ 1, & \text{otherwise} \end{cases} \quad (2.3)$$

and decides whether to allow a data transmission, where δ is a tuning parameter that determines the sensitivity of the event-based scheduler. Only when $\gamma_k = 1$, the sensor transmits y_k to the estimator. As a result, if $\gamma_k = 1$, the estimator knows the exact value of y_k ; otherwise it only knows that the value of y_k lies in a known region. The ultimate goal of the estimator is to provide an estimate $\hat{x}_{k|k}$ of x_k based on the known information. Notice that this type of feedback communication strategy is not energy-saving itself and an alternative strategy is to include a copy of the estimator in the scheduler, which instead adds to the computational burden of the scheduler. We will show that the obtained result in this work in fact does not require the feedback communication except when an event occurs.

In this work, the first objective is to determine, at time k , the optimal estimate $\hat{x}_{k|k}$ of x_k that maximizes the joint probability distribution function of $x_{0:k}$ and $y_{1:k}$:

$$f_{x_{0:k}, y_{1:k}}(\hat{x}_{0|0}, x_1, \dots, x_k, \hat{y}_1, \dots, \hat{y}_k) \quad (2.4)$$

where $x_{0:k}$ and $\hat{y}_{1:k}$ are the optimization parameters. If $\gamma_t = 1$, $\hat{y}_t = y_t$; otherwise the value of \hat{y}_t lies in $[\underline{y}_t, \bar{y}_t]$ at time instant t , where

$$\begin{aligned}\underline{y}_t &= C\hat{x}_{t|t-1} - \delta\mathbf{1}_m, \\ \bar{y}_t &= C\hat{x}_{t|t-1} + \delta\mathbf{1}_m,\end{aligned}$$

$\mathbf{1}_m = \underbrace{[1 \ 1 \ \dots \ 1]}_{m \text{ times}}^\top$, $t = 1, 2, \dots, k$. Consequently, at time instant k , the estimator solves the following optimization problem:

$$\begin{aligned}\max_{x_{1:k}, \hat{y}_{1:k}} & f_{x_{0:k}, y_{1:k}}(\hat{x}_{0|0}, x_1, \dots, x_k, \hat{y}_1, \dots, \hat{y}_k) \\ \text{s.t.} & x_t = Ax_{t-1} + w_{t-1}, \\ & y_t = Cx_t + v_t. \\ & \hat{y}_t = y_t, \quad \text{if } \gamma_t = 1; \\ & \hat{y}_t \in [\underline{y}_t, \bar{y}_t], \quad \text{if } \gamma_t = 0. \\ & t \in \{1, 2, \dots, k\}.\end{aligned}\tag{2.5}$$

The objective function in (2.5) is the joint probability distribution function of $x_{0:k}$ and $y_{1:k}$, which is always Gaussian regardless of the event-driven communication. Therefore the additional information introduced by the event-based scheduler is not reflected in the objective function and is only exploited in the constraints in (2.5).

Based on the solution to (2.5), we further look into a simpler yet more interesting one-step event-based ML estimate problem by taking the determined values of x_t and \hat{y}_t (at time instant $t < k$) into account, namely, by fixing the values of x_t and \hat{y}_t to the one determined at time instant t for $t = 1, 2, \dots, k - 1$ and only considering x_k and \hat{y}_k as optimization variables:

$$\begin{aligned}\max_{x_k, \hat{y}_k} & f_{x_{0:k}, y_{1:k}}(\hat{x}_{0|0}, \dots, \hat{x}_{k-1|k-1}, x_k, \hat{y}_1, \dots, \hat{y}_k) \\ \text{s.t.} & x_k = A\hat{x}_{k-1|k-1} + w_{k-1}, \\ & y_k = Cx_k + v_k. \\ & \hat{y}_k = y_k, \quad \text{if } \gamma_k = 1; \\ & \hat{y}_k \in [\underline{y}_k, \bar{y}_k], \quad \text{if } \gamma_k = 0.\end{aligned}\tag{2.6}$$

For this problem, we show that the solution has a simple recursive form, and communication rate is possible to be analyzed in terms of upper and lower bounds from the process side.

2.3 Solution to the event-based ML estimation problem

In this section, the solution to problem (2.5) is derived. From Lemma 9.3.1 of [25], we have

$$\begin{aligned} f_{x_{0:k}, y_{1:k}}(x_{0:k}, y_{1:k}) &= \alpha \cdot \exp \left\{ -\frac{1}{2} \sum_{t=0}^{k-1} w_t^\top Q^{-1} w_t \right\} \\ &\quad \cdot \exp \left\{ -\frac{1}{2} \sum_{t=1}^k v_t^\top R^{-1} v_t \right\} \\ &\quad \cdot \exp \left\{ -\frac{1}{2} (x_0 - \mu_0)^\top P_0^{-1} (x_0 - \mu_0) \right\} \end{aligned} \quad (2.7)$$

where α is a positive constant, and w_t and v_t satisfy $w_t = x_{t+1} - Ax_t$ and $v_t = y_t - Cx_t$, respectively. As a result, the estimation problem that needs to be solved at time k is equivalent to

$$\begin{aligned} \min_{w_{0:k-1}, v_{1:k}} \quad & \sum_{t=0}^{k-1} w_t^\top Q^{-1} w_t + \sum_{t=1}^k v_t^\top R^{-1} v_t \\ & + (x_0 - \mu_0)^\top P_0^{-1} (x_0 - \mu_0) \\ \text{s.t.} \quad & x_t = Ax_{t-1} + w_{t-1}, \\ & Cx_t + v_t = y_t, \quad \text{if } \gamma_t = 1; \\ & Cx_t + v_t \leq \bar{y}_t, \quad \text{if } \gamma_t = 0; \\ & -Cx_t - v_t \leq -\underline{y}_t, \quad \text{if } \gamma_t = 0. \\ & t \in \{1, 2, \dots, k\}. \end{aligned} \quad (2.8)$$

Before continuing, let us define the value function $V(w_{0:k-1}, v_{1:k})$ as

$$\begin{aligned} V(w_{0:k-1}, v_{1:k}) &:= \sum_{t=0}^{k-1} w_t^\top Q^{-1} w_t + \sum_{t=1}^k v_t^\top R^{-1} v_t \\ &\quad + (x_0 - \mu_0)^\top P_0^{-1} (x_0 - \mu_0). \end{aligned} \quad (2.9)$$

For brevity, we use V_k^* to denote the optimal value function at time k , namely, $V_k^* := V(w_{0:k-1}^*, v_{1:k}^*)$. In the following, an active-set approach will be utilized to characterize the structure of the optimal solution to (2.8). To maintain the simplicity in the description and derivation of the results, we assume C and v_t can be decomposed as¹

$$C = [\tilde{C}_t^\top \quad \bar{C}_t^\top]^\top \quad (2.10)$$

and $v_t = [\tilde{v}_t^\top \quad \hat{v}_t^\top]^\top$, where \tilde{C}_t and \tilde{v}_t correspond to the set of active constraints

$$\tilde{v}_t + \tilde{C}_t \tilde{x}_t = \tilde{y}_t \quad (2.11)$$

¹Notice that when this decomposition assumption does not hold, the results can be proved following the same argument but at the cost of more complicated notations.

at time t that lead to the optimal solution to (2.8). Correspondingly the covariance matrix R is decomposed as

$$R = \begin{bmatrix} \tilde{R}_t & \hat{R}_t \\ \hat{R}_t^\top & \bar{R}_t \end{bmatrix}^{-1}. \quad (2.12)$$

Define $R_t^* := (\tilde{R}_t - \hat{R}_t \bar{R}_t^{-1} \hat{R}_t^\top)^{-1}$. Utilizing these notations, for the problem in (2.8), we have the following results.

Theorem 2.1. *The optimal solution to problem (2.8) has the following properties:*

1. *The optimal prediction satisfies*

$$\hat{x}_{t+1|t} = A\hat{x}_{t|t};$$

The optimal estimation $\hat{x}_{t|t}$ satisfies

$$\hat{x}_{t|t} = \begin{cases} \hat{x}_{t|t-1} + P_{t|t-1} C^\top (R + C P_{t|t-1} C^\top)^{-1} \\ \quad (y_t - C \hat{x}_{t|t-1}), & \text{if } \gamma_t = 1; \\ \hat{x}_{t|t-1} + P_{t|t-1} \tilde{C}_t^\top (R_t^* + \tilde{C}_t P_{t|t-1} \tilde{C}_t^\top)^{-1} \\ \quad (\tilde{y}_t - \tilde{C}_t \hat{x}_{t|t-1}), & \text{if } \gamma_t = 0. \end{cases} \quad (2.13)$$

with $\hat{x}_{0|0} = \mu_0$, $P_{t|t-1} = A P_{k-1|k-1} A^\top + Q$, $P_{0|0} = P_0$; if $\gamma_t = 1$,

$$P_{t|t} = P_{t|t-1} - P_{t|t-1} C^\top (R + C P_{t|t-1} C^\top)^{-1} C P_{t|t-1};$$

if $\gamma_t = 0$,

$$P_{t|t} = P_{t|t-1} - P_{t|t-1} \tilde{C}_t^\top (R_t^* + \tilde{C}_t P_{t|t-1} \tilde{C}_t^\top)^{-1} \tilde{C}_t P_{t|t-1}.$$

2. *The optimal value function V_t^* satisfies*

$$V_t^* = (x_t - \hat{x}_{t|t})^\top P_{t|t}^{-1} (x_t - \hat{x}_{t|t}) + \Upsilon_t, \quad (2.14)$$

with

$$V_0^* = (x_0 - \hat{x}_{0|0})^\top P_{0|0}^{-1} (x_0 - \hat{x}_{0|0}) + \Upsilon_0,$$

$\Upsilon_0 = 0$; *if $\gamma_t = 1$,*

$$\Upsilon_t = \Upsilon_{t-1} + (y_t - C \hat{x}_{t|t-1})^\top (R + C P_{t|t-1} C^\top)^{-1} (y_t - C \hat{x}_{t|t-1});$$

if $\gamma_t = 0$,

$$\Upsilon_t = \Upsilon_{t-1} + (\tilde{y}_t - \tilde{C}_t \hat{x}_{t|t-1})^\top (R_t^* + \tilde{C}_t P_{t|t-1} \tilde{C}_t^\top)^{-1} (\tilde{y}_t - \tilde{C}_t \hat{x}_{t|t-1}).$$

Proof. The problem in (2.8) is a quadratic optimization problem with linear equality and inequality constraints. According to the first-order Karush-Kuhn-Tucker conditions (which is necessary and sufficient for local optimality in this case) [10], the global optimizer of this problem can be obtained by enumerating all sets of active constraints and testing the feasibility with respect to problem (2.8) of the solution to the corresponding quadratic optimization problem with equality constraints. Therefore, to characterize the structure of the optimal solution, it suffices to consider the set of optimal active constraints given in (2.11).

Without loss of generality, we first claim that the optimal value function V_{t-1}^* at time instant $k-1$ has the following form:

$$V_{t-1}^* = \Upsilon_{t-1} + (x_{t-1} - \hat{x}_{t-1|t-1})^\top P_{t-1|t-1}^{-1} (x_{t-1} - \hat{x}_{t-1|t-1}), \quad (2.15)$$

and then we provide an inductive proof for it. Note that this is satisfied at $k=1$ with $\Upsilon_0 = 0$ and $V_0^* = (x_0 - \hat{x}_{0|0})^\top P_{0|0}^{-1} (x_0 - \hat{x}_{0|0}) + \Upsilon_0$, where $\hat{x}_{0|0} = \mu_0$, $P_{0|0} = P_0$.

If $\gamma_t \neq 0$, following a similar argument as that in the proof of Lemma 9.6.1 of [25], we have

$$\begin{aligned} V_t^* &= (x_t - \hat{x}_{t|t})^\top P_{t|t}^{-1} (x_t - \hat{x}_{t|t}) + \Upsilon_t, \\ \Upsilon_t &= \Upsilon_{t-1} + (y_t - C\hat{x}_{t|t-1})^\top (R + CP_{t|t-1}C^\top)^{-1} \\ &\quad (y_t - C\hat{x}_{t|t-1}), \\ \hat{x}_{t|t-1} &= A\hat{x}_{t-1|t-1}, \\ \hat{x}_{t|t} &= \hat{x}_{t|t-1} + P_{t|t-1}C^\top (R + CP_{t|t-1}C^\top)^{-1} \\ &\quad (y_t - C\hat{x}_{t|t-1}), \\ P_{t|t-1} &= AP_{t-1|t-1}A^\top + Q, \\ P_{t|t} &= P_{t|t-1} - P_{t|t-1}C^\top (R + CP_{t|t-1}C^\top)^{-1} CP_{t|t-1}. \end{aligned} \quad (2.16)$$

If $\gamma_t = 0$, at time instant t , we solve

$$\begin{aligned} \min_{w_{t-1}, v_t} \quad & w_{t-1}^\top Q^{-1} w_{t-1} + v_t^\top R^{-1} v_t \\ & + (A^{-1}x_t - A^{-1}Bw_{t-1})^\top P_{t-1}^{-1} \\ & (A^{-1}x_t - A^{-1}Bw_{t-1}) \\ \text{s.t.} \quad & \tilde{v}_t + \tilde{C}_t \tilde{x}_t = \tilde{y}_t, \end{aligned} \quad (2.17)$$

where Υ_{t-1} is independent of w_{t-1} and v_t , and the relationship $x_{t-1} = A^{-1}x_t - A^{-1}w_{t-1}$ is used.

From (2.12), the resulting optimization problem with equality constraints can be written as

$$\begin{aligned} \min_{w_{t-1}, \tilde{v}_t, \hat{v}_t} \quad & w_{t-1}^\top Q^{-1} w_{t-1} + \tilde{v}_t^\top \tilde{R}_t \tilde{v}_t + \hat{v}_t^\top \hat{R}_t \hat{v}_t \\ & + \hat{v}_t^\top \hat{R}_t^\top \tilde{v}_t + \tilde{v}_t^\top \tilde{R}_t^\top \hat{v}_t + \Upsilon_{t-1} \\ & + (A^{-1}x_t - A^{-1}Bw_{t-1})^\top P_{t-1}^{-1} \\ & (A^{-1}x_t - A^{-1}Bw_{t-1}) \\ \text{s.t.} \quad & \tilde{v}_t + \tilde{C}_t \tilde{x}_t = \tilde{y}_t. \end{aligned} \quad (2.18)$$

This problem can be solved in two steps:

1. Optimal prediction. In this step, we identify the optimal w_{t-1} . Due to the structure of the problem, we obtain the same result as that in the case $\gamma_t = 1$. The optimizer is

$$\begin{aligned} w_{t-1}^* &= (A^{-\top} P_{t-1|t-1}^{-1} A^{-1} + Q^{-1})^{-1} \\ &\quad A^{-\top} P_{t-1|t-1}^{-1} (A^{-1} x_t - \hat{x}_{t-1|t-1}), \end{aligned} \quad (2.19)$$

and the optimal prediction and the corresponding value function are $\hat{x}_{t|t-1} = A\hat{x}_{t-1|t-1}$ and

$$\begin{aligned} V_t^* &= (x_t - \hat{x}_{t|t-1})^\top P_{t|t-1}^{-1} (x_t - \hat{x}_{t|t-1}) \\ &\quad + \tilde{v}_t^\top \tilde{R}_t \tilde{v}_t + \tilde{v}_t^\top \hat{R}_t \hat{v}_t + \hat{v}_t^\top \hat{R}_t^\top \tilde{v}_t + \hat{v}_t^\top \bar{R}_t \hat{v}_t \\ &\quad + \Upsilon_{t-1}, \end{aligned} \quad (2.20)$$

respectively, where $P_{t|t-1} = AP_{t-1|t-1}A^\top + Q$.

2. Measurement update. In this step, we optimize V_t^* with respect to \tilde{v}_t and \hat{v}_t subject to the active constraints. To do this, we include the constraints into the objective function and differentiate V_t^* with respect to x_t and \hat{v}_t , respectively, which leads to

$$\begin{aligned} P_{t|t-1}^{-1} (x_t - \hat{x}_{t|t-1}) - \tilde{C}_t^\top \tilde{R}_t (\tilde{y}_t - \tilde{C}_t x_t) - \tilde{C}_t \hat{R}_t \hat{v}_t &= 0, \\ \hat{R}_t^\top (\tilde{y}_t - \tilde{C}_t x_t) + \bar{R}_t \hat{v}_t &= 0. \end{aligned} \quad (2.21)$$

Some further matrix manipulations lead to

$$\begin{aligned} V_t^* &= (x_t - \hat{x}_{t|t})^\top P_{t|t}^{-1} (x_t - \hat{x}_{t|t}) + \Upsilon_t, \\ \Upsilon_t &= \Upsilon_{t-1} + (\tilde{y}_t - \tilde{C}_t \hat{x}_{t|t-1})^\top \\ &\quad (R_t^* + \tilde{C}_t P_{t|t-1} \tilde{C}_t^\top)^{-1} (\tilde{y}_t - \tilde{C}_t \hat{x}_{t|t-1}), \\ \hat{x}_{t|t} &= \hat{x}_{t|t-1} + P_{t|t-1} \tilde{C}_t^\top (R_t^* + \tilde{C}_t P_{t|t-1} \tilde{C}_t^\top)^{-1} \\ &\quad (\tilde{y}_t - \tilde{C}_t \hat{x}_{t|t-1}), \\ P_{t|t} &= P_{t|t-1} - P_{t|t-1} \tilde{C}_t^\top (R_t^* + \tilde{C}_t P_{t|t-1} \tilde{C}_t^\top)^{-1} \\ &\quad \tilde{C}_t P_{t|t-1}, \\ R_t^* &= (\tilde{R}_t - \hat{R}_t \bar{R}_t^{-1} \hat{R}_t^\top)^{-1}. \end{aligned} \quad (2.22)$$

This completes the proof. \square

Remark 2.1. The above result provides insights into the structure of the optimal solution to (2.5). However, to find the optimal solution to (2.5) at time k , we need to consider all possible (3^{mk}) combinations of active constraint sets considering $t = 1, 2, \dots, k$ and compare the corresponding value functions according to (2.16) and (2.22). As a result, the computation burden will increase exponentially with respect to the time horizon. Alternatively, since the problem is a Quadratic Programming (QP) problem subject to linear constraints, standard QP solvers can be applied to find the optimal solution as well. However, the issue is that the dimension of the optimization parameters in the QP problem increases linearly with respect to k , due to the lack of a recursive structure of the optimal solution from time k to $k + 1$ (This follows from the observation that the set of optimal active constraints for problem (2.5) at time k may not be part of the set of optimal active constraints for the problem at time $k + 1$).

2.4 One-step event-based ML state estimation

Motivated by the implementation issues of the optimal solution to problem (2.5) discussed in Remark 2.1, we further look into the one-step event-based ML estimation problem in (2.6) in this section. As will be shown later, this formulation allows us to obtain a recursive solution, and is a consequence of the compromise between optimality and implementability. The solution to this problem is first presented, and then a communication rate analysis is delivered by deriving upper and lower bounds for the communication rate from the process side.

2.4.1 Solution to the problem

In this case, the problem that needs to be solved at time k is equivalent to

$$\begin{aligned}
V(w_{k-1}^\dagger, v_k^\dagger) &:= \\
&\min_{w_{k-1}, v_k} \quad \sum_{t=0}^{k-1} w_t^\top Q^{-1} w_t + \sum_{t=1}^k v_t^\top R^{-1} v_t \\
&\quad + (x_0 - \mu_0)^\top P_0^{-1} (x_0 - \mu_0) \\
&s.t. \quad x_k = Ax_{k-1} + w_{k-1}, \\
&\quad \quad Cx_k + v_k = y_k, \quad \text{if } \gamma_k = 1; \\
&\quad \quad Cx_k + v_k \leq \bar{y}_k, \quad \text{if } \gamma_k = 0; \\
&\quad \quad -Cx_k - v_k \leq -\underline{y}_k, \quad \text{if } \gamma_k = 0. \\
&\quad \quad w_{t-1} = w_{t-1}^\dagger, \quad v_t = v_t^\dagger, \\
&\quad \quad t \in \{1, 2, \dots, k-1\}.
\end{aligned} \tag{2.23}$$

For notational simplicity, define $V_k^\dagger := V(w_{k-1}^\dagger, v_k^\dagger)$. For this problem, we have the following result.

Theorem 2.2. *The optimal solution to problem (2.23) has the following properties:*

1. *The optimal prediction is unbiased and satisfies $\hat{x}_{k+1|k} = A\hat{x}_{k|k}$; the optimal estimation $\hat{x}_{k|k}$ is also unbiased and satisfies*

$$\hat{x}_{k|k} = \begin{cases} A\hat{x}_{k-1|k-1} + P_{k|k-1}C^\top \\ (R + CP_{k|k-1}C^\top)^{-1}(y_k - CA\hat{x}_{k-1|k-1}), & \text{if } \gamma_k = 1; \\ A\hat{x}_{k-1|k-1}, & \text{if } \gamma_k = 0. \end{cases} \quad (2.24)$$

$$\begin{aligned} \text{with } \hat{x}_{0|0} = \mu_0, P_{k|k-1} = AP_{k-1|k-1}A^\top + Q, P_{0|0} = P_0; \text{ if } \gamma_k = 1, \\ P_{k|k} = P_{k|k-1} - P_{k|k-1}C^\top(R + CP_{k|k-1}C^\top)^{-1} \\ CP_{k|k-1}; \\ \text{if } \gamma_k = 0, \\ P_{k|k} = AP_{k-1|k-1}A^\top + Q. \end{aligned}$$

2. *The optimal value function V_k^\dagger satisfies*

$$V_k^\dagger = (x_k - \hat{x}_{k|k})^\top P_{k|k}^{-1}(x_k - \hat{x}_{k|k}) + \Upsilon_k, \quad (2.25)$$

with

$$V_0^\dagger = (x_0 - \hat{x}_{0|0})^\top P_{0|0}^{-1}(x_0 - \hat{x}_{0|0}) + \Upsilon_0,$$

$\Upsilon_0 = 0$; if $\gamma_k = 1$,

$$\begin{aligned} \Upsilon_k = \Upsilon_{k-1} + (y_k - C\hat{x}_{k|k-1})^\top \\ (R + CP_{k|k-1}C^\top)^{-1}(y_k - C\hat{x}_{k|k-1}); \\ \text{if } \gamma_k = 0, \Upsilon_k = \Upsilon_{k-1}. \end{aligned}$$

Proof. The proof of this result follows from a similar argument as that in Theorem 2.1. In particular, when $\gamma_k = 0$ and no constraint is active, the counter-part of results in (2.22) reduces to:

$$\begin{aligned} V_k^\dagger &= (x_k - \hat{x}_{k|k})^\top P_{k|k}^{-1}(x_k - \hat{x}_{k|k}) + \Upsilon_k, \\ \Upsilon_k &= \Upsilon_{k-1}, \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1}, \\ P_{k|k} &= P_{k|k-1}. \end{aligned} \quad (2.26)$$

The optimizer for this unconstrained case satisfies all constraints in (2.23). In addition, since $R > 0$, by Schur complement, we have $R_k^* > 0$ in (2.22), which further implies $\Upsilon_k \geq \Upsilon_{k-1}$ and thus the solution to the constrained case leads to a cost larger or equal than the solution to the unconstrained case. Therefore, when $\gamma_k = 0$, the optimization problem in (2.23) is solved by (2.26).

Finally, the unbiasedness of the optimal prediction and estimation follows directly from their structure and the fact that $\hat{x}_{0|0} = \mathbf{E}(x_0)$, which completes the proof. \square

The above result indicates that when the exact value of the measurement is unavailable and the information of the set-valued measurement $\hat{y}_t \in [\underline{y}_t, \bar{y}_t]$ is exploited instead, the one-step optimal state prediction also serves the optimal estimation in the sense of one-step maximum likelihood. Note that this does not hold in general for the event-based ML estimation problem in (2.5). Since the conditional distribution $f_{x_{1:k}|y_{1:k}}(x_{1:k}|y_{1:k})$ is no longer Gaussian due to the additional information of the set-valued measurement, the ML estimate does not necessarily coincide with the MMSE estimate for event-based estimation, which is different from the case of periodic state estimation of linear Gaussian systems.

Remark 2.2. In [73], when $\gamma_k = 1$, $P_{k|k}$ has the same update equation; but when $\gamma_k = 0$, $P_{k|k}$ evolves rather differently. The resultant estimate here has a much simpler form, which does not require to solve the integrations in [73]. Notice that for original ML estimation problems in (2.5), $P_{k|k}$ is not the estimation error covariance matrix for the estimate $\hat{x}_{k|k}$, but rather a time-varying parameter that helps to generate the ML estimate subject to the event-triggering rule. As a result, the obtained update equations have essentially different meanings compared with those in [64].

2.4.2 Analysis of the communication rate

Based on the proposed state estimation strategy, the average communication rate is analyzed in this section. Viewed from the process side, the resultant state estimator behaves exactly like the standard Kalman filter with intermittent observations [45, 64]: When $\gamma_k = 1$, the optimal estimator considers both time and measurement updates of the Kalman filter; When $\gamma_k = 0$, the optimal estimator only performs the time update. Therefore on the process side the resultant prediction error $\hat{e}_{k|k-1} := x_k - \hat{x}_{k|k-1}$ is zero-mean Gaussian with covariance $P_{k|k-1}$. Denote $z_k := y_k - C\hat{x}_{k|k-1}$. Since $y_k - C\hat{x}_{k|k-1} = C\hat{e}_{k|k-1} + v_k$, we have $\mathbf{E}(z_k) = 0$ and $\mathbf{E}(z_k z_k^\top) := \Phi_k = CP_{k|k-1}C^\top + R$. Define

$$\Omega := \{z \in \mathbb{R}^m \mid \|z\|_\infty \leq \delta\},$$

which is an m -dimensional cube with edge length 2δ . We have

$$\mathbf{E}(\gamma_k) = 1 - \int_{\Omega} f_{z_k}(z) dz, \quad (2.27)$$

where $f_{z_k}(z) = (2\pi)^{-m/2} (\det \Phi_k)^{-1/2} \exp(-\frac{1}{2} z^\top \Phi_k^{-1} z)$. Note that due to the Gaussian kernel and the structure of Ω , analytical calculation of the integration in (2.27) is not

possible. In the following, we provide lower and upper bounds for $\mathbf{E}(\gamma_k)$ in terms of incomplete Gamma functions.

We first focus on the calculation of integrals of the following form, which plays an important part in the derivation of the bounds:

$$\int_{\Omega_0} f_{z_k}(z) dz,$$

where $\Omega_0 := \{z \mid z^\top \Phi_k^{-1} z \leq r^2\}$. Define $\Omega_0^\perp := \{z \mid z^\top \Phi_k^{-1} z > r^2\}$. Since $\Omega_0 \cup \Omega_0^\perp = \mathbb{R}^m$, $\int_{\Omega_0} f_{z_k}(z) dz = 1 - \int_{\Omega_0^\perp} f_{z_k}(z) dz$. For the integration over Ω_0^\perp , we have the following result.

Lemma 2.1. $\int_{\Omega_0^\perp} f_{z_k}(z) dz = \Gamma(m/2, r^2/2) / \Gamma(m/2)$.

Proof.

$$\begin{aligned} & \int_{\Omega_0^\perp} f_{z_k}(z) dz \\ &= \int_{z^\top \Phi_k^{-1} z > r^2} (2\pi)^{-\frac{m}{2}} (\det \Phi_k)^{-\frac{1}{2}} \exp(-\frac{1}{2} z^\top \Phi_k^{-1} z) dz \\ &= (2\pi)^{-\frac{m}{2}} \int_{p^\top p > r^2} \exp(-\frac{1}{2} p^\top p) dp \\ &= (2\pi)^{-\frac{m}{2}} \frac{2\pi^{m/2}}{\Gamma(m/2)} \int_r^\infty v^{m-1} \exp(-v^2/2) dv \\ &= \frac{1}{\Gamma(m/2)} \int_{r^2/2}^\infty t^{m/2-1} \exp(-t) dt \\ &= \frac{\Gamma(m/2, r^2/2)}{\Gamma(m/2)}, \end{aligned} \tag{2.28}$$

where the second equality is obtained by using $p = \Phi_k^{-1/2} z$ and

$$dp = (\det \Phi_k)^{-1/2} dz,$$

the third equality is obtained by converting the Cartesian coordinates $p = [p_1, p_2, \dots, p_m]^\top$ to polar coordinates $[v, \theta_1, \theta_2, \dots, \theta_{m-1}]^\top$ and

$$dp = v^{m-1} \sin^{m-2} \theta_1 \sin^{m-3} \theta_2 \cdots \sin \theta_{m-2} dv d\theta_1 d\theta_2 \cdots d\theta_{m-1},$$

the fourth one is obtained by using $t = v^2/2$ and the surface area formula for an $(m-1)$ -dimensional unit sphere $S_{m-1} = 2\pi^{m/2} / \Gamma(m/2)$, where $\Gamma(m/2) := \int_0^\infty t^{m/2-1} \exp(-t) dt$, and the fifth one follows from the definition of the incomplete Gamma function $\Gamma(a, b) := \int_b^\infty t^{a-1} \exp(-t) dt$. \square

Note that the Gamma functions and incomplete Gamma functions can be iteratively calculated according to $\Gamma(z+1) = z\Gamma(z)$, $\Gamma(1/2) = \sqrt{\pi}$ and $\Gamma(a, b) = (a-1)\Gamma(a-1, b) + b^{a-1} \exp(-b)$, $\Gamma(1/2, b) = 2\sqrt{\pi}[1 - Q(\sqrt{2b})]$, respectively, where $Q(z) := \int_z^\infty \frac{1}{\sqrt{2\pi}} \exp(-\frac{t^2}{2}) dt$ is the standard Q -function.

Now we focus on the tightest inner and outer ellipsoidal approximations of Ω . Define $\underline{\Omega}_{k,1}$ as the largest ellipsoid that is contained in Ω and satisfies

$$\underline{\Omega}_{k,1} := \{z \in \mathbb{R}^m \mid z^\top \Phi_k^{-1} z \leq \underline{\delta}_{k,1}^2\}. \quad (2.29)$$

Define $\bar{\Omega}_{k,1}$ as the smallest ellipsoid that contains Ω and satisfies

$$\bar{\Omega}_{k,1} := \{z \in \mathbb{R}^m \mid z^\top \Phi_k^{-1} z \leq \bar{\delta}_{k,1}^2\}. \quad (2.30)$$

The relationship of $\bar{\Omega}_{k,1}$, $\underline{\Omega}_{k,1}$ and Ω for the case of $m = 2$ is shown in Fig. 2.2. To determine $\bar{\Omega}_{k,1}$ and $\underline{\Omega}_{k,1}$, the calculation of $\underline{\delta}_{k,1}$ and $\bar{\delta}_{k,1}$ are presented in the following.

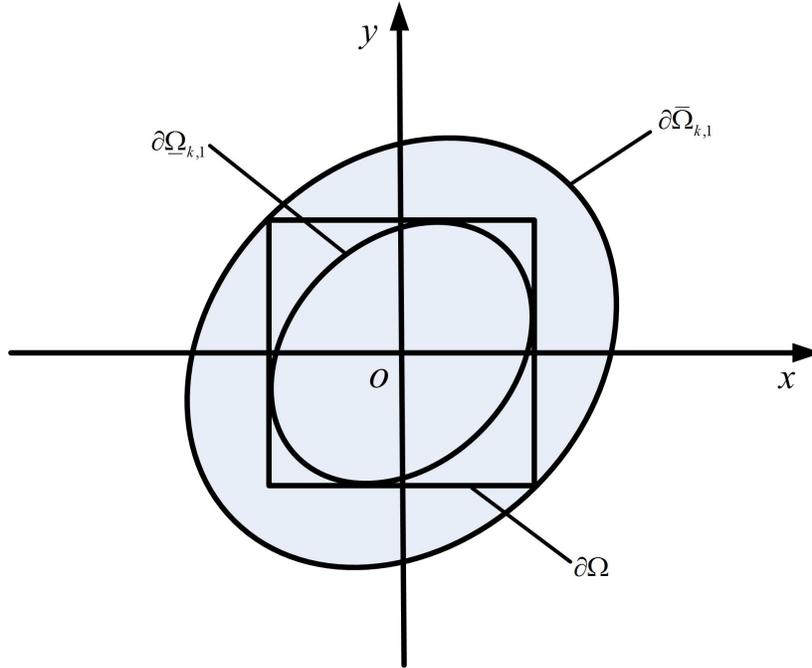


Figure 2.2: Relationship of inner and outer ellipsoidal approximations of Ω for the case of $m = 2$.

Since $\Phi_k > 0$, there exists an orthogonal matrix U_k such that $U_k^\top \Phi_k^{-1} U_k = \Lambda_k$, where

$$\Lambda_k := \text{diag}\{\lambda_{k,1}, \lambda_{k,2}, \dots, \lambda_{k,m}\}$$

and $0 < \lambda_{k,1} \leq \lambda_{k,2} \leq \dots \leq \lambda_{k,m}$. The value of $\bar{\delta}_{k,1}$ is relatively easier to determine. By convexity, at least two of the vertices of the m -dimensional cube Ω lie on the boundary of $\bar{\Omega}_{k,1}$, and the other vertices are either contained in $\bar{\Omega}_{k,1}$ or on the boundary of $\bar{\Omega}_{k,1}$ as well. Therefore, the value of $\bar{\delta}_{k,1}$ can be calculated as

$$\bar{\delta}_{k,1} = \max_{z_i \in \{\delta, -\delta\}, i \in \{1, 2, \dots, m\}} \sqrt{z^\top \Phi_k^{-1} z}, \quad (2.31)$$

where $z = [z_1, z_2, \dots, z_m]^\top$.

To calculate $\underline{\delta}_{k,1}$, the following bi-level optimization problem needs to be solved:

$$\begin{aligned} \max_i \quad & z_i^* \\ \text{s.t.} \quad & z_i^* = \max_z z_i \\ & \text{s.t. } z^\top (\Phi_k^{-1}) z = 1. \end{aligned} \quad (2.32)$$

The major difficulty is in the lower level problem, namely,

$$\begin{aligned} \max_z \quad & z_i \\ \text{s.t.} \quad & z^\top (\Phi_k^{-1}) z = 1. \end{aligned} \quad (2.33)$$

To solve this problem, we consider an alternative parameterization of z . For each $z \in \mathbb{R}^m$, there exists a unique $p_k = [p_{k,1}, p_{k,2}, \dots, p_{k,m}]^\top \in \mathbb{R}^m$ such that $z = U_k p_k$. For z satisfies $z^\top (\Phi_k^{-1}) z = 1$, the corresponding p_k satisfies $\sum_{i=1}^m \lambda_{k,i} (p_{k,i})^2 = 1$, and thus an equivalent parameterization of p_k is

$$p_k = \left[\frac{\sin \theta_{k,1}}{\sqrt{\lambda_{k,1}}}, \frac{\sin \theta_{k,2} \cos \theta_{k,1}}{\sqrt{\lambda_{k,2}}}, \dots, \frac{\sin \theta_{k,m-1} \prod_{i=1}^{m-2} \cos \theta_{k,i}}{\sqrt{\lambda_{k,m-1}}}, \frac{\prod_{i=1}^{m-1} \cos \theta_{k,i}}{\sqrt{\lambda_{k,m}}} \right]^\top. \quad (2.34)$$

According to this parameterization, problem (2.33) is equivalent to

$$\begin{aligned} z_i^* = \quad & \max_{\theta_{k,1}, \theta_{k,2}, \dots, \theta_{k,m-1}} \alpha_{k,i,m} \prod_{t=1}^{m-1} \cos \theta_{k,t} \\ & + \sum_{t=1}^{m-1} \alpha_{k,i,t} \sin \theta_{k,t} \prod_{j=1}^{t-1} \cos \theta_{k,j} \\ \text{s.t.} \quad & \theta_{k,t} \in [0, 2\pi], \quad t \in \{1, 2, \dots, m-1\}, \end{aligned} \quad (2.35)$$

where $\alpha_{k,i,j} = \frac{u_{k,i,j}}{\sqrt{\lambda_{k,j}}}$, $u_{k,i,j}$ being the element in the i th row and j th column of U_k . Notice that we allow all $\theta_{k,i}$'s to take value in $[0, 2\pi]$ in the above problem, which would lead to redundant parameterization of the ellipsoid. However, this does not affect the results presented here since each combination of $\theta_{k,i} \in [0, 2\pi]$ corresponds to a unique point on the ellipsoid.

For the above problem, we have the following result.

Lemma 2.2. *The optimal solution to problem (2.35) equals $z_i^* = \sqrt{\sum_{j=1}^m \alpha_{k,i,j}^2}$.*

Proof. First, for $\theta_{k,m-1}$, we have

$$\begin{aligned}
& \sum_{t=1}^{m-2} \alpha_{k,i,t} \sin \theta_{k,t} \prod_{j=1}^{t-1} \cos \theta_{k,j} \\
& + \alpha_{k,i,m-1} \sin \theta_{k,m-1} \prod_{t=1}^{m-2} \cos \theta_{k,t} + \alpha_{k,i,m} \prod_{t=1}^{m-1} \cos \theta_{k,t} \\
& = \sum_{t=1}^{m-2} \alpha_{k,i,t} \sin \theta_{k,t} \prod_{j=1}^{t-1} \cos \theta_{k,j} \\
& \quad + (\alpha_{k,i,m-1} \sin \theta_{k,m-1} + \alpha_{k,i,m} \cos \theta_{k,m-1}) \prod_{t=1}^{m-2} \cos \theta_{k,t} \\
& = \sum_{t=1}^{m-2} \alpha_{k,i,t} \sin \theta_{k,t} \prod_{j=1}^{t-1} \cos \theta_{k,j} \\
& \quad + \sqrt{\alpha_{k,i,m-1}^2 + \alpha_{k,i,m}^2} \left(\frac{\alpha_{k,i,m-1}}{\sqrt{\alpha_{k,i,m-1}^2 + \alpha_{k,i,m}^2}} \sin \theta_{k,m-1} \right. \\
& \quad \left. + \frac{\alpha_{k,i,m}}{\sqrt{\alpha_{k,i,m-1}^2 + \alpha_{k,i,m}^2}} \cos \theta_{k,m-1} \right) \prod_{t=1}^{m-2} \cos \theta_{k,t} \\
& \leq \sum_{t=1}^{m-2} \alpha_{k,i,t} \sin \theta_{k,t} \prod_{j=1}^{t-1} \cos \theta_{k,j} \\
& \quad + \operatorname{sgn}(\prod_{t=1}^{m-2} \cos \theta_{k,t}) \sqrt{\alpha_{k,i,m-1}^2 + \alpha_{k,i,m}^2} \prod_{t=1}^{m-2} \cos \theta_{k,t},
\end{aligned} \tag{2.36}$$

where the last inequality holds with equality for the maximizing $\theta_{k,m-1}$. The rest of the proof follows recursively, which completes the proof. \square

Based on the above result, the optimal solution to problem (2.32) can be written as $\max_i \sqrt{\sum_{j=1}^m \alpha_{k,i,j}^2}$. As a result, we have

$$\underline{\delta}_{k,1} = \frac{\delta}{\max_i \sqrt{\sum_{j=1}^m \alpha_{k,i,j}^2}}.$$

We are now in a position to present the first set of upper and lower bounds on the expectation of the communication rates.

Theorem 2.3. *For the state estimation scheme in Fig. 2.1 and the event-based scheduler in (2.3), the expected sensor to estimator communication rate $\mathbf{E}(\gamma_k)$ is bounded by*

$$\frac{\Gamma(m/2, \bar{\delta}_{k,1}^2/2)}{\Gamma(m/2)} \leq \mathbf{E}(\gamma_k) \leq \frac{\Gamma(m/2, \underline{\delta}_{k,1}^2/2)}{\Gamma(m/2)}, \tag{2.37}$$

with $\bar{\delta}_{k,1} = \max_{z_i \in \{\delta, -\delta\}, i \in \{1, 2, \dots, m\}} \sqrt{z^\top \Phi_k^{-1} z}$ and $\underline{\delta}_{k,1} = \frac{\delta}{\max_{i \in \{1, 2, \dots, m\}} \sqrt{\sum_{j=1}^m \alpha_{k,i,j}^2}}$.

Proof. From Lemma 2.1, we have

$$\int_{\Omega_0} f_{z_k}(z) dz = 1 - \Gamma(m/2, r^2/2) / \Gamma(m/2).$$

Since $\underline{\Omega}_{k,1} \subset \Omega \subset \bar{\Omega}_{k,1}$,

$$1 - \frac{\Gamma(m/2, \bar{\delta}_{k,1}^2/2)}{\Gamma(m/2)} \leq \int_{\Omega} f_{z_k}(z) dz \leq 1 - \frac{\Gamma(m/2, \underline{\delta}_{k,1}^2/2)}{\Gamma(m/2)}.$$

From (2.27), we have

$$\frac{\Gamma(m/2, \bar{\delta}_{k,1}^2/2)}{\Gamma(m/2)} \leq \mathbf{E}(\gamma_k) \leq \frac{\Gamma(m/2, \underline{\delta}_{k,1}^2/2)}{\Gamma(m/2)}.$$

□

Although the computation effort required to calculate the Gamma and incomplete Gamma functions is small, the determination of $\bar{\delta}_{k,1}$ and $\underline{\delta}_{k,1}$ can be computationally expensive, especially for the case when m is large. In fact, the number of vertices considered in computing $\bar{\delta}_{k,1}$ equals 2^m . In the following, upper and lower bounds with lower computational burden are further explored.

Define $\underline{S} \subset \mathbb{R}^m$ as the largest sphere contained in Ω :

$$\underline{S} := \{z \in \mathbb{R}^m \mid z^\top z \leq \delta^2\}, \quad (2.38)$$

and define $\bar{S} \subset \mathbb{R}^m$ as the smallest sphere that contains Ω :

$$\bar{S} := \{z \in \mathbb{R}^m \mid z^\top z \leq \delta^2 m\}. \quad (2.39)$$

Based on \bar{S} and \underline{S} , we further define $\underline{\Omega}_{k,2} \subset \underline{S}$ as the largest ellipsoid that is contained in \underline{S} and satisfies

$$\underline{\Omega}_{k,2} := \{z \in \mathbb{R}^m \mid z^\top \Phi_k^{-1} z \leq \underline{\delta}_{k,2}^2\}, \quad (2.40)$$

and define $\bar{\Omega}_{k,2}$ as the smallest ellipsoid that contains \bar{S} and satisfies:

$$\bar{\Omega}_{k,2} := \{z \in \mathbb{R}^m \mid z^\top \Phi_k^{-1} z \leq \bar{\delta}_{k,2}^2\}. \quad (2.41)$$

The relationship of \bar{S} , \underline{S} , $\bar{\Omega}_{k,2}$, $\underline{\Omega}_{k,2}$ and Ω for the case of $m = 2$ is shown in Fig. 2.3. In the following, we show that $\underline{\delta}_{k,2}$ and $\bar{\delta}_{k,2}$ can be analytically calculated based on the eigenvalues of the prediction error covariance matrix. Since $\underline{\Omega}_{k,2}$ and $\bar{\Omega}_{k,2}$ are convex, the major effort in determining $\underline{\delta}_{k,2}$ and $\bar{\delta}_{k,2}$ is to calculate the maximal and minimal values of $z^\top z$ over the set $\{z \in \mathbb{R}^m \mid z^\top \Phi_k^{-1} z = 1\}$, for which we have the following result.

Lemma 2.3. *For all $z \in \mathbb{R}^m$ satisfying $z^\top \Phi_k^{-1} z = 1$, $1/\bar{\lambda}_k \leq z^\top z \leq 1/\underline{\lambda}_k$ holds, where $\underline{\lambda}_k$ and $\bar{\lambda}_k$ are the smallest and largest eigenvalues of Φ_k^{-1} , respectively.*

Proof. By the properties of the orthogonal matrices, we have $z^\top z = p_k^\top p_k$. From (2.34),

$$p_k^\top p_k = \frac{\prod_{i=1}^{m-1} \cos^2 \theta_{k,i}}{\lambda_{k,m}} + \sum_{i=1}^{m-1} \frac{\sin^2 \theta_{k,i} \prod_{j=1}^{i-1} \cos^2 \theta_{k,j}}{\lambda_{k,i}},$$

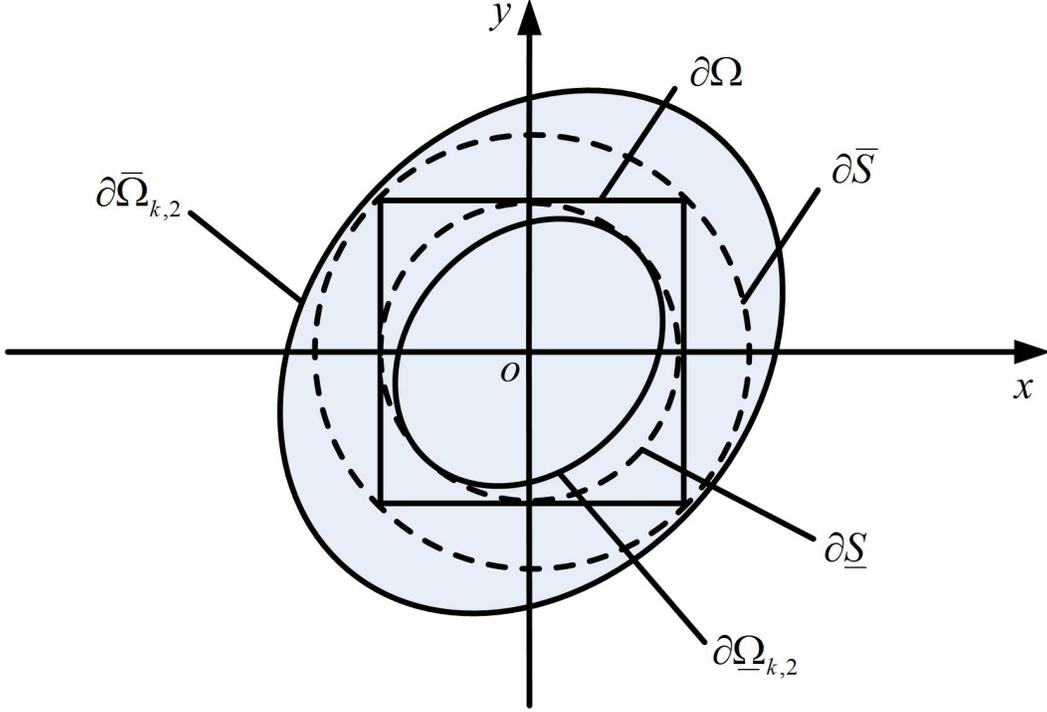


Figure 2.3: Relationship of inner and outer ellipsoidal approximations of Ω for the case of $m = 2$.

and we have

$$\begin{aligned}
\frac{1}{\lambda_{k,m}} &= \frac{1}{\lambda_{k,m}} \left(\prod_{i=1}^{m-1} \cos^2 \theta_{k,i} + \sum_{i=1}^{m-1} \sin^2 \theta_{k,i} \prod_{j=1}^{i-1} \cos^2 \theta_{k,j} \right) \\
&\leq \frac{\prod_{i=1}^{m-1} \cos^2 \theta_{k,i}}{\lambda_{k,m}} + \sum_{i=1}^{m-1} \frac{\sin^2 \theta_{k,i} \prod_{j=1}^{i-1} \cos^2 \theta_{k,j}}{\lambda_{k,i}} \\
&\leq \frac{1}{\lambda_{k,1}} \left(\prod_{i=1}^{m-1} \cos^2 \theta_{k,i} + \sum_{i=1}^{m-1} \sin^2 \theta_{k,i} \prod_{j=1}^{i-1} \cos^2 \theta_{k,j} \right) \\
&= \frac{1}{\lambda_{k,1}},
\end{aligned} \tag{2.42}$$

where equality is achieved in the first inequality for $\theta_{k,1} = \theta_{k,2} = \dots = \theta_{k,m-1} = 0$, and strict equality holds for the second inequality if $\theta_{k,1} = \pi/2$, $\theta_{k,2} = \theta_{k,3} = \dots = \theta_{k,m-1} = 0$. In conclusion, we have $1/\bar{\lambda}_k \leq z^\top z \leq 1/\underline{\lambda}_k$. \square

From the above lemma, it is straightforward to obtain that for $z \in \{z \in \mathbb{R}^m | z^\top \Phi_k^{-1} z = r^2\}$, $r^2/\bar{\lambda}_k \leq z^\top z \leq r^2/\underline{\lambda}_k$ holds. Therefore we have $\underline{\delta}_{k,2} = \sqrt{\underline{\lambda}_k} \delta$ and $\bar{\delta}_{k,2} = \sqrt{\bar{\lambda}_k} m \delta$.

Based on the above discussions, we have the following result on the bounds of the communication rate.

Theorem 2.4. *For the state estimation scheme in Fig. 2.1 and the event-based scheduler in (2.3), the expected sensor to estimator communication rate $\mathbf{E}(\gamma_k)$ is bounded*

by

$$\frac{\Gamma(m/2, \bar{\delta}_{k,2}^2/2)}{\Gamma(m/2)} \leq \mathbf{E}(\gamma_k) \leq \frac{\Gamma(m/2, \underline{\delta}_{k,2}^2/2)}{\Gamma(m/2)}, \quad (2.43)$$

with $\bar{\delta}_{k,2} = \sqrt{m\bar{\lambda}_k}\delta$ and $\underline{\delta}_{k,2} = \sqrt{\underline{\lambda}_k}\delta$.

Proof. The proof of this result is the same as that of Theorem 2.3 and thus is omitted. \square

Notice that when the sensor has only one measurement channel, i.e., $m = 1$, all proposed upper and lower bounds in Theorems 2.3 and 2.4 coincide. In this case, the exact value of $\mathbf{E}(\gamma_k)$ can be obtained by either of the results. The above results provide time-varying lower and upper bounds on the communication rate, which depend on the eigenvalues of the prediction error covariance matrices. If the system is stable, uniform upper and lower bounds (with respect to k) can be provided as well.

Corollary 2.1. If the system in (2.1) is stable, the communication rate is bounded by

$$\frac{\Gamma(m/2, \bar{\delta}^2/2)}{\Gamma(m/2)} \leq \mathbf{E}(\gamma_k) \leq \frac{\Gamma(m/2, \underline{\delta}^2/2)}{\Gamma(m/2)}, \quad (2.44)$$

as $k \rightarrow \infty$, where $\bar{\delta} = \sqrt{m\bar{\lambda}_1}\delta$, $\underline{\delta} = \sqrt{\underline{\lambda}_2}\delta$, $\lambda_1 = \max\{\text{eig}[(C\underline{P}C^\top + R)^{-1}]\}$, \underline{P} being the stabilizing solution to the Riccati equation

$$P = APA^\top - APC^\top[CPC^\top + R]^{-1}CPA^\top + Q,$$

and $\lambda_2 = \min\{\text{eig}[(C\bar{P}C^\top + R)^{-1}]\}$, \bar{P} being the stabilizing solution to the algebraic Lyapunov equation

$$P = APA^\top + Q.$$

Proof. This result follows from Theorem 2.4 and the monotonicity properties of the solutions to the time-varying Riccati difference equations [21]. \square

In addition, due to the structure of the optimal estimate, the estimator does not need to send the optimal prediction to the remote scheduler when no event occurs, since the same prediction can be generated by the scheduler based on the previous prediction (which is also the optimal estimation) with little additional computation cost. In this way, the communication cost is further reduced by the proposed state estimation method.

Remark 2.3. Notice that viewed from the remote side (namely, the remote estimator), the prediction error $\hat{e}_{k|k-1}$ is not Gaussian conditioned on the information provided by the event-triggering set. This issue has been addressed in [73] and is one of the main difficulties in event-triggered estimator design problems. On the other hand, for the problem of communication-rate analysis, we need to move back to the process side and study the problem from an unconditional perspective. The main reason of doing this is that the constraints are introduced by the event-triggering conditions, which have a different nature from the hard constraints that indeed restrict the supports of the random processes (e.g., input/output saturation, logic mechanisms). Therefore, viewed from the process side, x_k and $\hat{x}_{k|k-1}$ still evolve as unconstrained random processes, and are still Gaussian due to their linear dependence on the Gaussian noise and initial states. The benefit of this choice is that the problem is slightly simplified and allows us to derive upper and lower bounds on the communication rates, which coincide for the case $m = 1$.

Remark 2.4. For the case of the ML estimation problem in (2.5), the communication rate of the corresponding solution is difficult to analyze. The main problem lies in that the unconditional distribution of the resultant prediction error is unknown, due to the existence of the inequality constraints in (2.5). As a result, it is only possible to discuss the communication rate by analyzing the distribution of the prediction error given the condition $\hat{y}_t = \tilde{y}_t$ for the event time instants. However, we do not look into the details of this issue since the calculation burden of determining the resultant estimate itself goes beyond the limited calculation capability of the smart sensor.

2.5 Examples

In this section, simulation examples are presented to illustrate the proposed results. Apart from the proposed one-step event-based estimator, three other estimators are also implemented to compare the tradeoff between average communication rate and estimation performance, including the Kalman filter with periodic packet dropouts², Kalman filter with intermittent observations [64], and the event-based MMSE estimator [73].

²To implement this filter, we assume the first L measurement are lost in each period T ; when no measurement is available, only prediction is performed.

2.5.1 Example 1

Consider a third-order system of the form in (2.1) measured by a sensor with:

$$\begin{aligned}
 A &= \begin{bmatrix} 0.6559 & -0.1689 & 0.2196 \\ 0.0241 & 0.5864 & -0.0379 \\ 0.0378 & -0.0452 & 0.6206 \end{bmatrix}, \\
 Q &= \begin{bmatrix} 2.2861 & 0 & 0 \\ 0 & 2.3106 & 0 \\ 0 & 0 & 2.2683 \end{bmatrix}, \\
 C &= \begin{bmatrix} 0.0216 & 0.8383 & 0.0164 \\ 0.6359 & 0.0034 & 0.5174 \end{bmatrix}, \\
 R &= \begin{bmatrix} 0.0312 & 0 \\ 0 & 0.0385 \end{bmatrix}.
 \end{aligned}$$

First, δ is set to 0.8. The proposed one-step event-based ML sensor data scheduling and state estimation strategy is applied and the estimation performance is shown in Fig. 2.4. The plot of γ_k is presented in Fig. 2.5, and the proposed bounds on $\mathbf{E}(\gamma_k)$ are shown in Fig. 2.6. The shape of the time-varying bounds is due to the converging and monotonic properties of the solutions to the Riccati equations. To further compare the tightness of the different upper and lower bounds, the absolute distances between the upper and lower bounds calculated according to Theorems 2.3 and 2.4 are shown in Fig. 2.7. In addition, the actual average communication rate calculated according to Fig. 2.5 is 51%.

Second, the tradeoff between estimation performance and communication rate is further analyzed by changing the values of δ and compared with the three estimators mentioned above, and the results are shown in Fig. 2.8, where the average communication rate is defined by

$$\tilde{\gamma} := \frac{1}{N} \sum_{k=1}^N \gamma_k, \tag{2.45}$$

N being the simulation horizon, and the average estimation error is defined by

$$\epsilon := \frac{1}{N} \sum_{k=1}^N \|x_k - \hat{x}_{k|k}\|^2. \tag{2.46}$$

It is shown that at the same average communication rate, the performance of the Kalman filter with intermittent observations is better than that with periodic packet dropouts, and the one-step event-based ML estimator achieves very close estimation performance to the event-based MMSE estimator, which is much improved compared

with the Kalman filter with periodic packet dropouts and intermittent observations. On the other hand, at the same average communication rate $\tilde{\gamma}$, since the MMSE estimator needs feedback communication (including the state prediction and normalizing matrix) at each time instant, the amount of communication burden is much smaller for the proposed one-step ML estimator.

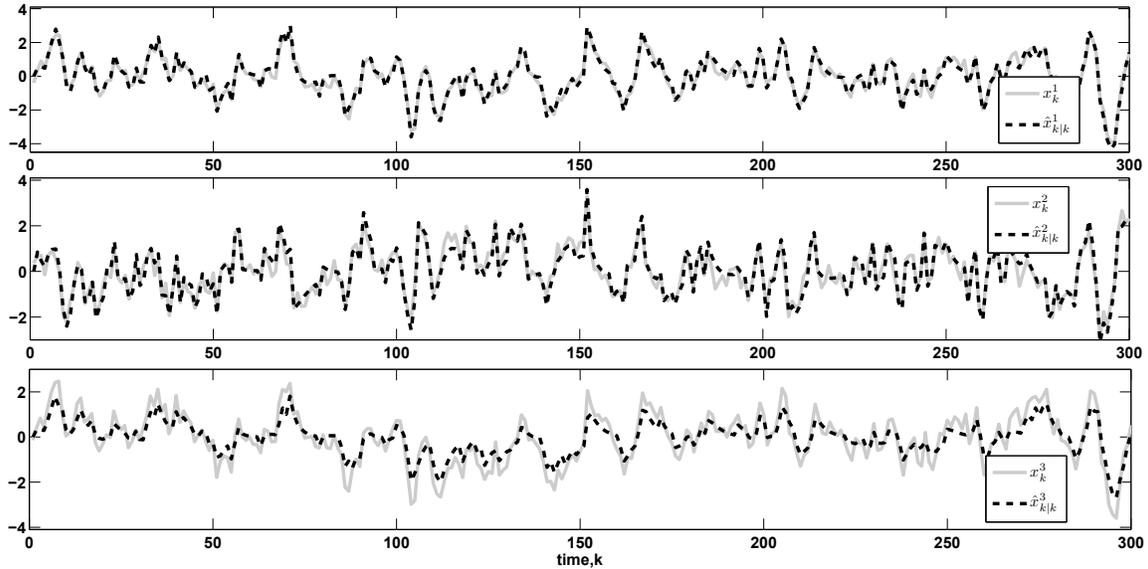


Figure 2.4: Performance of the state estimation strategy (x_k^i denotes the trajectory of the i th state at time instant k , and \hat{x}_k^i denotes the estimate of the i th state at time instant k).

2.5.2 Example 2: Sensorless event-based estimation of a DC motor system

In this subsection, we further illustrate the proposed results with a sensorless remote estimation problem involving a DC motor. The mechanical and electrical dynamics of the DC motor system are given by [20]:

$$\begin{aligned} J_m \frac{d^2\theta_m}{dt^2} + b \frac{d\theta_m}{dt} + T_L &= K_t i_a, \\ L_a \frac{di_a}{dt} + R_a i_a &= v_a - K_e \frac{d\theta_m}{dt}, \end{aligned}$$

where J_m is the rotor inertia, θ_m is the shaft rotational position, T_L is the load torque, b is the viscous friction coefficient, K_t is the torque constant, K_e is the electric constant,

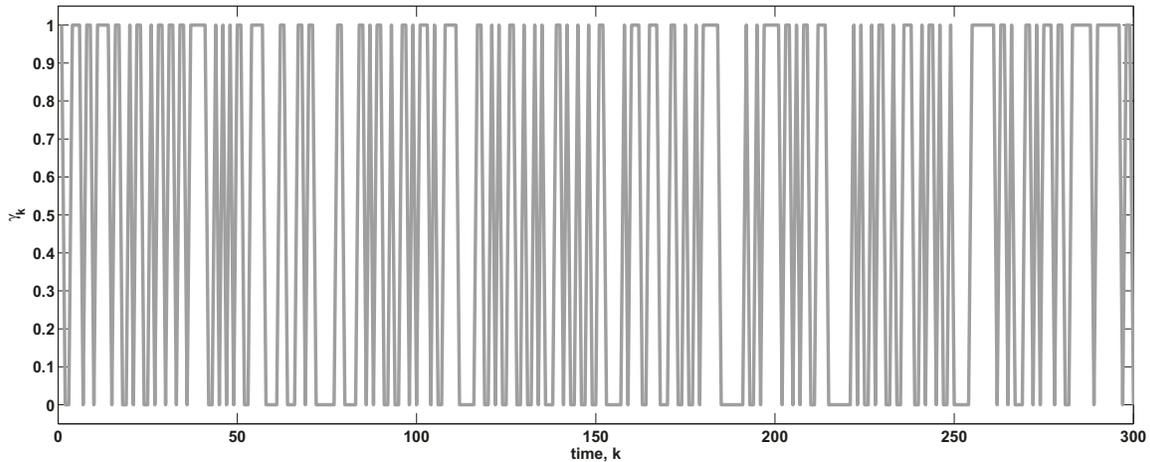


Figure 2.5: Plot of γ_k .

L_a is the armature inductance, R_a is the armature resistance, and v_a is the DC voltage input. The motor parameters are summarized in Tab. 2.1, which are obtained based on experimental measurements of a 500W permanent magnet DC motor with rated speed, current and voltage equal to 314.16rad/s, 3.5A and 180V, respectively [13].

The objective here is to estimate the shaft rotational position θ_m , shaft rotational speed $\dot{\theta}_m$ and armature current i_a with a current sensor (e.g., a Hall effect sensor). This is called the sensorless control/estimation technique³ in the industrial electronics community [29, 30, 67]. The estimation is performed by a remote estimator collecting the measurement information through a battery-powered wireless channel. In this work, we consider the load type to be piecewise constant, which can be provided by a synchronous machine. Since both the load torque and DC voltage are only subject to step changes, it is reasonable to assume that these signals are known/generated by the remote estimator.

To implement the estimator, a state-space model is first derived. Since the direct consideration of shaft rotational position will introduce an undetectable mode (in fact the corresponding eigenvalue equals 1) to the system, we choose the state vector as $x := [\dot{\theta}_m \ i_a]^\top$, the input vector as $u := [T_L \ v_a]^\top$, and the measurement output as

³Here ‘sensorless’ means the elimination of the speed sensor.

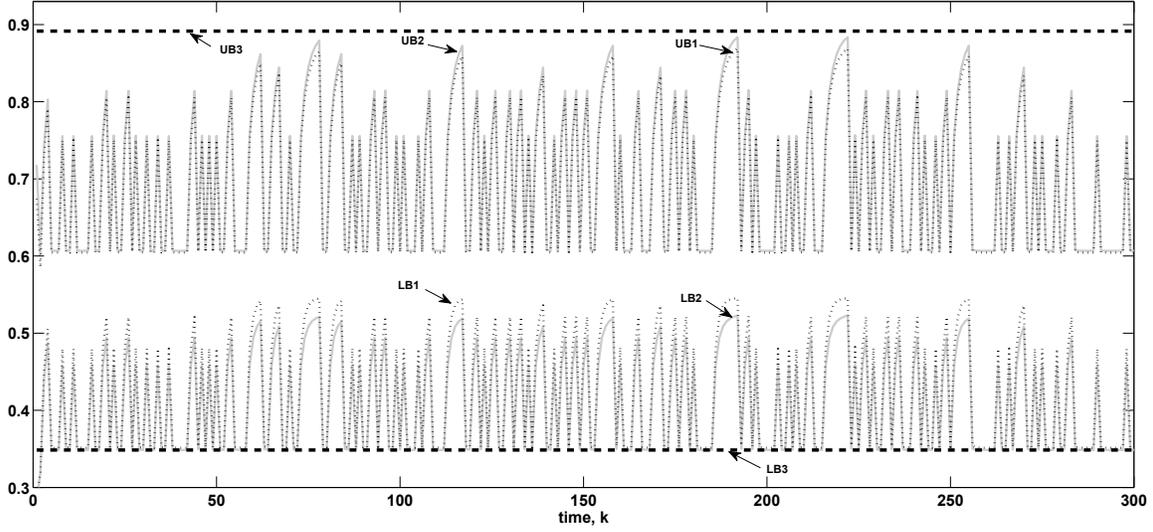


Figure 2.6: Bounds on $\mathbf{E}(\gamma_k)$ (**UB1** and **LB1** respectively denote the upper and lower bounds derived in Theorem 2.3, **UB2** and **LB2** respectively denote the upper and lower bounds derived in Theorem 2.4, and **UB3** and **LB3** respectively denote the upper and lower bounds derived in Corollary 2.1).

$y := i_a$, which lead to the state-space model:

$$\dot{x}(t) = \begin{bmatrix} -\frac{b}{J_m} & \frac{K_t}{J_m} \\ -\frac{K_e}{L_a} & -\frac{R_a}{L_a} \end{bmatrix} x(t) + \begin{bmatrix} -\frac{1}{J_m} & 0 \\ 0 & \frac{1}{L_a} \end{bmatrix} u(t), \quad (2.47)$$

$$y(t) = \begin{bmatrix} 0 & 1 \end{bmatrix} x(t). \quad (2.48)$$

Notice that based on the estimation of the rotational speed, the shaft rotational position can be estimated as well. With these parameter settings, a discrete-time model is obtained with sampling time chosen as $T_s = 0.001$ s:

$$x_{k+1} = \begin{bmatrix} 0.9951 & 0.2289 \\ -0.0177 & 0.8672 \end{bmatrix} x_k + \begin{bmatrix} -0.4158 & 0.0038 \\ 0.0038 & 0.0301 \end{bmatrix} u_k + w_k, \quad (2.49)$$

$$y_k = \begin{bmatrix} 0 & 1 \end{bmatrix} x_k + v_k, \quad (2.50)$$

where w_k and v_k are further introduced to model the noisy operating environment. Specifically, $w_k := [w_k^1 \ w_k^2]^\top$ with w_k^1 characterizing the mechanical noise that couples into the speed-loop and w_k^2 modelling the electrical noise that couples into the voltage input, and v_k models the measurement noise. The covariance matrices of w_k and v_k are assumed to be $Q = \begin{bmatrix} 0.2013 & 0.0430 \\ 0.0430 & 0.0363 \end{bmatrix}$ and $R = 0.03$, respectively. Since both inputs are

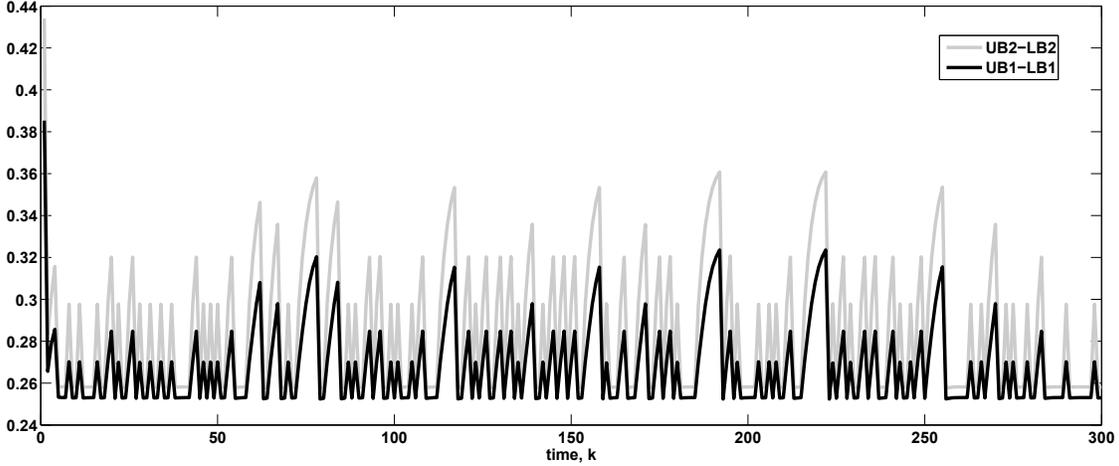


Figure 2.7: Comparison of the relative tightness of different bounds (**UB1-LB1** denotes the distance between the upper and lower bounds derived in Theorem 2.3, and **UB2-LB2** denotes the distance between the upper and lower bounds derived in Theorem 2.4).

Table 2.1: Motor parameters

Parameter	Value	Unit
L_a	20.25	H
R_a	16.4	Ω
K_e	0.0233	V/(rad/sec)
K_t	0.0183	N·m/A
J_m	9	g·cm ²
b	0.0064	N·m/(rad/sec)

known to the remote estimator, the proposed results can be applied by only modifying the prediction as $\hat{x}_{k|k-1} = A\hat{x}_{k-1|k-1} + Bu_k$, B being the discretized input matrix.

First, the event-triggering level is set to $\delta = 0.4A$. The input signals utilized are plotted in Fig. 2.9. The proposed event-based sensor data scheduling and state estimation strategy is applied and the estimation performance is shown in Fig. 2.10. For this case, the proposed upper and lower bounds on $\mathbf{E}(\gamma_k)$ in Theorems 2.3 and 2.4 coincide, and therefore the exact value of $\mathbf{E}(\gamma_k)$ can be determined, which is shown in Fig. 2.11. In addition, the actual average communication rate is 19.35%.

Second, by varying the event-triggering threshold δ , the relationship between estimation performance and average communication rate is further analyzed and compared with the other three methods. The results are shown in Fig. 2.12, where the aver-

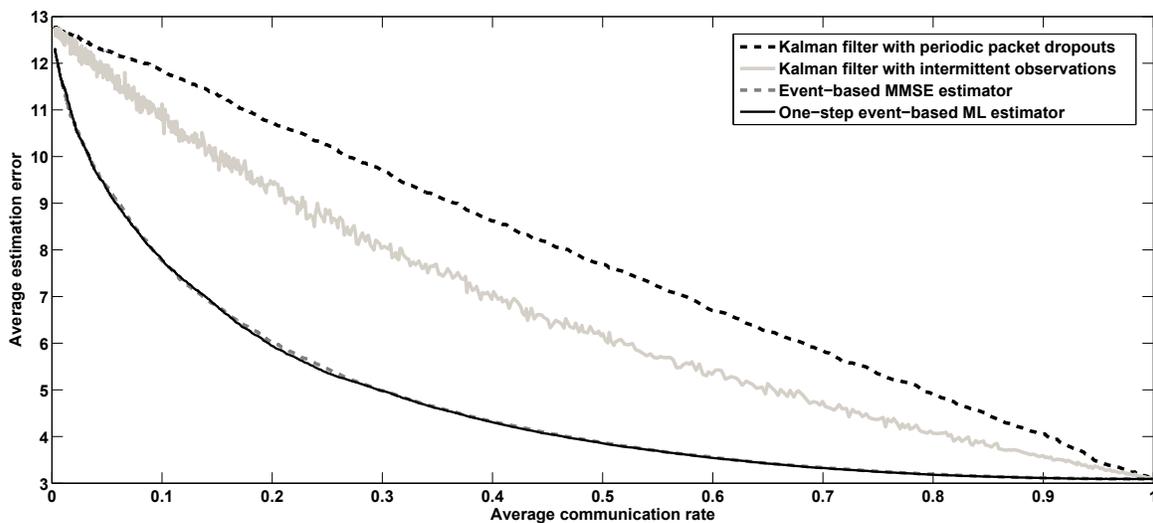


Figure 2.8: Tradeoff between estimation performance and communication rate.

age estimation error and average communication rate are defined in (2.46) and (2.45), respectively. Again it is shown that the performance of the one-step event-based ML estimator is very close to that of the event-based MMSE estimator, at a much decreased communication burden (due to the fact that no feedback communication is required for predicted state when $\gamma_k = 0$ and no normalizing matrix needs to be transmitted to the scheduler at all time instants). In this sense, the ML estimator has greater applicability in this wireless communication scenario with satisfactory estimation performance and potentially prolonged battery life.

2.6 Summary

In this chapter, an event-based state estimation problem is studied in the framework of ML estimation. We show that the optimal estimator is parameterized by a time-varying Riccati equation associated with an exponential computational complexity. A one-step event-based estimation problem with reduced computation burden is also studied and a recursive solution similar to the Kalman filter with intermittent observations is obtained. Results on the communication rate are obtained for this problem. For the case of scalar sensors, the exact value of the expectation of the communication rate can be determined; for the case of sensors with vector-valued measurements, upper and lower bounds are provided on the expectation of communication rate by exploring the inner and outer ellipsoidal approximations of the m -dimensional cube. An

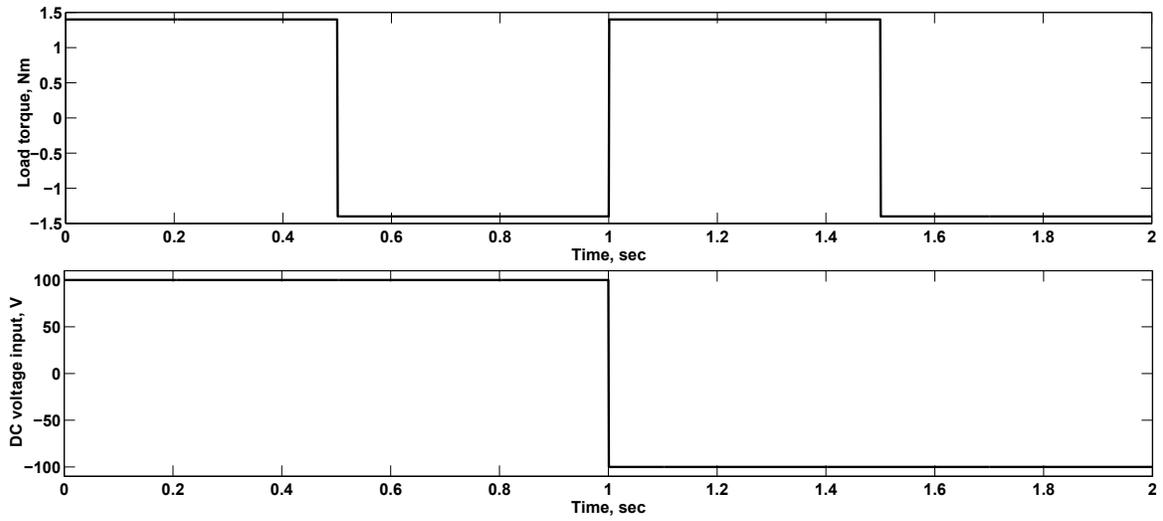


Figure 2.9: Plot of the input signals.

alternative approach to reduce the computational burden of the general event-based ML estimation problems is to consider the formulation of receding horizon estimation [4, 25, 48, 57], which points out the topic for possible future research.

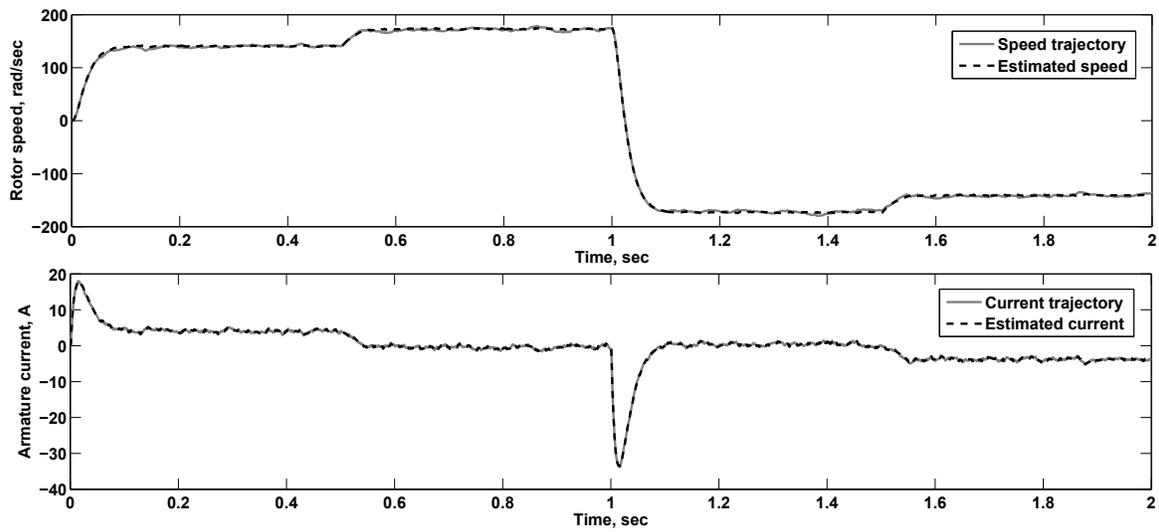


Figure 2.10: Performance of the state estimation strategy.

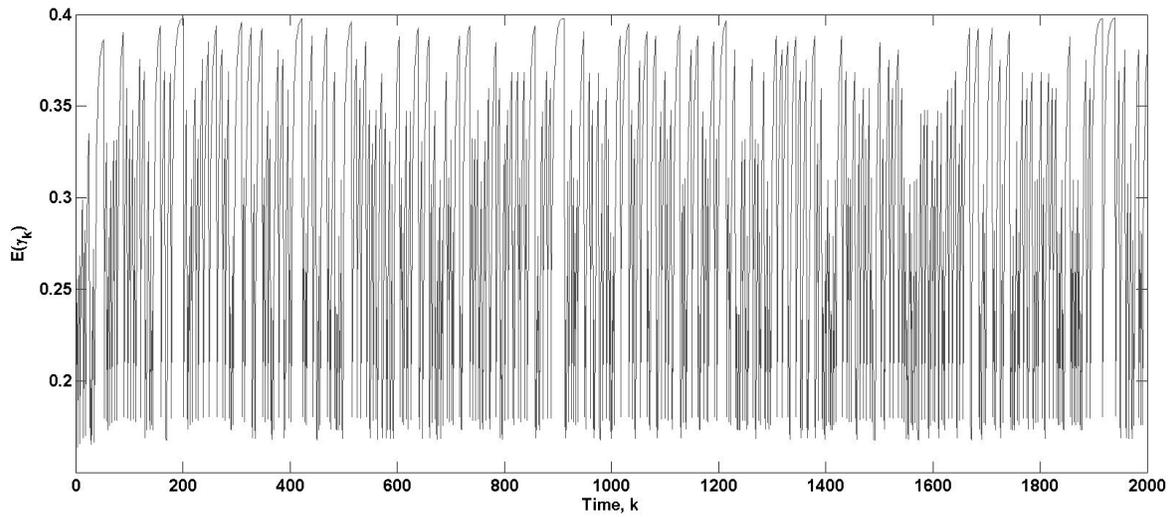


Figure 2.11: Plot of $\mathbf{E}(\gamma_k)$.

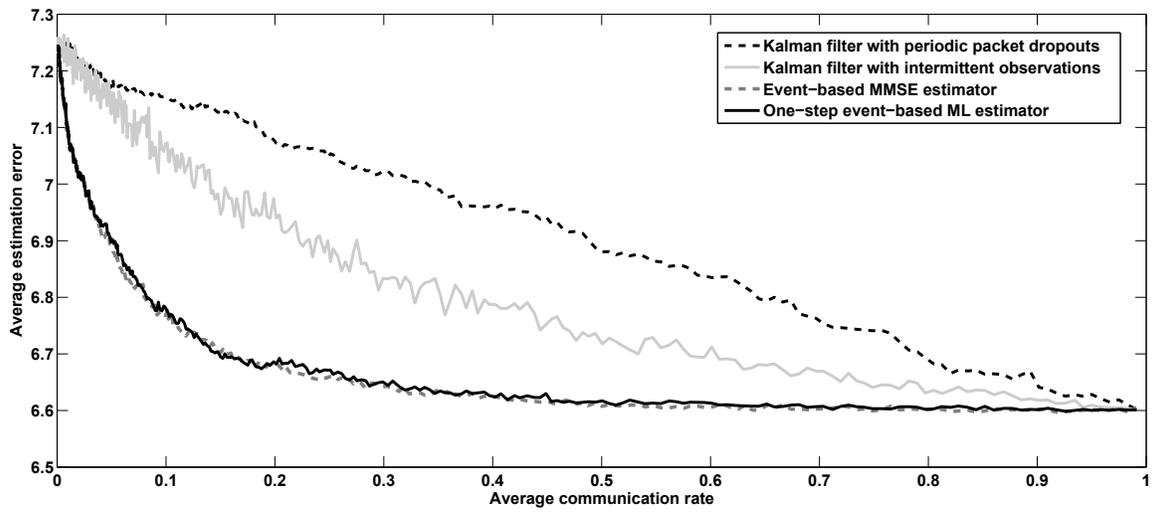


Figure 2.12: Tradeoff between estimation performance and communication rate.

Chapter 3

An approximate Gaussian approach to event-triggered state estimation*

3.1 Introduction

In this chapter, we consider the scenario that the process is measured by a network of sensors and that each sensor chooses to provide its latest measurement update according to its own event-triggering condition. In this case, hybrid information is provided by the whole group of sensors as well as the event-triggering sets. For the sensors whose event-triggering conditions are satisfied, the exact values of the sensor outputs are known, providing “point-valued measurement information” to the estimator; for sensors that the event-triggering conditions are not satisfied, some information contained in the event-triggering sets is known to the estimator as well, to which we refer as “set-valued measurement information” in this work. The basic goal is to find the MMSE estimate given the hybrid measurement information. As will be addressed later, the main issues arise from the computational aspect, due to the non-Gaussianity of the *a posteriori* distributions. Therefore we focus on the derivation of an approximate (due to the Gaussian assumption) MMSE estimate that possesses a simple structure but still inherits the important properties of the exact optimal estimate. The main contributions are summarized as follows:

1. An approximate MMSE estimate induced by the hybrid measurement information provided by a sequence of sensors has been derived. We show that the

*Parts of the results in this chapter appeared in *Automatica*, vol. 50, no. 6, pp. 1641-1648, 2014.

estimate is determined by the conditional mean and covariance of the innovations. The results are valid for general event-triggering schemes, and reduce to the results obtained in [73] if only one sensor and the level-based event-triggering conditions are considered.

2. Insights on the optimal estimate when each sensor has only one channel are provided. In this case, closed-form recursive state estimate update equations are obtained. Utilizing the recent results on the partial order of uncertainty and information [12], we show that the exploration of the set-valued information guarantees improved estimation performance in terms of smaller estimation error covariance. The results are equally applicable to multiple-channel sensors with uncorrelated/correlated measurement noises but separate event-triggering conditions on each channel.
3. Extensive Monte Carlo experiments are performed to test the effectiveness of the proposed estimator. Compared with the Kalman filter that only exploits the received point-valued measurements, the proposed estimator provides almost-guaranteed improved performance, which is not sensitive to the sensor sequence used.

Notation: \mathbb{N} and \mathbb{N}^+ denote the sets of nonnegative and positive integers, respectively. Let $i, j \in \mathbb{N}$. Denote $\mathbb{N}_{i:j} := \{i, i+1, \dots, j\}$. \mathbb{R} denotes the set of real numbers. For $m, n \in \mathbb{N}^+$, $\mathbb{R}^{m \times n}$ denotes the set of m by n real-valued matrices, whereas \mathbb{R}^m is short for $\mathbb{R}^{m \times 1}$. \mathbb{B} denotes the set $\{0, 1\}$. For a random variable x , $\mathbf{E}(x)$ denotes its expectation, $\mathbf{Cov}(x)$ denotes its covariance; for a univariate random variable, $\mathbf{Cov}(\cdot)$ is also used to denote the variance.

3.2 System description and problem setup

Consider a linear time-invariant process that evolves in discrete time driven by white noise:

$$x_{k+1} = Ax_k + w_k, \quad (3.1)$$

where $x_k \in \mathbb{R}^n$ is the state, and $w_k \in \mathbb{R}^n$ is the process noise, which is zero-mean Gaussian with covariance $Q \geq 0$. The initial value x_0 of the state is Gaussian with $\mathbf{E}(x_0) = \mu_0$, and covariance P_0 . The state information is measured by a number of battery-powered sensors, which communicate with the state estimator through a

wireless channel, and the output equations are

$$y_k^i = C^i x_k + v_k^i, \quad (3.2)$$

where $v_k^i \in \mathbb{R}^m$ is zero-mean Gaussian with covariance $R^i > 0$. In addition, x_0 , w_k and v_k^i are uncorrelated with each other. We assume the number of sensors equals M . Considering limitation in sensor battery capacity and the communication costs, an event-based data scheduler is equipped with each sensor i . At each time instant k , sensor i produces a measurement y_k^i , and the scheduler of sensor i tests the event-triggering condition

$$\gamma_k^i = \begin{cases} 0, & \text{if } y_k^i \in \Xi_k^i \\ 1, & \text{otherwise} \end{cases} \quad (3.3)$$

where Ξ_k^i denotes the event-triggering set of sensor i at time k , and decides whether to allow a data transmission. If $\gamma_k^i = 1$, sensor i sends y_k^i to the estimator through the wireless channel. Notice that the event-triggering scheme in (3.3) is fairly general and covers most schemes considered in the literature and industrial applications, e.g., the “send on delta” strategy and the level-based triggering conditions (not necessarily being symmetric). For many previously considered event-triggering schemes (e.g., the level-based event-triggering conditions in [73] and Chapter 2 of this thesis), feedback communication from the estimator to the sensor is needed at certain time instants as the event is related to the innovation; however, since the event-triggering sets Ξ_k^i can be designed offline, the remote estimator will have full knowledge of them without communication. In this way, the proposed results are applicable to battery-powered WSNs, where it is normally too costly to use feedback communication.

Since the main task is to study event-based estimation and sensor fusion, we assume the capacity of the channel is greater than M so that it is possible for the sensors to communicate with the estimator at the same time.

Let \hat{x}_k^i denote the optimal estimate of x_k after updating the measurement of the i th sensor, and denote P_k^i as the corresponding covariance matrix.¹ Denote \mathbb{S}_+^n as the set of symmetric positive semidefinite matrices. Define the functions $h(\cdot): \mathbb{S}_+^n \rightarrow \mathbb{S}_+^n$ and $\tilde{g}_i(\cdot, \cdot): \mathbb{S}_+^n \times \mathbb{R} \rightarrow \mathbb{S}_+^n$ as follows:

$$\begin{aligned} h(X) &:= AXA^\top + Q, \\ \tilde{g}_i(X, \vartheta) &:= X - \vartheta X(C^i)^\top [C^i X(C^i)^\top + R^i]^{-1} C^i X. \end{aligned} \quad (3.4)$$

For brevity, we denote $\tilde{g}_i(X, 1)$ as $\tilde{g}_i(X)$. Denote $\mathcal{Y}_k := \{\mathcal{Y}_k^1, \mathcal{Y}_k^2, \dots, \mathcal{Y}_k^M\}$ as the collection of measurement information received by the estimator. Notice that if $\gamma_k^i = 1$,

¹Here we denote the 0th sensor as the case that no sensor information has been fused, namely, the prediction case.

$\mathcal{Y}_k^i = \{y_k^i\}$; otherwise, $\mathcal{Y}_k^i = \{y_k^i \mid y_k^i \in \Xi_k^i\}$. In the latter case, although y_k^i is unknown, it is still jointly Gaussian with x_k . Further define

$$\mathcal{I}_k^i := \{\mathcal{Y}_1, \mathcal{Y}_2, \dots, \mathcal{Y}_{k-1}, \{\mathcal{Y}_k^1, \mathcal{Y}_k^2, \dots, \mathcal{Y}_k^i\}\} \quad (3.5)$$

for $i \in \mathbb{N}_{1:M}$, and in this way, we are able to summarize all the information we have in \mathcal{I}_k^i before considering the additional information \mathcal{Y}_k^{i+1} from sensor $i+1$ at time k . The objective of our work is to explore the MMSE estimate of the process state (namely, $\mathbf{E}(x_k \mid \mathcal{I}_k^M)$) by taking into account all given information, namely, the set- and point-valued measurements provided by the sensor network as well as the event-triggering schemes.

When the state information is contained in combined point- and set-valued measurements, following a standard Bayesian argument, the exact MMSE estimate is the mean of the distribution of x_k conditioned on \mathcal{I}_k^M ,

$$\mathbf{E}(x_k \mid \mathcal{I}_k^M) = \int_{\mathbb{R}^n} x f_{x_k}(x \mid \mathcal{I}_k^M) dx. \quad (3.6)$$

The major problem of this estimate arises from the computational aspect, due to the fact that the conditional distribution of x_k in (3.6) is no longer Gaussian when set-valued measurements are provided. This conditional distribution can be updated recursively by fusing the information sequentially

$$f_{x_k}(x \mid \mathcal{I}_k^i) = \frac{f_{x_k}(x \mid \mathcal{I}_k^{i-1}) \int_{\Xi_k^i} f_{y_k^i}(y \mid \mathcal{I}_k^{i-1}, x_k = x) dy}{\int_{\mathbb{R}^n} f_{x_k}(x \mid \mathcal{I}_k^{i-1}) \int_{\Xi_k^i} f_{y_k^i}(y \mid \mathcal{I}_k^{i-1}, x_k = x) dy dx}, \quad (3.7)$$

and the final result does not depend on the sensor sequence utilized during the fusion procedure (since the distribution is unique). However, analytical solutions to the integrations in (3.7) rarely exist and the only method to implement this estimate is numerical integration, which is inevitably expensive in computation.

On the other hand, one notices that

$$\begin{aligned} & f_{x_k}(x \mid \mathcal{I}_k^i) \\ &= \frac{\int_{\mathcal{Y}_k^i} f_{x_k}(x \mid y_k^i = y, \mathcal{I}_k^{i-1}) f_{y_k^i}(y \mid \mathcal{I}_k^{i-1}) dy}{\int_{\mathcal{Y}_k^i} f_{y_k^i}(y \mid \mathcal{I}_k^{i-1}) dy} \end{aligned} \quad (3.8)$$

$$= \int_{\mathbb{R}^m} f_{x_k}(x \mid y_k^i = y, \mathcal{I}_k^{i-1}) f_{y_k^i}(y \mid y \in \mathcal{Y}_k^i, \mathcal{I}_k^{i-1}) dy, \quad (3.9)$$

where $f_{y_k^i}(y \mid y \in \mathcal{Y}_k^i, \mathcal{I}_k^{i-1})$ satisfies $f_{y_k^i}(y \mid y \in \mathcal{Y}_k^i, \mathcal{I}_k^{i-1}) = 0$, $y \notin \mathcal{Y}_k^i$ and

$$\int_{\mathbb{R}^m} f_{y_k^i}(y \mid y \in \mathcal{Y}_k^i, \mathcal{I}_k^{i-1}) dy = 1,$$

and behaves similarly as the Dirac function $\delta(y)$, which equals to 0 except for $y = 0$ and satisfies $\int_{\mathbb{R}^m} \delta(y) dy = 1$. If point-valued measurements are always available, equation (3.9) becomes $\int_{\mathbb{R}^m} f_{x_k}(x|y_k^i = y, \mathcal{I}_k^{i-1}) \delta(y_k^i - y) dy = f_{x_k}(x|y_k^i = y, \mathcal{I}_k^{i-1})$, which maintains Gaussianity. Motivated by these observations, we introduce the following assumption:

Assumption 3.1. *The conditional distribution of x_k given \mathcal{I}_k^i can be approximated by a Gaussian distribution with the same mean and covariance.*

This assumption is also a commonly used technique in designing nonlinear Gaussian filters [7, 8, 33, 35]. To further illustrate the above assumption in the context of event-based estimation, we present the following numerical example.

Example 3.1. *Consider a linear system measured by one sensor and assume x_{k-1} is Gaussian with*

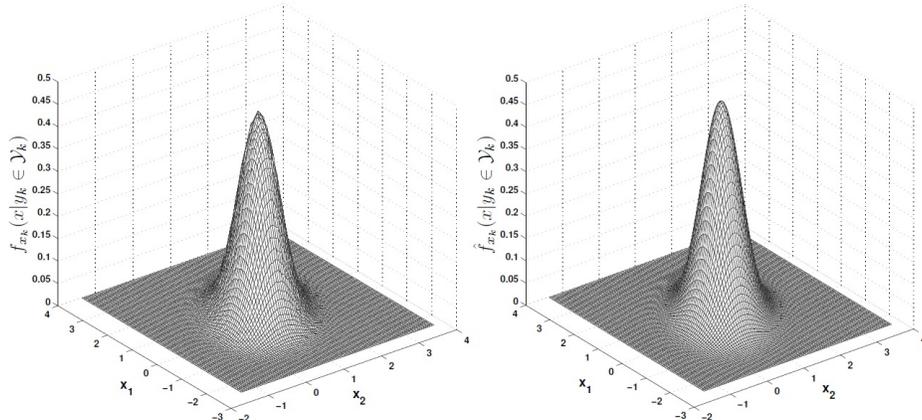
$$A = \begin{bmatrix} 1.5 & 0.7 \\ 0.8 & 1.6 \end{bmatrix}, Q = \begin{bmatrix} 0.5 & 0.1 \\ 0.1 & 0.6 \end{bmatrix}, \mathbf{Cov}(x_{k-1}) = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.8 \end{bmatrix},$$

$C = [1.2 \ 0.3]$, $R = 0.3$, and $\mathbf{E}(x_{k-1}) = [0.5 \ 0.5]^\top$ respectively. We study the distribution of x_k conditioned on set-valued measurement information. We perform Monte Carlo simulation and collect the realizations of x_k 's such that $y_k \in \mathcal{Y}_k := [CAE(x_{k-1}) - \delta, CAE(x_{k-1}) + \delta]$, and estimate the resulting distribution. Different values of δ are considered, and 10 million realizations of x_k satisfying $y_k \in \mathcal{Y}_k$ are used to estimate the conditional pdf $f_{x_k}(x|y_k \in \mathcal{Y}_k)$ for each δ . The pdf of Gaussian distributions $\hat{f}_{x_k}(x|y_k \in \mathcal{Y}_k)$ with equal first two moments are also included for comparison in the plots (see also the KL-divergences $D_{KL}(f||\hat{f})$ and $D_{KL}(\hat{f}||f)$ of the distributions). From Fig. 3.1, it is reasonable to approximate the conditional distributions as Gaussian distributions with acceptable approximation errors. \square

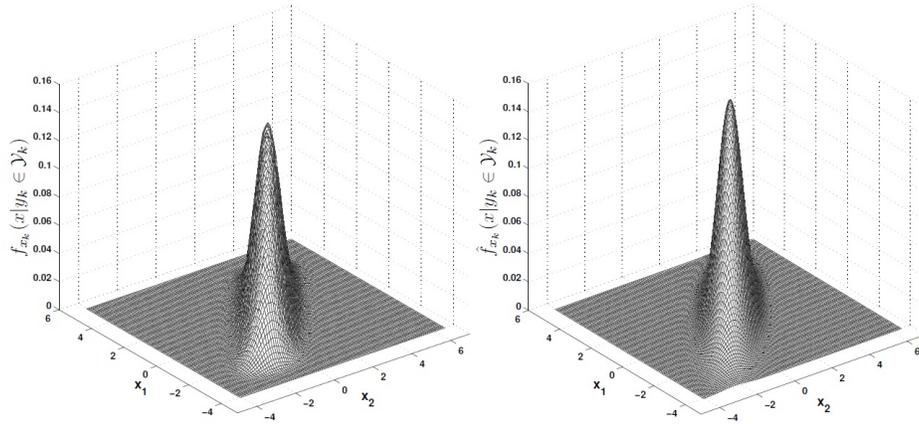
Now we are in a position to state the main problem considered in this work:

Problem 3.1. *At time k , given a sequence of measurement information $\{\mathcal{Y}_k^i | i \in \mathbb{N}_{1:M}\}$ of x_k and under Assumption 3.1, is it possible to find a simple approximate MMSE estimator in the recursive form? Does the exploration of the set-valued information lead to improved estimation performance in terms of estimation error covariance?*

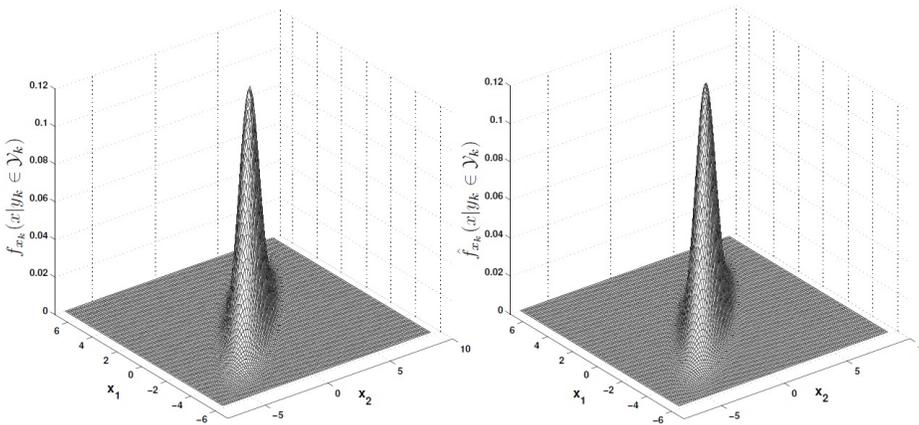
Meanwhile, since the exact MMSE estimate is the same for all fusion sequences under the Bayesian decision framework (by the uniqueness of the conditional distribution), when an approximate solution of a simple form is obtained, an additional



(a) $\delta = 1$. $D_{KL}(\hat{f}||f) = 0.0032$, $D_{KL}(f||\hat{f}) = 0.0033$.



(b) $\delta = 5$. $D_{KL}(\hat{f}||f) = 0.0267$, $D_{KL}(f||\hat{f}) = 0.0319$.



(c) $\delta = 10$. $D_{KL}(\hat{f}||f) = 0.0006$, $D_{KL}(f||\hat{f}) = 0.0004$.

Figure 3.1: Plot of the conditional distributions.

question to ask is whether the estimation performance is sensitive to the fusion sequence (due to the Gaussian assumption); we will further address this issue in the experimental verification section, where we test the performance of the proposed results extensively by Monte Carlo simulations.

3.3 Optimal fusion of sequential event-triggered measurement information

In this section, Problem 3.1 is studied in detail. Define $z_k^i = y_k^i - C^i \hat{x}_k^0$. Since \hat{x}_k^0 is known at time k by the estimator, this relationship maps the set Ξ_k^i to a unique set $\Omega_k^i := \{z_k^i : z_k^i = y_k^i - C^i \hat{x}_k^0, y_k^i \in \Xi_k^i\}$. Define $L_k^{i+1} := P_k^i (C^{i+1})^\top [C^{i+1} P_k^i (C^{i+1})^\top + R^{i+1}]^{-1}$, and $e_k^i := x_k - \hat{x}_k^i$. We have the following result:

Theorem 3.1.

1. *The optimal prediction \hat{x}_k^0 of the state x_k and the corresponding covariance P_k^0 are given by*

$$\begin{aligned}\hat{x}_k^0 &= A \hat{x}_{k-1}^M, \\ P_k^0 &= h(P_{k-1}^M).\end{aligned}$$

2. *For $i \in \mathbb{N}_{0:M-1}$, the fusion of information from the $(i+1)$ th sensor leads to the following recursive state estimation equations:*

- (a) *If $\gamma_k^{i+1} = 1$,*

$$\hat{x}_k^{i+1} = \hat{x}_k^i + L_k^{i+1} (z_k^{i+1} - \bar{z}_k^{i+1|i}), \quad (3.10)$$

$$P_k^{i+1} = \tilde{g}_{i+1}(P_k^i); \quad (3.11)$$

- (b) *If $\gamma_k^{i+1} = 0$,*

$$\hat{x}_k^{i+1} = \hat{x}_k^i + L_k^{i+1} (\bar{z}_k^{i+1|i+1} - \bar{z}_k^{i+1|i}), \quad (3.12)$$

$$P_k^{i+1} = \tilde{g}_{i+1}(P_k^i) + L_k^{i+1} \mathbf{Cov}(z_k^{i+1} | \mathcal{I}_k^{i+1}) (L_k^{i+1})^\top, \quad (3.13)$$

where $\bar{z}_k^{i+1|i} := C^{i+1}(\hat{x}_k^i - \hat{x}_k^0)$, and $\bar{z}_k^{i+1|i+1} := \mathbf{E}(z_k^{i+1} | \mathcal{I}_k^{i+1})$.

Proof. The first part of the result follows from Assumption 3.1. The proof of the second part is given in two steps.

(1) Proof of a few instrumental equalities:

$$\begin{aligned}
& \mathbf{E}[e_k^i(z_k^{i+1} - \bar{z}_k^{i+1|i})^\top | \mathcal{I}_k^{i+1}] \\
&= L_k^{i+1} \mathbf{E}[(z_k^{i+1} - \bar{z}_k^{i+1|i})(z_k^{i+1} - \bar{z}_k^{i+1|i})^\top | \mathcal{I}_k^{i+1}], \\
& \mathbf{E}[(e_k^i - L_k^{i+1}(z_k^{i+1} - \bar{z}_k^{i+1|i})) (z_k^{i+1} - \bar{z}_k^{i+1|i})^\top | \mathcal{I}_k^{i+1}] = 0, \\
& \mathbf{E}[(e_k^i - L_k^{i+1}(z_k^{i+1} - \bar{z}_k^{i+1|i})) \\
&\quad (e_k^i - L_k^{i+1}(z_k^{i+1} - \bar{z}_k^{i+1|i}))^\top | \mathcal{I}_k^i, z_k^{i+1} = z] = \tilde{g}_{i+1}(P_k^i), \\
& \mathbf{E}[(e_k^i - L_k^{i+1}(z_k^{i+1} - \bar{z}_k^{i+1|i})) \\
&\quad (e_k^i - L_k^{i+1}(z_k^{i+1} - \bar{z}_k^{i+1|i}))^\top | \mathcal{I}_k^{i+1}] = \tilde{g}_{i+1}(P_k^i).
\end{aligned}$$

Since $y_k^{i+1} = C^{i+1}x_k + v_k^{i+1}$, we have

$$\mathbf{E}(y_k^{i+1} | \mathcal{I}_k^i) = C^{i+1} \mathbf{E}(x_k | \mathcal{I}_k^i) = C^{i+1} \hat{x}_k^i. \quad (3.14)$$

$$\begin{aligned}
& \mathbf{Cov}[y_k^{i+1} | \mathcal{I}_k^i] \\
&= \mathbf{E}[(y_k^{i+1} - \mathbf{E}(y_k^{i+1} | \mathcal{I}_k^i))(y_k^{i+1} - \mathbf{E}(y_k^{i+1} | \mathcal{I}_k^i))^\top | \mathcal{I}_k^i], \\
&= \mathbf{E}[(C^{i+1}e_k^i + v_k^{i+1})(C^{i+1}e_k^i + v_k^{i+1})^\top | \mathcal{I}_k^i] \\
&= C^{i+1}P_k^i(C^{i+1})^\top + R^{i+1},
\end{aligned} \quad (3.15)$$

where $P_k^i = \mathbf{Cov}[x_k | \mathcal{I}_k^i]$. Since $z_k^{i+1} = y_k^{i+1} - C^{i+1}\hat{x}_k^0$,

$$\mathbf{E}(z_k^{i+1} | \mathcal{I}_k^i) = C^{i+1}\hat{x}_k^i - C^{i+1}\hat{x}_k^0. \quad (3.16)$$

$$\begin{aligned}
& \mathbf{Cov}[z_k^{i+1} | \mathcal{I}_k^i] \\
&= \mathbf{E}[(z_k^{i+1} - \mathbf{E}(z_k^{i+1} | \mathcal{I}_k^i))(z_k^{i+1} - \mathbf{E}(z_k^{i+1} | \mathcal{I}_k^i))^\top | \mathcal{I}_k^i], \\
&= \mathbf{E}[(C^{i+1}e_k^i + v_k^{i+1})(C^{i+1}e_k^i + v_k^{i+1})^\top | \mathcal{I}_k^i] \\
&= C^{i+1}P_k^i(C^{i+1})^\top + R^{i+1}.
\end{aligned} \quad (3.17)$$

Similarly, we have

$$\mathbf{Cov}[y_k^{i+1}x_k^\top | \mathcal{I}_k^i] = C^{i+1}P_k^i. \quad (3.18)$$

Thus

$$\mathbf{Cov}[x_k | \mathcal{I}_k^i, y_k^{i+1} = y] = \tilde{g}_{i+1}(P_k^i), \quad (3.19)$$

$$\mathbf{E}[x_k | \mathcal{I}_k^i, y_k^{i+1} = y] = \hat{x}_k^i + L_k^{i+1}(y_k^{i+1} - C^{i+1}\hat{x}_k^i). \quad (3.20)$$

Define $p_k^{i+1} := \mathbf{Pr}[z_k^{i+1} \in \Omega_k^{i+1} | \mathcal{I}_k^i] = \int_{z \in \Omega_k^{i+1}} f_{z_k^{i+1}}(z | \mathcal{I}_k^i) dz$. We have the conditional pdf

$$f_{z_k^{i+1}}(z | \mathcal{I}_k^{i+1}) = \begin{cases} f_{z_k^{i+1}}(z | \mathcal{I}_k^i) / p_k^{i+1}, & \text{if } z \in \Omega_k^{i+1}; \\ 0, & \text{otherwise.} \end{cases} \quad (3.21)$$

The rest of the proof follows from similar arguments as those below equation (27) of [73].

(2) Proof of the theorem: The case of $\gamma_k^{i+1} = 1$ follows from (3.19) and (3.20). Now we focus on the case of $\gamma_k^{i+1} = 0$. If the information provided by sensor $i + 1$ is given as a set \mathcal{Y}_k^{i+1} , \hat{x}_k^{i+1} should evolve according to

$$\begin{aligned} \hat{x}_k^{i+1} &= \mathbf{E}[x_k | \mathcal{I}_k^{i+1}] \\ &= \int_{z \in \Omega_k^{i+1}} \mathbf{E}[x_k | \mathcal{I}_k^i, z_k^{i+1} = z] f_{z_k^{i+1}}(z | \mathcal{I}_k^i) dz \Big/ p_k^{i+1} \\ &= \frac{1}{p_k^{i+1}} \int_{z \in \Omega_k^{i+1}} [\hat{x}_k^i + L_k^{i+1} z + L_k^{i+1} C^{i+1} (\hat{x}_k^0 - \hat{x}_k^i)] \\ &\quad f_{z_k^{i+1}}(z | \mathcal{I}_k^i) dz \\ &= \hat{x}_k^i - L_k^{i+1} \bar{z}_k^{i+1|i} + L_k^{i+1} \bar{z}_k^{i+1|i+1}, \end{aligned} \quad (3.22)$$

where $\bar{z}_k^{i+1|i+1} := \frac{1}{p_k^{i+1}} \int_{z \in \Omega_k^{i+1}} z f_{z_k^{i+1}}(z | \mathcal{I}_k^i) dz = \mathbf{E}(z_k^{i+1} | \mathcal{I}_k^{i+1})$. Finally we calculate the covariance of x_k conditioned on \mathcal{I}_k^{i+1} :

$$\begin{aligned} P_k^{i+1} &= \mathbf{E}[(x_k - \hat{x}_k^{i+1})(x_k - \hat{x}_k^{i+1})^\top | \mathcal{I}_k^{i+1}] \\ &= \mathbf{E}[(e_k^i - L_k^{i+1}(\bar{z}_k^{i+1|i+1} - \bar{z}_k^{i+1|i})) \\ &\quad (e_k^i - L_k^{i+1}(\bar{z}_k^{i+1|i+1} - \bar{z}_k^{i+1|i}))^\top | \mathcal{I}_k^{i+1}] \\ &= \tilde{g}_{i+1}(P_k^i) \\ &\quad + L_k^{i+1} \mathbf{E}[(z_k^{i+1} - \bar{z}_k^{i+1|i})(z_k^{i+1} - \bar{z}_k^{i+1|i})^\top | \mathcal{I}_k^{i+1}] (L_k^{i+1})^\top \\ &\quad - L_k^{i+1} (\bar{z}_k^{i+1|i+1} - \bar{z}_k^{i+1|i}) (\bar{z}_k^{i+1|i+1} - \bar{z}_k^{i+1|i})^\top (L_k^{i+1})^\top \end{aligned} \quad (3.23)$$

$$= \tilde{g}_{i+1}(P_k^i) + L_k^{i+1} \mathbf{Cov}(z_k^{i+1} | \mathcal{I}_k^{i+1}) (L_k^{i+1})^\top, \quad (3.24)$$

where equation (3.23) follows from the instrumental equalities as well as the equation

$$\begin{aligned} &\mathbf{E}[e_k^i | \mathcal{I}_k^{i+1}] \\ &= \mathbf{E}[x_k - \hat{x}_k^i | \mathcal{I}_k^{i+1}] \\ &= \int_{z \in \Omega_k^{i+1}} \mathbf{E}[x_k - \hat{x}_k^i | \mathcal{I}_k^i, z_k^i = z] f_{z_k^i}(z | \mathcal{I}_k^i) dz \Big/ p_k^{i+1} \\ &= \int_{z \in \Omega_k^{i+1}} L_k^{i+1} (z - \bar{z}_k^{i+1|i}) f_{z_k^i}(z | \mathcal{I}_k^i) dz \Big/ p_k^{i+1} \\ &= L_k^{i+1} (\bar{z}_k^{i+1|i+1} - \bar{z}_k^{i+1|i}), \end{aligned} \quad (3.25)$$

and equation (3.24) follows from the relation

$$\begin{aligned}
& \mathbf{E}[(z_k^{i+1} - \bar{z}_k^{i+1|i})(z_k^{i+1} - \bar{z}_k^{i+1|i})^\top | \mathcal{I}_k^{i+1}] \\
&= \mathbf{E}[(z_k^{i+1} - \bar{z}_k^{i+1|i+1}) + (\bar{z}_k^{i+1|i+1} - \bar{z}_k^{i+1|i}) \\
&\quad ((z_k^{i+1} - \bar{z}_k^{i+1|i+1}) + (\bar{z}_k^{i+1|i+1} - \bar{z}_k^{i+1|i}))^\top | \mathcal{I}_k^{i+1}] \\
&= \mathbf{Cov}[z_k^{i+1} | \mathcal{I}_k^{i+1}] \\
&\quad + (\bar{z}_k^{i+1|i+1} - \bar{z}_k^{i+1|i})(\bar{z}_k^{i+1|i+1} - \bar{z}_k^{i+1|i})^\top.
\end{aligned}$$

□

From the above result, the first and second moments of the truncated Gaussian distributions, namely, $\mathbf{E}(z_k^{i+1} | \mathcal{I}_k^{i+1})$ and $\mathbf{Cov}(z_k^{i+1} | \mathcal{I}_k^{i+1})$ need to be calculated to implement the event-based estimator. Fortunately, the moment evaluation problems of truncated Gaussian distributions have been extensively studied in the literature of statistical analysis; explicit formulae and efficient implementation methods have been proposed for a variety of truncation sets, see [42, 68, 69] and the references therein. Also, the estimate in (3.10) and (3.12) can be written in terms of the sum of series of random variables with Gaussian and non-Gaussian distributions. According to the asymptotic distribution theory for state estimate from a Kalman filter in the absence of Gaussian assumptions [5, 65], the central limit theorem for the estimates is still valid, which helps explain the rationality of Assumption 3.1.

The above result provides an acceptable answer to the first part of Problem 3.1; The second part of the problem, however, is difficult to answer for general event-triggering schemes. In the following, we consider $m = 1$, namely, when each sensor has only one channel. Notice that this scenario is equivalent to that the sensors have multiple channels, but each channel has uncorrelated measurement noise and separate event-triggering conditions, which is easy to implement in most prevailing embedded systems. Furthermore, the results can be equally applied to the case of multiple-channel sensors with correlated measurement noise but separate event-triggering conditions. To do this, it suffices to first transform each sensor measurement y_k^i to $\hat{y}_k^i = U^i y_k^i$ (where U^i is an orthogonal matrix satisfying $R^i = (U^i)^\top \Lambda^i U^i$, Λ^i being a diagonal matrix containing the eigenvalues of R^i), and then design the event-triggering conditions for each channel of \hat{y}_k^i .

When $m = 1$, without loss of generality, the event-triggering sets can be parameterized as $\Omega_k^i = \{z_k^i | a_k^i \leq z_k^i \leq b_k^i\}$, for $i \in \mathbb{N}_{1:M}$. For this type of sets, we have the following well known result [34].

Lemma 3.1. *For a univariate Gaussian random variable*

$$z_k^{i+1} | \mathcal{I}_k^i \sim \mathcal{N}(\bar{z}_k^{i+1|i}, Q_{z_k^{i+1}}),$$

its truncated mean and variance over $\Omega_k^{i+1} = \{z_k^{i+1} | a_k^{i+1} \leq z_k^{i+1} \leq b_k^{i+1}\}$ satisfy

$$\mathbf{E}(z_k^{i+1} | \mathcal{I}_k^{i+1}) = \bar{z}_k^{i+1|i} + \hat{z}_k^{i+1}, \quad (3.26)$$

$$\mathbf{Cov}[z_k^{i+1} | \mathcal{I}_k^{i+1}] = (1 - \vartheta_k^{i+1}) Q_{z_k^{i+1}}, \quad (3.27)$$

where $\phi(z) := \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}z^2)$,

$$\hat{z}_k^{i+1} = \frac{\phi\left(\frac{a_k^{i+1} - \bar{z}_k^{i+1|i}}{Q_{z_k^{i+1}}^{1/2}}\right) - \phi\left(\frac{b_k^{i+1} - \bar{z}_k^{i+1|i}}{Q_{z_k^{i+1}}^{1/2}}\right)}{Q\left(\frac{a_k^{i+1} - \bar{z}_k^{i+1|i}}{Q_{z_k^{i+1}}^{1/2}}\right) - Q\left(\frac{b_k^{i+1} - \bar{z}_k^{i+1|i}}{Q_{z_k^{i+1}}^{1/2}}\right)} Q_{z_k^{i+1}}^{1/2}, \quad (3.28)$$

$$\begin{aligned} \vartheta_k^{i+1} = & \left[\frac{\phi\left(\frac{a_k^{i+1} - \bar{z}_k^{i+1|i}}{Q_{z_k^{i+1}}^{1/2}}\right) - \phi\left(\frac{b_k^{i+1} - \bar{z}_k^{i+1|i}}{Q_{z_k^{i+1}}^{1/2}}\right)}{Q\left(\frac{a_k^{i+1} - \bar{z}_k^{i+1|i}}{Q_{z_k^{i+1}}^{1/2}}\right) - Q\left(\frac{b_k^{i+1} - \bar{z}_k^{i+1|i}}{Q_{z_k^{i+1}}^{1/2}}\right)} \right]^2 - \\ & \frac{\frac{a_k^{i+1} - \bar{z}_k^{i+1|i}}{Q_{z_k^{i+1}}^{1/2}} \phi\left(\frac{a_k^{i+1} - \bar{z}_k^{i+1|i}}{Q_{z_k^{i+1}}^{1/2}}\right) - \frac{b_k^{i+1} - \bar{z}_k^{i+1|i}}{Q_{z_k^{i+1}}^{1/2}} \phi\left(\frac{b_k^{i+1} - \bar{z}_k^{i+1|i}}{Q_{z_k^{i+1}}^{1/2}}\right)}{Q\left(\frac{a_k^{i+1} - \bar{z}_k^{i+1|i}}{Q_{z_k^{i+1}}^{1/2}}\right) - Q\left(\frac{b_k^{i+1} - \bar{z}_k^{i+1|i}}{Q_{z_k^{i+1}}^{1/2}}\right)}, \end{aligned} \quad (3.29)$$

$Q(\cdot)$ denotes the standard Q -function.

Based on this result, we will show that the optimal estimate subject to a given sequence of measurement information reduces to a simple closed form, and that the exploration of set-valued information could lead to guaranteed enhanced performance. To do this, we introduce the following lemmas.

Lemma 3.2 (Theorem 2 in [12]). *Let z be an absolutely continuous random variable with cumulative distribution function $F(z)$. The conditional variance $\mathbf{Cov}(z | a \leq z \leq b)$ is increasing in b if and only if*

$$\int_{a \leq z_1 \leq z_2 \leq b} \{F(z_1) - F(a)\} dz_1 dz_2 \quad (3.30)$$

is log-concave in b , and it is decreasing in a if and only if

$$\int_{a \leq z_1 \leq z_2 \leq b} \{F(b) - F(z_1)\} dz_1 dz_2 \quad (3.31)$$

is log-concave in a . When both conditions in (3.30) and (3.31) are satisfied for all $a, b \in \mathcal{C}$ for some convex set \mathcal{C} , then $\mathbf{Cov}(z|z \in \mathcal{A})$ is partially monotonic in an interval \mathcal{A} such that $\mathcal{A} \subset \mathcal{C}$.

Lemma 3.3 (Lemma 1 in [12]). *If a function $f(z)$ is log-concave for $z \in (a, b)$, then the antiderivative $F(x) = \int_a^z f(t)dt$ is also log-concave for $z \in (a, b)$ whenever it is well defined.*

Now we are ready to present the following result.

Theorem 3.2.

1. The optimal prediction \hat{x}_k^0 of the state x_k and the corresponding covariance P_k^0 are given by

$$\begin{aligned} \hat{x}_k^0 &= A\hat{x}_{k-1}^M, \\ P_k^0 &= h(P_{k-1}^M). \end{aligned} \quad (3.32)$$

2. For $i \in \mathbb{N}_{0:M-1}$, the fusion of information from the $(i+1)$ th sensor leads to the following recursive state estimation equations:

- (a) If $\gamma_k^{i+1} = 1$,

$$\hat{x}_k^{i+1} = \hat{x}_k^i + L_k^{i+1}(z_k^{i+1} - \bar{z}_k^{i+1|i}), \quad (3.33)$$

$$P_k^{i+1} = \tilde{g}_{s_{i+1}}(P_k^i). \quad (3.34)$$

- (b) If $\gamma_k^{i+1} = 0$,

$$\hat{x}_k^{i+1} = \hat{x}_k^i + L_k^{i+1}\hat{z}_k^{i+1}, \quad (3.35)$$

$$P_k^{i+1} = \tilde{g}_{s_{i+1}}(P_k^i, \vartheta_k^{i+1}), \quad (3.36)$$

where \hat{z}_k^{i+1} is given in (3.28), and ϑ_k^{i+1} is given in (3.29) and in particular, satisfies $\vartheta_k^{i+1} \in (0, 1)$.

Proof. It suffices to prove equations (3.35) and (3.36). Equation (3.35) follows from (3.26) and (3.12). From (3.13),

$$\begin{aligned} P_k^{i+1} &= \tilde{g}_{s_{i+1}}(P_k^i) + L_k^{i+1}\mathbf{Cov}(z_k^{i+1}|\mathcal{I}_k^{i+1})(L_k^{i+1})^\top \\ &= \tilde{g}_{s_{i+1}}(P_k^i) + (1 - \vartheta_k^{i+1})L_k^{i+1} \\ &\quad [C^{i+1}P_k^i(C^{i+1})^\top + R^{i+1}](L_k^{i+1})^\top \\ &= \tilde{g}_{s_{i+1}}(P_k^i, \vartheta_k^{i+1}). \end{aligned}$$

Finally we show $\vartheta_k^{i+1} \in (0, 1)$. Since $\mathbf{Cov}[z_k^{i+1}|\mathcal{I}_k^{i+1}] > 0$, we have $\vartheta_k^{i+1} < 1$. We consider the case $\bar{z}_k^{i+1|i} \in [a_k^{i+1}, b_k^{i+1}]$. In this case, $(a_k^{i+1} - \bar{z}_k^{i+1|i})/Q_{z_k^{i+1}}^{1/2} \leq 0$ and $(b_k^{i+1} - \bar{z}_k^{i+1|i})/Q_{z_k^{i+1}}^{1/2} \geq 0$ hold. From (3.29), we have $\vartheta_k^{i+1} > 0$. This implies that a pair (a_k^{i+1}, b_k^{i+1}) such that $a_k^{i+1} \leq \bar{z}_k^{i+1|i} \leq b_k^{i+1}$ will lead to $\mathbf{Cov}[z_k^{i+1}|\mathcal{I}_k^{i+1}] < Q_{z_k^{i+1}}$. Now consider the case that $\bar{z}_k^{i+1|i} \notin [a_k^{i+1}, b_k^{i+1}]$. There always exists a pair $(\underline{a}_k^{i+1}, \bar{b}_k^{i+1})$ such that $[a_k^{i+1}, b_k^{i+1}] \subset [\underline{a}_k^{i+1}, \bar{b}_k^{i+1}]$ and $\bar{z}_k^{i+1|i} \in [\underline{a}_k^{i+1}, \bar{b}_k^{i+1}]$. Since $\phi(z)$ is a logarithmically concave function, from Lemma 3.2 and Lemma 3.3, we have $\mathbf{Cov}[z_k^{i+1}|\mathcal{I}_k^{i+1}] \leq \mathbf{Cov}[z_k^{i+1}|\mathcal{I}_k^i, z_k^{i+1} \in [\underline{a}_k^{i+1}, \bar{b}_k^{i+1}]] < Q_{z_k^{i+1}}$. Thus we have $\vartheta_k^{i+1} > 0$, which completes the proof. \square

Since $\vartheta_k^i \in (0, 1)$ is guaranteed when $\gamma_k^i = 0$, smaller estimation error covariance can be obtained by exploiting the set-valued measurement information, which implies improved estimation performance. Also, we know that for a given sensor information sequence s , the resultant optimal estimate evolves according to (3.33) and (3.35). The calculation of ϑ_k^i mainly requires the calculation of the standard Q -functions, which is easy to implement. Therefore, theoretically, the derived event-based estimator enjoys both potentially improved performance and a simple closed form with low computational complexity. The actual effectiveness of the estimator will be further verified in the following section.

3.4 Experimental verification of the proposed results based on Monte Carlo simulations

In this section, we test the efficiency of the proposed results by Monte Carlo simulation. Specifically, we consider the practical ‘‘send on delta’’ communication strategy [44], namely, at time k , sensor i decides whether to send new measurement updates to the remote estimator according to the following condition:

$$\gamma_k^i = \begin{cases} 1 & \text{if } |y_k^i - y_{\tau_k^i}^i| \geq \delta^i, \\ 0 & \text{otherwise,} \end{cases} \quad (3.37)$$

where τ_k^i denotes the last instance when the measurement of sensor i is transmitted. To study the applicability of the results, we consider three categories of systems:

1. Category 1: $\text{trace}\{Q\}/n \gg \text{trace}\{R^i\}/m$.
2. Category 2: $\text{trace}\{Q\}/n \sim \text{trace}\{R^i\}/m$.

3. Category 3: $\text{trace}\{Q\}/n \ll \text{trace}\{R^i\}/m$.

For each category, we randomly generate 1,000 third-order stable discrete-time systems², the eigenvalues of which lie uniformly in $[-0.95, 0.95]$, and measure each system by 5 sensors with $m = 1$ and randomly generated parameters³. For each system, we perform the simulation for 1,000 time instants, and evaluate the performance of the proposed event-based estimator from two aspects:

(1) To study the possible performance improvement induced by exploring the set-valued information, comparison is made with the Kalman filter with intermittent observations exploring only the received point-valued measurement information. To quantify the performance difference, the estimation errors are normalized by the averaged norm of the original state:

$$\Delta_E := \frac{e_K - e_E}{\sqrt{\sum_{t=1}^{1000} \|x_t\|^2 / 1000}}, \quad (3.38)$$

where e_K denotes the root average squared estimation error of the Kalman filter with intermittent observations, e_E denotes the root average squared estimation error of the proposed event-based estimator, and x_t denotes the random generated state trajectory of the system. The distributions of Δ_E 's for different categories as well as the corresponding average communication rates⁴ are plotted in Fig. 3.2. From this figure, it is observed that the proposed event-based estimator obtained almost guaranteed improved performance compared with the Kalman filter with intermittent observations, indicating the efficient exploitation of the set-valued information. The only few cases that the event-based estimator slightly deteriorates the estimation performance belong to Category 1 (see Fig. 3.2(a)), and from Fig. 3.2(b), it is observed that these cases have very low communication rates, which correspond to large δ^i 's; intuitively, the Gaussian assumptions sometimes may not be accurate enough to provide effective description of the *a priori* distributions for this case, thus resulting in less effective estimates.

(2) To test the sensitivity of the estimation performance to sensor fusion sequences, comparison is made between the estimates that are obtained according to different se-

²We do not consider unstable eigenvalues here to avoid errors introduced by the unbounded state trajectories.

³The Q and R^i matrices are obtained by first enumerating a set of positive real numbers satisfying the same uniform distributions, and then decreasing (increasing) those corresponding to R^i 's by one magnitude for Category 1 (Category 3); the δ^i 's are also randomly generated positive real numbers to allow for different communication rates.

⁴The average communication rates are calculated as $\frac{1}{5 \cdot 1000} \sum_{i=1}^5 \sum_{k=1}^{1000} \gamma_k^i$, which are nonnegative by definition.

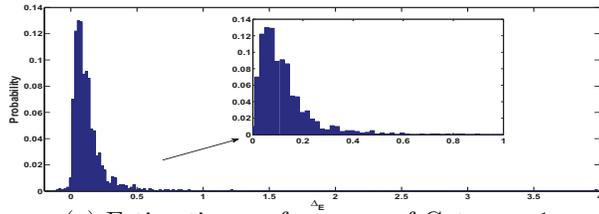
quences of sensors. The first one is obtained by the sequences that minimize the estimation error variances at each time instant, while the second one is obtained by sequences that maximize the estimation error variance at each time instant. To quantify the performance difference, define the normalized performance difference as

$$\Delta_F := \frac{e_W - e_B}{e_B}, \quad (3.39)$$

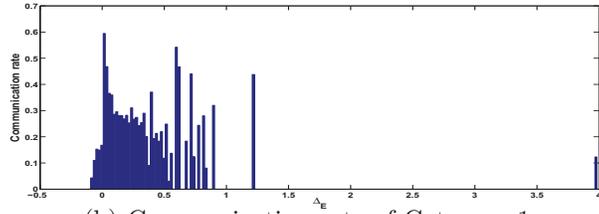
where e_B and e_W denote the root average squared estimation errors of the fusion sequences obtained by minimizing and maximizing the error variance, respectively. The distribution of Δ_F 's and the corresponding communication rates are shown in Fig. 3.3. It is observed that the difference is always relatively small, and becomes smaller as the system becomes more measurement-noise dominant. Since the difference should be zero for the MMSE estimate without the Gaussian assumption, the results indicate that the proposed estimator represents the exact MMSE estimator to a satisfactory extent.

3.5 Summary

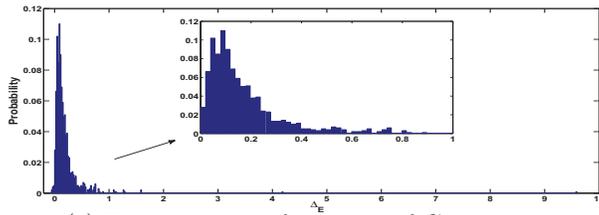
In this chapter, the problem of optimal fusion of hybrid measurement information for event-based estimation is studied. For a fixed sensor sequence, we show that the optimal MMSE estimate depends on the conditional mean and variance of the innovations. When each sensor has only one channel, a closed-form representation for the MMSE estimate is developed, and it is proved that exploring the set-valued information always improves estimation performance. The results are equally applicable to multiple-channel sensors with separate event-triggering conditions. Extensive simulation results show that the proposed estimator provides improved performance for most cases and is not sensitive to the fusion sequence.



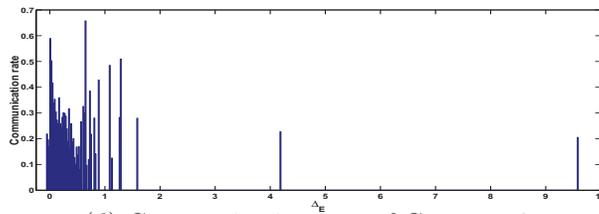
(a) Estimation performance of Category 1



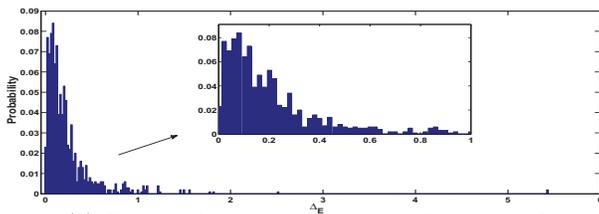
(b) Communication rate of Category 1



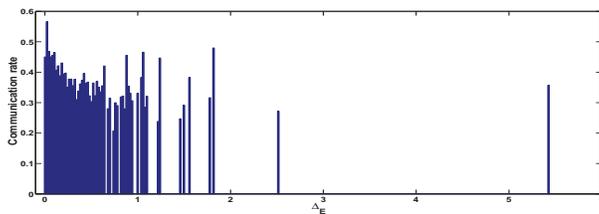
(c) Estimation performance of Category 2



(d) Communication rate of Category 2

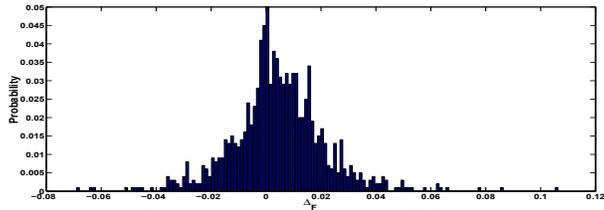


(e) Estimation performance of Category 3

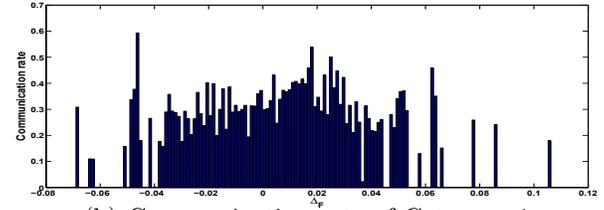


(f) Communication rate of Category 3

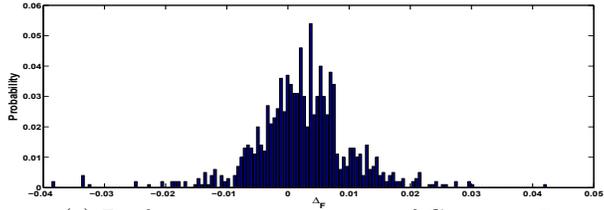
Figure 3.2: Performance validation of the proposed event-based estimator.



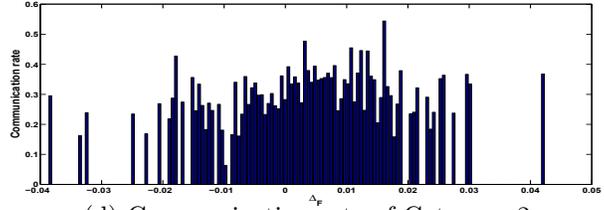
(a) Performance comparison of Category 1



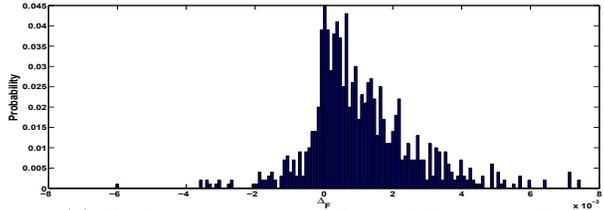
(b) Communication rate of Category 1



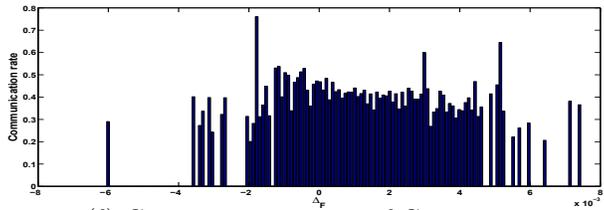
(c) Performance comparison of Category 2



(d) Communication rate of Category 2



(e) Performance comparison of Category 3



(f) Communication rate of Category 3

Figure 3.3: Performance comparison between different fusion sequences.

Chapter 4

On set-valued Kalman filtering and its application to event-based state estimation*

4.1 Introduction

The event-based estimation problem can be considered as a set-valued Kalman filtering problem in the convex Bayesian decision framework [66], which takes the difference and separation between “stochastic uncertainty” and “non-stochastic uncertainty” into account. In [47], it was pointed out that “Ignorance, in its root meaning, means lack of knowledge; uncertainty, on the other hand, typically means lack of precision. By specifying a probability distribution for a random variable, we attempt to characterize uncertainty. If the correct distribution function is unknown, that is a manifestation of ignorance.” From this perspective, the statistical information of the noise and the initial states are regarded as “uncertainty” (or “stochastic uncertainty”), while the ambiguous information contained in the event-triggering sets can be considered as “ignorance” (or “non-stochastic uncertainty”), since the estimator’s inability of knowing the point-valued measurement information during non-event time instants can be regarded as “lack of knowledge”, which is caused by the subjective choice of the event-triggering conditions.

Compared with the existing results in event-based state estimation, the set-valued filtering approach provides an alternative way of exploiting and understanding the

*Parts of the results in this chapter have been conditionally accepted by IEEE Transactions on Automatic Control as a full paper, 2014.

additional information contained in the event-triggering conditions. The set-valued Kalman filter was originally introduced by [47], where the standard Kalman filter was extended to the case that a convex set of initial estimate distributions was considered. Although the filters bear good asymptotic properties, they are not applicable to the event-based estimation scenario, since only point-valued measurements are considered. Recently, further relaxation of the assumptions on uniqueness for the *a posteriori* probability distributions was considered in [51] by allowing set-valued measurements and the multiple sensor fusion problem was considered in [52] utilizing the information filter approach. When the set-valued measurements are treated as non-stochastic uncertainty, the choice of different points in the measurement set at each time instant only leads to different values of the estimation mean (which we refer to as “the set of estimation means” hereafter), while the estimation error covariance remains unaffected. These results allow set-valued event-based estimators to be designed; however, several problems remain unexplored with respect to these new set-valued filters, which are of fundamental importance for the study of event-based estimation:

1. For multiple point-valued measurements, the performance is quantified only in terms of the estimation error covariance, and it is known that the fusion sequence used to update the sensor measurement information at the same time instant does not affect the resultant centralized Kalman filter. For set-valued Kalman filters, the overall performance is measured in two terms, the estimation error covariance and the size of the set of estimation means (e.g., the size of an ellipsoidal set can be quantified by the trace of a positive semidefinite matrix defining the shape of the set in this work). Apparently, the fusion sequence still does not affect the error covariance, but its effect on the size of estimation means is not known.
2. In [47], it was shown that the set of estimation means converges towards a singleton as time goes to infinity for point-valued measurements. However, the asymptotic behavior of the size of the set of estimation means is not clear when set-valued measurements are considered. In addition, due to the properties of the Minkowski sum of ellipsoids, the set of estimation means can only be calculated approximately [51]; in this regard, the asymptotic property of the size of the approximate set of estimation means is of importance as well.
3. In set-valued Kalman filtering, the performance of the estimator is measured by not only the estimation error covariance that quantifies stochastic uncertainty, but the size of the set of estimation means quantifying non-stochastic uncertainty

as well. For standard Kalman filters, it is known that increasing the number of sensors can reduce the estimation error covariance; this result is still valid for the set-valued case. The effect of adding more sensors on the size of the set of estimation means is, however, still unknown.

In this chapter, we seek to explore the above problems for linear time-invariant systems with an emphasis on event-based estimation. The estimation problem is considered in the multiple sensor scenario, where each sensor is allowed to provide its own set-valued measurement parameterized by ellipsoids. The main contributions are summarized as follows.

1. The exact set of estimation means is shown to be invariant with respect to the fusion sequences. Since the exact set-valued filter is normally not implementable, a two-step approximate set-valued estimator is proposed and is shown to be unaffected by the fusion sequences. The approximate estimator proposed here is different from that in [52], which was given in the information filtering form.
2. The boundedness of the size of the set of estimation means for the exact set-valued filter is proved. For the approximate estimator, we show that if the closed-loop matrix is contractive at steady state, then the boundedness of the size of the set of estimation means is guaranteed; otherwise, there exists an invertible linear transformation such that the size of the set of estimation means of the approximate estimator after the transformation is bounded.
3. An upper bound on the steady-state performance in terms of the size of the set of estimation means is proposed, and conditions are characterized for smaller size of the set of estimation means. For scalar systems, a sufficient condition is provided for guaranteed performance improvement. Based on the developed results, an optimal event-triggering condition design problem is further formulated and solved.

Notation: We use *Italic* letters to denote vector- or matrix-valued variables, and use calligraphic letters to denote sets (except for the sets of real numbers and integers). \mathbb{R} denotes the set of real numbers. \mathbb{N} denotes the set of nonnegative integers. \mathbb{N}^+ denotes the set of positive integers. Let $m, n \in \mathbb{N}^+$; $\mathbb{R}^{m \times n}$ denotes the set of m by n real-valued matrices. For brevity, denote $\mathbb{R}^m := \mathbb{R}^{m \times 1}$. For $v \in \mathbb{R}^m$, let $\|v\|$ denote its Euclidean norm. For $Z \in \mathbb{R}^{m \times n}$, Z^\top denotes the transpose of Z , and $\|Z\|_2$ denotes the spectral norm of Z . The symbol I denotes the identity matrix with a context-dependent size.

For $X, Y \in \mathbb{R}^{n \times n}$, $X > (\geq) Y$ means $X - Y$ is positive definite (positive semidefinite). For two convex sets $\mathcal{X}, \mathcal{Y} \subseteq \mathbb{R}^n$, let $\mathcal{X} \oplus \mathcal{Y}$ denote their Minkowski sum, namely, $\mathcal{X} \oplus \mathcal{Y} := \{x + y | x \in \mathcal{X}, y \in \mathcal{Y}\}$. Also, $\bigoplus_{i=1}^n \mathcal{X}_i := \mathcal{X}_1 \oplus \mathcal{X}_2 \oplus \dots \oplus \mathcal{X}_n$. For $T \in \mathbb{R}^{m \times n}$ and $\mathcal{X} \subseteq \mathbb{R}^n$, define $T\mathcal{X}$ as

$$T\mathcal{X} := \{Tx \in \mathbb{R}^m | x \in \mathcal{X}\}.$$

Given $Y > 0$, an ellipsoidal set (or an ellipsoid) $\mathcal{Y} = \mathcal{E}(c, Y)$ in \mathbb{R}^m is defined as

$$\begin{aligned} \mathcal{Y} &:= \mathcal{E}(c, Y) \\ &= \{y \in \mathbb{R}^m \mid (y - c)^\top (Y)^{-1} (y - c) \leq 1, Y > 0\}; \end{aligned}$$

if Y is singular and $Y \geq 0$, \mathcal{Y} is parameterized as¹

$$\mathcal{Y} = \{y \in \mathbb{R}^m \mid \langle l, y \rangle \leq \langle l, c \rangle + \langle l, Yl \rangle^{1/2}, \forall l \in \mathbb{R}^m\}.$$

In this work, we define the size of an ellipsoidal set \mathcal{Y} as $\text{Tr} Y$, and we say set \mathcal{Y} has a bounded size if $\text{Tr} Y$ is bounded².

Let $m, n, p, q \in \mathbb{N}$ satisfying $m \leq n$ and $p \leq q$; $\mathbb{N}_{m:n}$ denotes the set of integers $\{m, \dots, n\}$; letting $\{s_i \in \mathbb{N} | i \in \mathbb{N}_{1:r}, r \in \mathbb{N}^+\}$ be an indexed set of integers, $y^{s_{m:n}}$ denotes the set $\{y^{s_m}, \dots, y^{s_n}\}$, and $y_{p:q}^{s_{m:n}}$ denotes the set $\{y_p^{s_{m:n}}, \dots, y_q^{s_{m:n}}\}$; similarly, $y^{s_{m:n}} \in \mathcal{Y}^{s_{m:n}}$ denotes the relationship $y^{s_m} \in \mathcal{Y}^{s_m}, \dots, y^{s_n} \in \mathcal{Y}^{s_n}$, and $y_{p:q}^{s_{m:n}} \in \mathcal{Y}_{p:q}^{s_{m:n}}$ denotes the relationship

$$y_p^{s_{m:n}} \in \mathcal{Y}_p^{s_{m:n}}, \dots, y_q^{s_{m:n}} \in \mathcal{Y}_q^{s_{m:n}}.$$

For a vector-valued random variable x , we use $\mathbf{E}(x)$ and $\mathbf{Cov}(x)$ to denote its mean and covariance, respectively.

4.2 Problem setup

The process is linear time-invariant and evolves in discrete time driven by white noise:

$$x_{k+1} = Ax_k + w_k, \tag{4.1}$$

where $x \in \mathbb{R}^n$ is the state, and $w \in \mathbb{R}^n$ is the noise input, which is zero-mean Gaussian with covariance $Q \geq 0$. We assume (A, Q) is stabilizable³. The initial value x_0 of

¹Note that the way of parameterizing an ellipsoidal set does not affect the results developed in this work.

²Notice that based on this definition, the boundedness of the size of \mathcal{Y} is independent of its centre $c \in \mathbb{R}^m$, since c only describes the relative position of \mathcal{Y} . Normally the size of an ellipsoid is given by the maximal eigenvalue of Y . In terms of boundedness, however, these two definitions are equivalent.

³Note that this is equivalent to the stabilizability of the pair (A, \sqrt{Q}) , which can be proved based on the PBH criteria.

the state is also zero-mean Gaussian, with covariance P_0 . The state information is measured using M different sensors, the measurement equations of which are

$$y_k^i = C_i x_k + v_k^i, \quad (4.2)$$

where $y^i \in \mathbb{R}^m$ denotes the output of the i th sensor, $v^i \in \mathbb{R}^m$ is zero-mean Gaussian with covariance R_i for $i \in \mathbb{N}_{1:M}$, and v^i and v^j are uncorrelated if $i \neq j$. In addition, x_0 , w and v^i are uncorrelated with each other. We assume (A, C) is detectable, where $C := [C_1^\top, \dots, C_M^\top]^\top$; define $R := \text{diag}\{R_1, R_2, \dots, R_M\}$.

We consider the scenario that the values of the measurement outputs y_k^i are not exactly known, but are only partially known in the sense that only the exact description of sets \mathcal{Y}_k^i is known such that $y_k^i \in \mathcal{Y}_k^i$ for all $i \in \mathbb{N}_{1:M}$. To some extent, this reflects the estimator's inability of telling a point measurement from an uncountable set of measurements, due to the lack of knowledge, e.g., the situation the remote estimator is facing during the non-event instances in an event-based estimation scenario [73]. As a result, the uniqueness of the posteriori probability distributions cannot be maintained, which gives rise to the set-valued Kalman filters [47]. Due to the set-valued measurements from the M sensors at each time instant, one feasible way to update the state estimate is to fuse the measurement information from the sensors sequentially piece by piece according to some sequence, which (can be chosen either arbitrarily or by design) is mathematically given as

$$s = [s_1, s_2, \dots, s_M],$$

where $s_i \in \mathbb{N}_{1:M}$ and $s_i \neq s_j$ unless $i = j$, for $i, j \in \mathbb{N}_{1:M}$. We refer to this sequence as ‘‘fusion sequence’’ in this work. Note that in a fusion sequence, each sensor appears once and only once, and the sequence is used to update the information from different sensors measured at the same time instant and does not affect the sensor measurement information in this work.

In standard Kalman filtering, the optimal state prediction $\tilde{x}_k^{s_0}$ that minimizes the estimation error covariance at time instant k is known to satisfy

$$\tilde{x}_k^{s_0} = \mathbf{E}(x_k | y_0^{s_{1:M}}, y_1^{s_{1:M}}, \dots, y_{k-1}^{s_{1:M}}), \quad (4.3)$$

where the superscript s_0 is used to indicate that no sensor information measured at time k has been updated; similarly, for $i \in \mathbb{N}_{1:M}$, the optimal state estimate $\tilde{x}_k^{s_i}$ after updating the measurement information from sensors s_1, s_2, \dots, s_i at time k satisfies

$$\tilde{x}_k^{s_i} = \mathbf{E}(x_k | y_0^{s_{1:M}}, \dots, y_{k-1}^{s_{1:M}}, y_k^{s_{1:i}}). \quad (4.4)$$

The corresponding estimation error covariance satisfies

$$\begin{aligned} P_k^{s_0} &:= \mathbf{Cov}(x_k | y_0^{s_{1:M}}, y_1^{s_{1:M}}, \dots, y_{k-1}^{s_{1:M}}) \\ &= AP_{k-1}^{s_M} A^\top + Q, \end{aligned} \quad (4.5)$$

$$\begin{aligned} P_k^{s_i} &:= \mathbf{Cov}(x_k | y_0^{s_{1:M}}, y_1^{s_{1:M}}, \dots, y_{k-1}^{s_{1:M}}, y_k^{s_{1:i}}) \\ &= P_k^{s_{i-1}} - P_k^{s_{i-1}} C_{s_i}^\top (C_{s_i} P_k^{s_{i-1}} C_{s_i}^\top + R_{s_i})^{-1} C_{s_i} P_k^{s_{i-1}} \end{aligned} \quad (4.6)$$

for $i \in \mathbb{N}_{1:M}$. In set-valued filtering [51, 52], the set-valued measurements are treated as non-stochastic uncertainty; as a result, the choice of different points in the measurement set only leads to different values of the estimation mean, while the estimation error covariance remains unaffected. Specifically, the set of estimation means is defined as

$$\mathcal{X}_k^{s_0} := \{ \mathbf{E}(x_k | y_{0:k-1}^{s_{1:M}}) | y_{0:k-1}^{s_{1:M}} \in \mathcal{Y}_{0:k-1}^{s_{1:M}} \}, \quad (4.7)$$

$$\mathcal{X}_k^{s_i} := \{ \mathbf{E}(x_k | y_{0:k-1}^{s_{1:M}}, y_k^{s_{1:i}}) | y_{0:k-1}^{s_{1:M}} \in \mathcal{Y}_{0:k-1}^{s_{1:M}}, y_k^{s_{1:i}} \in \mathcal{Y}_k^{s_{1:i}} \}, \quad (4.8)$$

for $i \in \mathbb{N}_{1:M}$, where $\mathcal{X}_k^{s_0}$ denotes the set of estimation means when no sensor information is fused at time k (namely, the prediction of the state), and for $i \in \mathbb{N}_{1:k}$, $\mathcal{X}_k^{s_i}$ denotes the set of estimation means after fusing the information of sensor s_1, s_2, \dots, s_i at time instant k . We assume $\mathcal{X}_0^{s_0} = \{0\}$, following the zero-mean Gaussian assumption of x_0 . The definition of estimation error covariance still follows that of the standard Kalman filters, which has been given in (4.5) and (4.6). In light of the results in [47], [51, 52], the exact set-valued Kalman filter with multiple sensor measurements is recursively given as

$$\mathcal{X}_k^{s_0} = A\mathcal{X}_{k-1}^{s_M}, \quad (4.9)$$

$$P_k^{s_0} = AP_{k-1}^{s_M} A^\top + Q, \quad (4.10)$$

and for $i \in \mathbb{N}_{0:M-1}$,

$$\mathcal{X}_k^{s_{i+1}} = (I - K_k^{s_{i+1}} C_{s_{i+1}}) \mathcal{X}_k^{s_i} \oplus K_k^{s_{i+1}} \mathcal{Y}_k^{s_{i+1}}, \quad (4.11)$$

where

$$K_k^{s_{i+1}} = P_k^{s_i} C_{s_{i+1}}^\top (C_{s_{i+1}} P_k^{s_i} C_{s_{i+1}}^\top + R_{s_{i+1}})^{-1},$$

$$P_k^{s_{i+1}} = P_k^{s_i} - P_k^{s_i} C_{s_{i+1}}^\top (C_{s_{i+1}} P_k^{s_i} C_{s_{i+1}}^\top + R_{s_{i+1}})^{-1} C_{s_{i+1}} P_k^{s_i}. \quad (4.12)$$

In Kalman filtering, the confidence on the estimate is fully characterized by the estimation error covariance; while in set-valued Kalman filtering, a set of probability

density functions with the same covariance is considered, and the unknown information caused by stochastic uncertainty and non-stochastic uncertainty is treated separately: the confidence on stochastic uncertainty is still quantified as covariance (equations (4.10) and (4.12)); the confidence on non-stochastic uncertainty is quantified as the size of the set of estimation means (equations (4.9) and (4.11)). As will be shown in this work, this separation can help provide new insights for event-based estimation problems.

In this work, we assume \mathcal{Y}_k^i are ellipsoidal sets parameterized as

$$\mathcal{Y}_k^i := \mathcal{E}(c_k^i, Y_k^i). \quad (4.13)$$

Notice that the parameters of \mathcal{Y}_k^i can be acknowledged by the estimator without communication from the sensor. For instance, in an event-based estimation scenario with the aforementioned “send-on-delta” triggering conditions in Section 4.1, c_k^i is the previously transmitted measurement, while Y_k^i is normally defined offline and thus can be known to the estimator beforehand. In the literature, there are alternative ways of describing set-valued measurements, e.g., in terms of parallelotopes and zonotopes [3, 14]. The properties of the resultant estimates are, however, difficult to characterize, due to the lack of intuitive mathematical description of the notion “sizes of the sets”. Although the Minkowski sum of ellipsoids (which may not be an ellipsoid) is difficult to calculate exactly [37], the ellipsoidal sets are very helpful in analysing the dynamic behavior of the estimates, since the size and shape of an ellipsoid are uniquely determined by a positive semidefinite matrix. At the same time, outer ellipsoidal approximations are conveniently employed to calculate the set that contains the set of means of the estimates at each time instant, which is calculated according to the following result.

Lemma 4.1 ([37]). *Let $p > 0$. We have*

$$\begin{aligned} & \mathcal{E}(c_1, X_1) \oplus \mathcal{E}(c_2, X_2) \\ & \subseteq \mathcal{E}(c_1 + c_2, (1 + p^{-1})X_1 + (1 + p)X_2). \end{aligned} \quad (4.14)$$

Normally p is calculated in some optimal sense. In this work, we take $p = \frac{(\text{Tr } X_1)^{1/2}}{(\text{Tr } X_2)^{1/2}}$, which minimizes the trace of $(1 + p^{-1})X_1 + (1 + p)X_2$. In this way, we are able to evaluate the outer ellipsoidal approximate estimates

$$\hat{\mathcal{X}}_k^{s_0} := \mathcal{E}(\hat{x}_k^{s_0}, X_k) \supseteq \mathcal{X}_k^{s_0} \quad (4.15)$$

of $\mathcal{X}_k^{s_0}$ according to Lemma 4.1 and equations (4.9)-(4.11), which will be formally introduced in the next section to take account of the effect from sensor fusion sequence.

Based on the introduced notations, we are now in the position to present the problems to be considered in this work:

1. Analyze the effect of the sensor fusion sequence s on the exact and approximate sets of estimation means $\mathcal{X}_k^{s_0}$ and $\hat{\mathcal{X}}_k^{s_0}$, respectively;
2. Analyze the asymptotic behavior of the sizes of $\mathcal{X}_k^{s_0}$ and $\hat{\mathcal{X}}_k^{s_0}$ subject to multiple sensor set-valued measurements.
3. Analyze the effect of including additional sensors on $\hat{\mathcal{X}}_k^{s_0}$.

In addition, after obtaining the solutions to these problems, we will apply them to the analysis and design in event-based state estimation.

4.3 Sensor fusion

In this section, we analyze the effect of the fusion sequence on the size of the set of estimation means, based on which a “sequence-independent” separate fusion principle of fusing multiple sensor measurements is proposed. This property is of fundamental importance for the analysis of asymptotical behavior and performance improvement in the multiple-sensor scenario, without which the whole set of fusion sequences (the cardinality of which equals $M!$) would have to be considered to analyze the worst-case behavior.

To aid the analysis, we first present the following lemma on the properties of Minkowski sum.

Lemma 4.2. *Let $\mathcal{X}, \mathcal{Y} \subseteq \mathbb{R}^n$, and let $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a linear transformation. Then $T(\mathcal{X} \oplus \mathcal{Y}) = (T\mathcal{X}) \oplus (T\mathcal{Y})$.*

Proof.

$$\begin{aligned}
 (T\mathcal{X}) \oplus (T\mathcal{Y}) &= \{Tx|x \in \mathcal{X}\} \oplus \{Ty|y \in \mathcal{Y}\} \\
 &= \{a + b|a \in \{Tx|x \in \mathcal{X}\}, b \in \{Ty|y \in \mathcal{Y}\}\} \\
 &= \{T(x + y)|x \in \mathcal{X}, y \in \mathcal{Y}\} \\
 &= T(\mathcal{X} \oplus \mathcal{Y}).
 \end{aligned}$$

□

Now we show that the fusion sequence does not affect the exact set of means of the estimates. Before that, we first present some insights into the structure of the filter gains and the closed-loop system matrix. For a given fusion sequence s , the closed-loop matrix $\bar{A}_k^{s_0}$ satisfies

$$\bar{A}_k^{s_0} := A \prod_{i=1}^M (I - K_k^{s_i} C_{s_i}) \quad (4.16)$$

and the filter gain $\bar{K}_k^{s_j}$ for the j th sensor satisfies

$$\bar{K}_k^{s_j} := A \left[\prod_{i=j+1}^M (I - K_k^{s_i} C_{s_i}) \right] K_k^{s_j}. \quad (4.17)$$

For these two matrices, we have the following equivalent representations.

Proposition 4.1. $\bar{A}_k^{s_0} P_k^{s_0} = A P_k^{s_M}$, $\bar{K}_k^{s_j} = A P_k^{s_M} C_{s_j}^\top R_{s_j}^{-1}$.

Proof. First, applying the matrix inversion lemma to equation (4.12), we have

$$P_k^{s_{i+1}} = (I + P_k^{s_i} C_{s_{i+1}}^\top R_{s_{i+1}}^{-1} C_{s_{i+1}})^{-1} P_k^{s_i}. \quad (4.18)$$

Similarly, for $K_k^{s_{i+1}}$, we have

$$\begin{aligned} K_k^{s_{i+1}} &= (I + P_k^{s_i} C_{s_{i+1}}^\top R_{s_{i+1}}^{-1} C_{s_{i+1}})^{-1} P_k^{s_i} C_{s_{i+1}}^\top R_{s_{i+1}}^{-1} \\ &= P_k^{s_{i+1}} C_{s_{i+1}}^\top R_{s_{i+1}}^{-1}. \end{aligned} \quad (4.19)$$

Also, from equation (4.12) and the fact that $K_k^{s_{i+1}} = P_k^{s_i} C_{s_{i+1}}^\top (C_{s_{i+1}} P_k^{s_i} C_{s_{i+1}}^\top + R_{s_{i+1}})^{-1}$, we have

$$P_k^{s_{i+1}} = (I - K_k^{s_{i+1}} C_{s_{i+1}}) P_k^{s_i}. \quad (4.20)$$

From equation (4.17), we have

$$\begin{aligned} \bar{K}_k^{s_j} &= A \left[\prod_{i=j+1}^M (I - K_k^{s_i} C_{s_i}) \right] K_k^{s_j} \\ &= A \left[\prod_{i=j+1}^M (I - K_k^{s_i} C_{s_i}) \right] P_k^{s_j} C_{s_j}^\top R_{s_j}^{-1} \\ &= A P_k^{s_M} C_{s_j}^\top R_{s_j}^{-1}, \end{aligned} \quad (4.21)$$

where the last equality is obtained by recursively applying equation (4.20). The relation for $\bar{A}_k^{s_0}$ can be obtained following a similar argument. \square

Remark 4.1. Notice that if $P_k^{s_0}$ is nonsingular⁴, we have $\bar{A}_k^{s_0} = AP_k^{s_M}(P_k^{s_0})^{-1}$. The above result implies that the filter gains can be updated either by calculating the Riccati equation in (4.12) corresponding to (C_{s_i}, R_{s_i}) sequentially or by lifting all sensor information matrices $\{(C_{s_i}, R_{s_i})\}$ as (C, R) and computing the Riccati equation by replacing C_{s_i} and R_{s_i} with C and R in (4.12). The calculation of $\bar{A}_k^{s_0}$ is straightforward as it is well known that it satisfies

$$\bar{A}_k^{s_0} = A - AP_k^{s_0}C^\top(CP_k^{s_0}C^\top + R)^{-1}. \quad (4.22)$$

The update of the set of estimation means, on the other hand, can only be updated by sequentially fusing the sensor information, although the fusion result is sequence independent, as is shown in the result below.

Theorem 4.1. *Let s^1, s^2 denote two different sensor fusion sequences. We have*

1. *If $P_{k-1}^{s_0^1} = P_{k-1}^{s_0^2}$, then $P_k^{s_0^1} = P_k^{s_0^2}$.*
2. *If $\mathcal{X}_{k-1}^{s_0^1} = \mathcal{X}_{k-1}^{s_0^2}$, then $\mathcal{X}_k^{s_0^1} = \mathcal{X}_k^{s_0^2}$.*

Proof. The first part of the result follows from the matrix inversion lemma and a few matrix manipulations. To prove the second part, first notice that according to Lemma 4.2, we have

$$\mathcal{X}_k^{s_0^r} = \bar{A}_{k-1}^{s_0^r} \mathcal{X}_{k-1}^{s_0^r} \oplus \bigoplus_{j=1}^M \bar{K}_{k-1}^{s_j^r} \mathcal{Y}_{k-1}^{s_j^r} \quad (4.23)$$

for $r \in \mathbb{N}_{1:2}$. From Proposition 4.1, \bar{K}_{k-1}^i only depends on C_i and R_i , which are not affected by the relative position of sensor i in the fusion sequence. Since for $i \in \mathbb{N}_{1:M}$, each sensor i appears once and only once in a fusion sequence, different fusion sequences will lead to different permutation of the same set of summands $\{\bar{K}_{k-1}^i \mathcal{Y}_{k-1}^i | i \in \mathbb{N}_{1:M}\}$ in the second term on the right hand side of equation (4.23). Finally, from (4.22), $\bar{A}_{k-1}^{s_0^r} \mathcal{X}_{k-1}^{s_0^r}$ is unaffected by the fusion sequence either, the conclusion now follows from the commutativity and associativity of Minkowski sums over convex bodies [41, 58]. \square

Furthermore, note that since (A, C) is detectable,

$$\bar{A} = \lim_{k \rightarrow \infty} A \left[\prod_{i=1}^M (I - K_{k-1}^{s_i} C_{s_i}) \right]$$

⁴This condition holds asymptotically if (A, Q) is reachable (see the corollary on page 710 of [11]).

exists and \bar{A} is stable [15], which will be used in the stability analysis in the next section. The above result shows that the estimation performance in terms of either the estimation error covariance or the size of the set of estimation means does not depend on the fusion sequence s for the exact set of means of the estimates. Unfortunately, the exact sets of means of the estimates either in the form (4.23) or the recursive form (4.9)-(4.11) are difficult to obtain analytically when the measurements are given in terms of ellipsoidal sets, since the summation of ellipsoids may not be ellipsoids at all [37], and consequently the analytical expression of the exact set-valued estimator cannot be maintained. Motivated from the above result, however, we propose the following procedure in Algorithm 1 to calculate the outer approximation of the set of estimation means.

Algorithm 1 Calculation of $\hat{\mathcal{X}}_k^{s_0}$

```

1:  $\hat{\mathcal{X}}_0^{s_0} = \mathcal{E}(0, 0)$ ;
2:  $P_0^{s_0} = P_0$ ;
3:  $k = 0$ ;
4: while  $k \geq 0$  do
5:    $P_k^{sM} = P_k^{s_0} - P_k^{s_0} C^\top (C P_k^{s_0} C^\top + R)^{-1} C P_k^{s_0}$ ;
6:    $P_{k+1}^{s_0} = A P_k^{sM} A^\top + Q$ ;
7:    $\bar{A}_k^{s_0} = A - A P_k^{s_0} C^\top (C P_k^{s_0} C^\top + R)^{-1}$ ;
8:   for  $i = 1 : M$  do
9:      $\bar{K}_k^{s_i} = A P_k^{sM} C_{s_i} R_{s_i}^{-1}$ ;
10:  end for
11:   $\bar{\mathcal{X}}_k^{s_0} := \mathcal{E}(\bar{x}_k^{s_0}, \bar{X}_k^{s_0}) = \bar{A}_k^{s_0} \hat{\mathcal{X}}_k^{s_0}$ 
12:     $= \mathcal{E}(\bar{A}_k^{s_0} \hat{x}_k^{s_0}, \bar{A}_k^{s_0} X_k(\bar{A}_k^{s_0})^\top)$ ;
13:  for  $i = 1 : M$  do
14:     $p_k^{s_i} = \sqrt{\text{Tr} \bar{X}_k^{s_{i-1}} / \text{Tr} \bar{K}_k^{s_i} Y_k^{s_i} (\bar{K}_k^{s_i})^\top}$ ;
15:     $\bar{x}_k^{s_i} = \bar{x}_k^{s_{i-1}} + \bar{K}_k^{s_i} c_k^{s_i}$ ;
16:     $\bar{X}_k^{s_i} = (1 + 1/p_k^{s_i}) \bar{X}_k^{s_{i-1}} + (1 + p_k^{s_i}) \bar{K}_k^{s_i} Y_k^{s_i} (\bar{K}_k^{s_i})^\top$ ;
17:     $\bar{\mathcal{X}}_k^{s_i} := \mathcal{E}(\bar{x}_k^{s_i}, \bar{X}_k^{s_i})$ ;
18:  end for
19:   $\hat{\mathcal{X}}_{k+1}^{s_0} := \bar{\mathcal{X}}_k^{sM}$ ;
20:   $k = k + 1$ ;
21: end while
22: end

```

Algorithm 1 indicates that for multiple-sensor set-valued filtering, the fusion of covariance and estimation means should be performed separately: the estimation error covariance is updated first (see lines 5 – 10), where the covariance updates are first

calculated by solving the Riccati equation for C and R , and $\bar{A}_k^{s_0}$ and $\bar{K}_k^{s_i}$ are respectively calculated according to (4.22) and (4.21); then the update of estimation means is performed (see lines 11-19), where the set of estimation means $\hat{\mathcal{X}}_k^{s_0}$ are calculated by iteratively fusing the summands in (4.23) in a two-by-two fashion based on Lemma 4.1 according to an arbitrary fusion sequence s . One may think that it is not necessary to do so, as is the case for classical Kalman filtering with multiple point-valued measurements. We show that, however, the proposed procedure bears the basic properties of the classical Kalman filter while enjoying the benefits of distributed implementation.

To see this, we look into the structure of the outer approximation of the set of the means of the estimates with the help of the following lemmas [37].

Lemma 4.3. *Let $\mathcal{E}(a, Q) \subseteq \mathbb{R}^n$. Then $x \in \mathcal{E}(a, Q)$ is equivalent to $Ax + b \in \mathcal{E}(Aa + b, AQA^\top)$.*

Lemma 4.4.

$$\bigoplus_{i=1}^l \mathcal{E}(c_i, X_i) \subseteq \mathcal{E}(c_0, X_0), \quad (4.24)$$

with $c_0 = \sum_{i=1}^l c_i$,

$$X_0 = \left(\sum_{i=1}^l q_i \right) \sum_{i=1}^l q_i^{-1} X_i \quad (4.25)$$

for all $q_i > 0$, $i \in \mathbb{N}_{1:l}$.

Following these lemmas and equation (4.23), we have the following updating equations of $\hat{\mathcal{X}}_k^{s_0}$ from $\hat{\mathcal{X}}_{k-1}^{s_0} = \mathcal{E}(\hat{x}_{k-1}^{s_0}, X_{k-1})$ and $\mathcal{Y}_k^{s_i} = \mathcal{E}(c_k^{s_i}, Y_k^{s_i})$:

$$\hat{\mathcal{X}}_k^{s_0} = \mathcal{E}(\hat{x}_k^{s_0}, X_k), \quad (4.26)$$

$$\hat{x}_k^{s_0} = \bar{A}_{k-1}^{s_0} \hat{x}_{k-1}^{s_0} + \sum_{j=1}^M \bar{K}_{k-1}^{s_j} c_{k-1}^{s_j}, \quad (4.27)$$

$$\begin{aligned} X_k = & \left(\sqrt{\text{Tr} \bar{A}_{k-1}^{s_0} X_{k-1} (\bar{A}_{k-1}^{s_0})^\top} + \sum_{j=1}^M \sqrt{\text{Tr} \bar{K}_{k-1}^{s_j} Y_{k-1} (\bar{K}_{k-1}^{s_j})^\top} \right) \\ & \left[\left(\sqrt{\text{Tr} \bar{A}_{k-1}^{s_0} X_{k-1} (\bar{A}_{k-1}^{s_0})^\top} \right)^{-1} \bar{A}_{k-1}^{s_0} X_{k-1} (\bar{A}_{k-1}^{s_0})^\top \right. \\ & \left. + \sum_{j=1}^M \left(\sqrt{\text{Tr} \bar{K}_{k-1}^{s_j} Y_{k-1} (\bar{K}_{k-1}^{s_j})^\top} \right)^{-1} \bar{K}_{k-1}^{s_j} Y_{k-1} (\bar{K}_{k-1}^{s_j})^\top \right]. \end{aligned} \quad (4.28)$$

Using a similar argument as that in Section IV.A of [52], equation (4.28) can be evaluated in an iterative way as in lines 11-19 of Algorithm 1 according to an arbitrary sequence. This not only helps to reduce the computational complexity at the fusion centre through distributed computation (the acknowledgement or computation of $\bar{A}_k^{s_0}$ and $\bar{K}_k^{s_i}$ at sensor i would be necessary), but also guarantees the invariance of outer ellipsoidal approximation of the set of estimation means with respect to the fusion sequence. In [52], to possess this property, a different filter form, namely, the information filter [50], was considered. On the other hand, the filter form utilized here inherits the original form of Kalman filter with multiple point-valued measurements, due to the separate covariance and estimate updating procedure in Algorithm 1. Notice that in fact, at steady state, only the update of the estimation means (lines 11-19) is necessary, since the solution to the Riccati equation converges to its unique stabilizing solution, and therefore the algorithm can be implemented in a completely distributive way without considering covariance update at the steady state.

4.4 Asymptotic properties of the set of means of the estimates

In this section, the objective is to discuss the asymptotical boundedness properties of both the exact and outer-approximate sets of estimation means for the multiple sensor case. We first focus on the single sensor case and analyze the asymptotic properties of the set-valued mean evolutions, and then extend the results to multiple sensor case. When there is only one sensor, the equations are given by

$$\mathcal{X}_k^0 = A\mathcal{X}_{k-1}^1, \quad (4.29)$$

$$\mathcal{X}_k^1 = (I - K_k C)\mathcal{X}_k^0 \oplus K_k \mathcal{Y}_k. \quad (4.30)$$

In the prediction form, we have

$$\mathcal{X}_{k+1}^0 = \bar{A}_k \mathcal{X}_k^0 \oplus \bar{K}_k \mathcal{Y}_k, \quad (4.31)$$

where $\bar{A}_k = A(I - K_k C)$, $\bar{K}_k = AK_k$. Correspondingly, let $\hat{\mathcal{X}}_k^0 = \mathcal{E}(\hat{x}_k^0, X_k)$ and $\mathcal{Y}_k = \mathcal{E}(c_k, Y_k)$, and the approximate estimate is given by

$$\hat{\mathcal{X}}_{k+1}^0 = \mathcal{E}(\bar{A}_k \hat{x}_k^0 + \bar{K}_k c_k, X_{k+1}), \quad (4.32)$$

$$X_{k+1} = \left(1 + \frac{\sqrt{\text{Tr } \bar{K}_k Y_k \bar{K}_k^\top}}{\sqrt{\text{Tr } \bar{A}_k X_k \bar{A}_k^\top}}\right) \bar{A}_k X_k \bar{A}_k^\top + \left(1 + \frac{\sqrt{\text{Tr } \bar{A}_k X_k \bar{A}_k^\top}}{\sqrt{\text{Tr } \bar{K}_k Y_k \bar{K}_k^\top}}\right) \bar{K}_k Y_k \bar{K}_k^\top. \quad (4.33)$$

The objective of this section is to show the boundedness of the sizes of the sequence of sets $\{\mathcal{X}_k^0\}$ and the possible boundedness of the sizes of the sequences of sets $\{\hat{\mathcal{X}}_k^0\}$. Before continuing, we give the following lemma.

Lemma 4.5. *Let $Q \geq 0$, and $0 \leq P < I$. Then $\text{Tr } QP \leq \text{Tr } Q$.*

Proof. Since $0 \leq P < I$, there exists a unitary matrix U such that $U^\top = U^{-1}$ and

$$U^\top P U = U^{-1} P U = \text{diag}\{p_1, p_2, \dots, p_n\}$$

satisfying $0 \leq p_i < 1$, p_i being the eigenvalues of P . Let $p^* = \max_{i \in \mathbb{N}_{1:n}} p_i < 1$. We have

$$\text{Tr } QP = \text{Tr } U^{-1} Q U U^{-1} P U = \text{Tr } U^{-1} Q U \text{diag}\{p_i\}.$$

Since $U^{-1} Q U = U^\top Q U \geq 0$, the diagonal elements of $U^{-1} Q U$ are nonnegative. Thus we have

$$\text{Tr } U^{-1} Q U \text{diag}\{p_i\} \leq \text{Tr } U^{-1} Q U p^* I = p^* \text{Tr } Q \leq \text{Tr } Q.$$

Notice that since $p^* < 1$, the equality holds if and only if $Q = 0$. □

Now we are ready to present the first result on the asymptotic properties of the sizes of the sets of the means.

Theorem 4.2. *Assume the pair (A, C) is detectable and (A, Q) is stabilizable. Let $\bar{A} := \lim_{k \rightarrow \infty} \bar{A}_k$ and $\bar{K} := \lim_{k \rightarrow \infty} \bar{K}_k$.*

1. *The sizes of the sequence of sets $\{\mathcal{X}_k^0\}$ are asymptotically bounded for all measurement set sequences $\{\mathcal{Y}_k\}$ with bounded sizes.*
2. *If $\|\bar{A}\|_2 < 1$, the sizes of the sequence of ellipsoids $\{\hat{\mathcal{X}}_k^0\}$ are asymptotically bounded for all measurement set sequences $\{\mathcal{Y}_k\}$ with bounded sizes.*
3. *If $\|\bar{A}\|_2 \geq 1$, there exists an invertible linear transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that the sizes of the sequence of the set of estimation means $\{\hat{\mathcal{X}}_k^0\}$ for the transformed state $\tilde{x}_k := T x_k$ are asymptotically bounded for all measurement set sequences $\{\mathcal{Y}_k\}$ with bounded sizes.*

Proof. Since (A, C) is detectable and (A, Q) is stabilizable, the Kalman filter is stable and the solution to the Riccati equation converges to the unique stabilizing solution.

Thus $\bar{A} = \lim_{k \rightarrow \infty} \bar{A}_k$ and $\bar{K} = \lim_{k \rightarrow \infty} \bar{K}_k$ are well defined, and satisfy $\bar{A} = A - \bar{K}C$ and $\bar{K} = A\bar{P}C^\top(C\bar{P}C^\top + R)^{-1}$, \bar{P} being the stabilizing solution to the Riccati equation

$$P = APA^\top - APC^\top(CPC^\top + R)^{-1}CPA^\top + Q.$$

We will prove parts (2) and (3) before proving the result in part (1).

First we show that if $\|\bar{A}\|_2 < 1$, the evolution of the size of the outer approximation of the Minkowski sum in (4.32-4.33) is asymptotically bounded. Since the evolution of equation (4.33) does not affect the evolution of the covariance and (A, C) is detectable, it suffices to consider the steady-state Kalman filter gain, which is equivalent to the consideration of \bar{A} . At steady state, taking traces on both sides of equation (4.33), we have

$$\begin{aligned} \text{Tr } X_{k+1} &= \left(1 + \frac{\sqrt{\text{Tr } \bar{K}Y_k\bar{K}^\top}}{\sqrt{\text{Tr } \bar{A}X_k\bar{A}^\top}}\right) \text{Tr } \bar{A}X_k\bar{A}^\top \\ &\quad + \left(1 + \frac{\sqrt{\text{Tr } \bar{A}X_k\bar{A}^\top}}{\sqrt{\text{Tr } \bar{K}Y_k\bar{K}^\top}}\right) \text{Tr } \bar{K}Y_k\bar{K}^\top \\ &= \left(\sqrt{\text{Tr } \bar{A}X_k\bar{A}^\top} + \sqrt{\text{Tr } \bar{K}Y_k\bar{K}^\top}\right)^2 \end{aligned} \quad (4.34)$$

Since $X_{k+1} \geq 0$, we have $\text{Tr } X_{k+1} \geq 0$. Thus

$$\begin{aligned} \sqrt{\text{Tr } X_{k+1}} &= \sqrt{\text{Tr } \bar{A}X_k\bar{A}^\top} + \sqrt{\text{Tr } \bar{K}Y_k\bar{K}^\top} \\ &= \sqrt{\text{Tr } X_k\bar{A}^\top\bar{A}} + \sqrt{\text{Tr } \bar{K}Y_k\bar{K}^\top} \\ &\leq \sqrt{a^*} \sqrt{\text{Tr } X_k} + \sqrt{\text{Tr } \bar{K}Y_k\bar{K}^\top}, \end{aligned} \quad (4.35)$$

for some $a^* \in (0, 1)$, which follows from $\|\bar{A}\|_2 < 1$ and Lemma 4.5. This implies the boundedness of $\{\sqrt{\text{Tr } X_k}\}$, given the boundedness of $\{Y_k\}$.

Now we consider the case $\|\bar{A}\|_2 \geq 1$. Since \bar{A} is stable, there exists $P_s > 0$ such that P_s is the solution to the Lyapunov equation $\bar{A}^\top P_s \bar{A} - P_s + I = 0$, which implies $P_s \geq I > 0$. Now we introduce a linear transformation $T = P_s^{1/2} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and let $\tilde{x}_k = Tx_k$. Apparently \tilde{x}_k evolves according to

$$\begin{aligned} \tilde{x}_{k+1} &= \tilde{A}\tilde{x}_k + \tilde{B}w_k, \\ y_k &= \tilde{C}\tilde{x}_k + v_k, \end{aligned}$$

where $\tilde{A} = TAT^{-1}$, $\tilde{B} = T$, $\tilde{C} = CT^{-1}$. Furthermore, it is easy to verify that $\tilde{K} := \tilde{A}\tilde{P}\tilde{C}^\top(\tilde{C}\tilde{P}\tilde{C}^\top + R)^{-1} = T\bar{K}$ and $\tilde{P} = T\bar{P}T^\top$, \tilde{P} being the stabilizing solution to the Riccati equation

$$P = \tilde{A}P\tilde{A}^\top - \tilde{A}\tilde{P}\tilde{C}^\top(\tilde{C}\tilde{P}\tilde{C}^\top + R)^{-1}\tilde{C}P\tilde{A}^\top + TQT^\top.$$

Define $\bar{\bar{A}} := \tilde{A} - \bar{K}\tilde{C}$, we have $\bar{\bar{A}} = T(A - \bar{K}C)T^{-1} = P_s^{1/2}\bar{A}P_s^{-1/2}$. Thus

$$\begin{aligned}\bar{\bar{A}}^\top \bar{\bar{A}} &= P_s^{-1/2}\bar{A}^\top P_s \bar{A} P_s^{-1/2} \\ &= P_s^{-1/2}(P_s - I)P_s^{-1/2} = I - P_s^{-1} < I,\end{aligned}$$

which implies $\|\bar{\bar{A}}\|_2 < 1$. Therefore the conclusion of part (3) follows from the same argument used for proof of part (2).

Finally we prove the result in part (1). The case of $\|\bar{A}\|_2 < 1$ follows from the result in part (2), since $\{\hat{\mathcal{X}}_k^0\}$ provides an outer approximation of $\{\mathcal{X}_k^0\}$. To prove the case of $\|\bar{A}\|_2 \geq 1$, we analyze the relationship between the exact Minkowski sum for the original state estimate and that of the transformed state estimate. Similar to equation (4.31), we have

$$\tilde{\mathcal{X}}_{k+1}^0 = \bar{\bar{A}}_k \tilde{\mathcal{X}}_k^0 \oplus \bar{\bar{K}}_k \mathcal{Y}_k, \quad (4.36)$$

where $\bar{\bar{A}}_k = \tilde{A} - \bar{\bar{K}}_k \tilde{C}$, $\bar{\bar{K}}_k = \tilde{A}\tilde{P}_{k-1}\tilde{C}^\top (\tilde{C}\tilde{P}_{k-1}\tilde{C}^\top + R)^{-1}$ and \tilde{P}_k being the solution to the Riccati equation $\tilde{P}_{k+1} = \tilde{A}\tilde{P}_k\tilde{A}^\top - \tilde{A}\tilde{P}_k\tilde{C}^\top (\tilde{C}\tilde{P}_k\tilde{C}^\top + R)^{-1}\tilde{C}\tilde{P}_k\tilde{A}^\top + TQT^\top$ subject to $\tilde{P}_0 = TP_0T^\top$. At time $t = 0$, $\tilde{\mathcal{X}}_0^0 = \{Tx_0\} = T\mathcal{X}_0^0$. Note that following a similar argument as that in the proof of part (3), $\bar{\bar{A}}_k = T\bar{A}_kT^{-1}$ and $\bar{\bar{K}}_k = T\bar{K}_k$. Now assume at time $t = k$, the relationship $\tilde{\mathcal{X}}_k^0 = T\mathcal{X}_k^0$ holds. We have

$$\tilde{\mathcal{X}}_{k+1}^0 = T\bar{A}_kT^{-1}\tilde{\mathcal{X}}_k^0 \oplus T\bar{K}_k\mathcal{Y}_k. \quad (4.37)$$

Following the definition Minkowski sum,

$$\begin{aligned}\tilde{\mathcal{X}}_{k+1}^0 &:= \left\{ T\bar{A}_kT^{-1}\tilde{x} + T\bar{K}_ky \mid \tilde{x} \in \tilde{\mathcal{X}}_k^0, y \in \mathcal{Y}_k \right\} \\ &= \left\{ T\bar{A}_kx + T\bar{K}_ky \mid x \in \mathcal{X}_k^0, y \in \mathcal{Y}_k \right\} \\ &= \left\{ T(a + b) \mid a \in \{\bar{A}_kx \mid x \in \mathcal{X}_k^0\}, b \in \{\bar{K}_ky \mid y \in \mathcal{Y}_k\} \right\} \\ &= T(\bar{A}_k\mathcal{X}_k^0 \oplus \bar{K}_k\mathcal{Y}_k) \\ &= T\mathcal{X}_{k+1}^0.\end{aligned}$$

Thus $\tilde{\mathcal{X}}_k^0 = T\mathcal{X}_k^0$ for all k . Since T is nonsingular, the boundedness of $\{\mathcal{X}_k^0\}$ is equivalent to that of $\{\tilde{\mathcal{X}}_k^0\}$. The conclusion follows from part (3) of the theorem and the fact that $\{\hat{\mathcal{X}}_k^0\}$ provides an outer approximation of $\{\mathcal{X}_k^0\}$. \square

Remark 4.2. From the above proof, a quantitative relationship of the size of the set of estimation means with the sets of measurements, the statistical properties of the

noises and the system matrices can be obtained. To see this, assume there exists an upper bound $\bar{Y} \geq Y_k$ for all $k \in \mathbb{N}$. For the case $\|\bar{A}\|_2 < 1$, from (4.35), we have

$$\sqrt{\text{Tr } X_{k+1}} \leq \|\bar{A}\|_2 \sqrt{\text{Tr } X_k} + \sqrt{\text{Tr } \bar{K} \bar{Y} \bar{K}^\top}. \quad (4.38)$$

Thus we have

$$\lim_{k \rightarrow \infty} \sqrt{\text{Tr } X_k} \leq \sqrt{\text{Tr } \bar{K} \bar{Y} \bar{K}^\top} / (1 - \|\bar{A}\|_2),$$

where the system and noise parameters are reflected in \bar{K} and \bar{A} (Recall that $\bar{K} = A\bar{P}C^\top(C\bar{P}C^\top + R)^{-1}$ and $\bar{A} = A - \bar{K}C$, respectively, \bar{P} being the stabilizing solution to the Riccati equation $P = APA^\top - APC^\top(CPC^\top + R)^{-1}CPA^\top + Q$). The same analysis applies to the case of $\|\bar{A}\|_2 \geq 1$ by introducing the linear transformation $T = P_s^{1/2}$. In addition, note that the boundedness of $\{\tilde{\mathcal{X}}_k^0\}$ does not imply that of $\{\hat{\mathcal{X}}_k^0\}$, since considering the trace operations in (4.33), the relationship between $\hat{\mathcal{X}}_k^0$ and $\tilde{\mathcal{X}}_k^0$ is not clear.

Another consequence of the above result is that for first-order systems with constant size of measurement set, we are able to exactly characterize the size of the set of means of the estimate at steady state.

Corollary 4.1. For $n = m = 1$, and $Y_k = Y$. The size of $\{\mathcal{X}_k\}$ converges to $|\bar{K}\sqrt{Y}|/(1 - |\bar{A}|)$.

Proof. The proof of this result follows from inequality (4.35) and the fact that $|\bar{A}| < 1$ always holds for $n = 1$. \square

The next result generalizes Theorem 4.2 to the multiple sensor case, utilizing the properties of the outer-approximate estimate set.

Corollary 4.2. Consider the exact and approximate multiple sensor set-valued estimators in (4.9)-(4.12) and (4.26)-(4.28), respectively. Assume (A, C) is detectable and (A, Q) is stabilizable. Let $\bar{A} := \lim_{k \rightarrow \infty} \bar{A}_k^{s_0}$.

1. The sizes of the sequence of sets $\{\mathcal{X}_k^{s_0}\}$ are asymptotically bounded for all measurement set sequences $\{\mathcal{Y}_k^{s_i}\}$ with bounded sizes.
2. If $\|\bar{A}\|_2 < 1$, the sizes of the sequence of ellipsoids $\{\hat{\mathcal{X}}_k^{s_0}\}$ are asymptotically bounded for all measurement set sequences $\{\mathcal{Y}_k^{s_i}\}$ with bounded sizes.
3. If $\|\bar{A}\|_2 \geq 1$, there exists an invertible linear transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that the sizes of the set of mean of the estimates $\{\hat{\mathcal{X}}_k^{s_0}\}$ for the transformed state $\tilde{x}_k := Tx_k$ are asymptotically bounded for all measurement set sequences $\{\mathcal{Y}_k^{s_i}\}$ with bounded sizes.

Proof. To show this result, we establish the relationship between the single sensor case and the multiple sensor case. By calculating the traces of both sides for (4.28), it is not difficult to verify that

$$\begin{aligned} \text{Tr } X_k = & \left(\sqrt{\text{Tr } \bar{A}_{k-1}^{s_0} X_{k-1} (\bar{A}_{k-1}^{s_0})^\top} \right. \\ & \left. + \sum_{i=1}^M \sqrt{\text{Tr } \bar{K}_{k-1}^{s_i} Y_{k-1}^{s_i} (\bar{K}_{k-1}^{s_i})^\top} \right)^2, \end{aligned} \quad (4.39)$$

and thus

$$\sqrt{\text{Tr } X_k} = \sqrt{\text{Tr } \bar{A}_{k-1}^{s_0} X_{k-1} (\bar{A}_{k-1}^{s_0})^\top} \quad (4.40)$$

$$+ \sum_{i=1}^M \sqrt{\text{Tr } \bar{K}_{k-1}^{s_i} Y_{k-1}^{s_i} (\bar{K}_{k-1}^{s_i})^\top}. \quad (4.41)$$

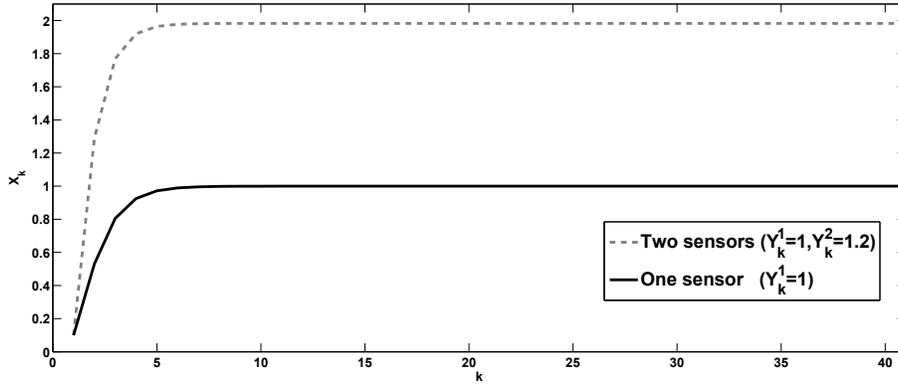
Noticing the relationship with equation (4.35) and the boundedness of $\{Y_k^{s_i}\}$, the results are proved with a similar argument as that in the proof of Theorem 4.2. \square

4.5 Performance improvement

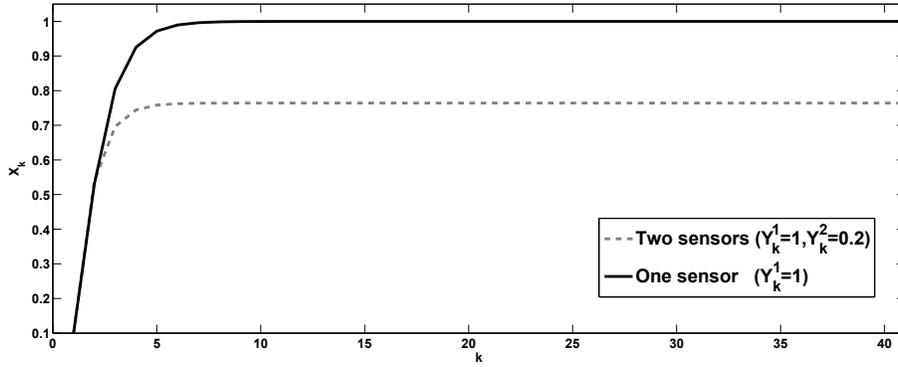
Now we analyze the effect of including more sensors on the estimation performance in the set-valued estimation framework. Adding sensors always reduces the estimation error covariance, following the monotonicity properties of the solutions to the Riccati equations. Adding sensors, however, does not necessarily reduce the size of the set of the means of the estimates, as is shown in the following example.

Example 4.1. Consider the system in (4.1) with $n = 1$, $m = 1$, $A = 1.3$, $Q = 1.2$, $C_1 = 1$, $C_2 = 0.6$, $R_1 = 1.9$, $R_2 = 0.7$. Assume $Y_k^1 = 1$, we consider two choices of Y_k^2 : (1) $Y_k^2 = 1.2$, (2) $Y_k^2 = 0.2$. The performance in terms of the size of the set of estimation means obtained by using sensor 1 alone and using sensor 1 and sensor 2 are shown in Fig. 4.1 (a) and (b), respectively. It is shown that when $Y_k^2 = 0.2$, the addition of sensor 2 helps to improve the estimation performance; the choice of $Y_k^2 = 1.2$, however, deteriorates the performance in terms of a larger size of the set of estimation means.

Motivated by the above example, given an existing sensor 1, it is interesting to characterize conditions on properties of sensor 2 such that improved performance can be guaranteed. Now we make the problem more explicit. Suppose we have a linear



(a) Performance comparison for $Y_k^1 = 1, Y_k^2 = 1.2$



(b) Performance comparison for $Y_k^1 = 1, Y_k^2 = 0.2$

Figure 4.1: Comparison of the sizes of the sets of estimation means for different choices of Y_k^2 .

system originally measured only by sensor 1, namely, equations (4.1) and (4.2) with $M = 1$. Now we introduce sensor 2 and measure the system state using two sensors. We want to compare the size of the set of estimation means obtained only using sensor 1 with that using sensors 1 and 2 together. First we need to quantify the performance. To do this, we focus on the steady-state behavior of the size of the set of estimation means by assuming that the closed-loop matrix \bar{A} under consideration satisfies $\|\bar{A}\|_2 < 1$ (Note that if this condition is not satisfied by the original system, we can introduce the linear transformation T in the proof of Theorem 4.2 such that the transformed closed-loop matrix satisfies this condition). For simplicity, we assume the shape of $\mathcal{Y}_k^{s_i}$ is asymptotically time invariant, namely, $\lim_{k \rightarrow \infty} Y_k^{s_i} = Y^{s_i}$, and further assume (A, Q) is reachable, which guarantees the positive definiteness of P^{s_0} (see the corollary on page 710 of [11]). From equation (4.41), we have

$$\sqrt{\text{Tr } X_k} \leq \|\bar{A}_{k-1}^{s_0}\|_2 \sqrt{\text{Tr } X_{k-1}} + \sum_{i=1}^M \sqrt{\text{Tr } \bar{K}_{k-1}^{s_i} Y_{k-1}^{s_i} (\bar{K}_{k-1}^{s_i})^\top}. \quad (4.42)$$

Since $X_k \geq 0$ and $Y^{s_i} \geq 0$, the solution to the following equation serves as an upper bound for the size of the set of means of estimates at steady state

$$\bar{x} = \|\bar{A}^{s_0}\|_2 \bar{x} + \sum_{i=1}^M \sqrt{\text{Tr } \bar{K}^{s_i} Y^{s_i} (\bar{K}^{s_i})^\top}. \quad (4.43)$$

From Proposition 4.1, this is equivalent to

$$\begin{aligned} \bar{x} = & \|AP^{s_M}(P^{s_0})^{-1}\|_2 \bar{x} \\ & + \sum_{i=1}^M \sqrt{\text{Tr } AP^{s_M} C_{s_i}^\top R_{s_i}^{-1} Y^{s_i} R_{s_i}^{-1} C_{s_i} P^{s_M} A^\top}, \end{aligned} \quad (4.44)$$

where P^{s_0} is the stabilizing solution to the algebraic Riccati equation

$$P = APA^\top - APC^\top (CPC^\top + R)^{-1} CPA^\top + Q, \quad (4.45)$$

and

$$P^{s_M} = P^{s_0} - P^{s_0} C^\top (C P^{s_0} C^\top + R)^{-1} C P^{s_0}. \quad (4.46)$$

Notice that $P^{s_0} = AP^{s_M} A^\top + Q$. Thus from Proposition 4.1, we have

$$\begin{aligned} \bar{A}^{s_0} &= AP^{s_M} (P^{s_0})^{-1} \\ &= AP^{s_M} (AP^{s_M} A^\top + Q)^{-1}. \end{aligned} \quad (4.47)$$

For the case of one sensor ($M = 1$), we denote $P^1 := P^{s_1}$ for brevity, and the steady-state performance is

$$\bar{x}^1 = \frac{\sqrt{\text{Tr } AP^1 C_1^\top R_1^{-1} Y^1 R_1^{-1} C_1 P^1 A^\top}}{1 - \|AP^1(AP^1 A^\top + Q)^{-1}\|_2}. \quad (4.48)$$

When sensor 2 is included (namely, $M = 2$), we denote $P^2 := P^{s_2}$ for brevity, and the steady-state performance becomes

$$\bar{x}^{1,2} = \left[\sqrt{\text{Tr } AP^2 C_1^\top R_1^{-1} Y^1 R_1^{-1} C_1 P^2 A^\top} + \sqrt{\text{Tr } AP^2 C_2^\top R_2^{-1} Y^2 R_2^{-1} C_2 P^2 A^\top} \right] / \quad (4.49)$$

$$[1 - \|AP^2(AP^2 A^\top + Q)^{-1}\|_2]. \quad (4.50)$$

Notice that for sensor i , the C_i and R_i matrices are fixed and cannot be adjusted; the only adjustable parameter⁵ is Y^i , which controls the size and shape of the set of measurement. Therefore, at steady state, the parameters A , C_i , R_i , P^i are constant. In this way, it is easier to check whether a choice of Y^2 will lead to improved performance in terms of the size of the set of estimation means. In particular, the condition becomes easier to verify when Y^i 's has special structures, e.g., $Y^i = \eta^i I$, which can be used for design purposes. Finally, we consider scalar systems, namely, $n = m = 1$, and have the following result.

Proposition 4.2. *For $n = m = 1$, if $\frac{Y^2}{Y^1} < \left[\frac{(P^1 - P^2)C_1 R_1^{-1}}{P^2 C_2 R_2^{-1}} \right]^2$, then adding sensor 2 improves the steady-state performance in terms of both estimation error covariance and size of the set of the means of the estimates.*

Proof. When $n = m = 1$, equations (4.48) and (4.50) reduce to

$$\begin{aligned} \bar{x}^1 &= \frac{|A|P^1 C_1^\top R_1^{-1} \sqrt{Y^1}}{1 - |A|P^1(P^0)^{-1}} \\ &= \frac{|A|P^1 C_1^\top R_1^{-1} \sqrt{Y^1}}{1 - |A|P^1(AP^1 A + Q)^{-1}}, \end{aligned} \quad (4.51)$$

and

$$\bar{x}^{1,2} = \frac{|A|P^2 \left(C_1^\top R_1^{-1} \sqrt{Y^1} + C_2^\top R_2^{-1} \sqrt{Y^2} \right)}{1 - |A|P^2(AP^2 A + Q)^{-1}}, \quad (4.52)$$

⁵This can be achieved by changing the event-triggering conditions in the microprocessors on the sensor side.

respectively. Since $C_1 R_1^{-1} C_1 < C^\top R^{-1} C$, from the monotonicity properties of the solutions to the Riccati equations (4.45) and (4.46) (Lemma 3 of [59]), we have $P^1 > P^2$. Therefore

$$\begin{aligned}
\frac{Q}{AP^1 A + Q} &< \frac{Q}{AP^2 A + Q} \\
&\Rightarrow \frac{|A|P^1}{AP^1 A + Q} > \frac{|A|P^2}{AP^2 A + Q} \\
&\Rightarrow 1 - \frac{|A|P^1}{AP^1 A + Q} < 1 - \frac{|A|P^2}{AP^2 A + Q} \\
&\Rightarrow \frac{1}{1 - \frac{|A|P^1}{AP^1 A + Q}} > \frac{1}{1 - \frac{|A|P^2}{AP^2 A + Q}}, \tag{4.53}
\end{aligned}$$

where the fact that $0 < |A|P^1/(AP^1 A + Q) < 1$ and $0 < |A|P^2/(AP^2 A + Q) < 1$ are utilized in the last line, due to equation (4.47) and the stability of the Kalman filter. Since $\frac{Y^2}{Y^1} < \left(\frac{(P^1 - P^2)C_1 R_1^{-1}}{P^2 C_2 R_2^{-1}}\right)^2$, we have

$$|A|P^1 C_1 R_1^{-1} \sqrt{Y^1} > |A|P^2 \left(C_1 R_1^{-1} \sqrt{Y^1} + C_2 R_2^{-1} \sqrt{Y^2} \right).$$

Combining with (4.53), we have $\bar{x}^1 > \bar{x}^{1,2}$, which completes the proof. \square

Remark 4.3. The intuition provided in the above result is that to achieve improved performance, the accuracy of sensor 2 should exceed certain level determined by that of sensor 1, although this does not require the accuracy of sensor 2 should be better compared with that of sensor 1.

Remark 4.4. As the confidence on stochastic and non-stochastic uncertainties is parameterized separately as covariance and the size of the set of estimation means, evaluation of the overall performance of a set-valued estimator is more complicated compared with its point-valued counterpart. Adding a sensor can always reduce the estimation error covariance, but can decrease, slightly or even severely increase the size of the set of estimation means. One possible approach of evaluating the overall or combined performance is to explore the equivalence relationship between stochastic and non-stochastic uncertainties in some sense, which can be potentially pursued based on ideas of the probabilistic approach or randomized algorithms utilized in control and estimation of uncertain systems [60, 70].

It is straightforward to observe that similar phenomenon exists for the multiple sensor case, and conditions for performance improvement can be obtained in a similar

way. On the other hand, when taking all Y^i 's as tuning parameters, the above analysis can be utilized to formulate design problems such that pre-specified performance can be achieved, as will be shown in the next section.

4.6 Application to event-based state estimation

In this section, we show how the results obtained in this work can be applied in remote event-based state estimation. Furthermore, an optimization problem is formulated and solved to design the event-triggering conditions by considering requirements on estimation performance and communication rates.

4.6.1 Analysis and parameter design

We consider the system in equation (4.1) measured by M sensors described in equation (4.2), which communicate with the remote state estimator through a wireless channel (see Fig. 4.2). To simplify the analysis, we assume the channel is reliable with no packet dropouts. Due to the event trigger, the sets \mathcal{Y}_k^i have more detailed parameterizations. At each time instant, the sensors measure the current state and decide whether to send the current measurement or not according to the values of binary decision variables γ_k^i 's that are determined by pre-specified triggering conditions. We consider a relatively general description of the triggering conditions:

$$\gamma_k^i = \begin{cases} 0, & \text{if } y_k^i \in \bar{\mathcal{Y}}_k^i \\ 1, & \text{if } y_k^i \notin \bar{\mathcal{Y}}_k^i \end{cases} \quad (4.54)$$

where

$$\bar{\mathcal{Y}}_k^i = \{y \in \mathbb{R}^m \mid (y - \bar{c}_k^i)^\top (\bar{Y}_k^i)^{-1} (y - \bar{c}_k^i) \leq 1\}. \quad (4.55)$$

Note that the necessity of transmitting \bar{Y}_k^i and \bar{c}_k^i to the estimator during the non-event instants depends on the specific triggering conditions under consideration, as will be shown in Section 4.6.2. In this case, when $\gamma_k^i = 1$, the remote estimator receives the point-valued measurement information from sensor i , and thus the set of measurement information is given by a singleton $\mathcal{Y}_k^i = \{y_k^i\}$; when $\gamma_k^i = 0$, the set of measurement information is implicitly given by $\mathcal{Y}_k^i = \bar{\mathcal{Y}}_k^i$.

From the results obtained in Sections 4.3 and 4.4, it is known that

1. The performance of the exact and approximate set-valued event-based estimators are invariant with respect to the fusion sequence. Notice that the counter part for either the exact or approximate MMSE event-based estimator is very difficult to be theoretically verified (see Chapter 3).

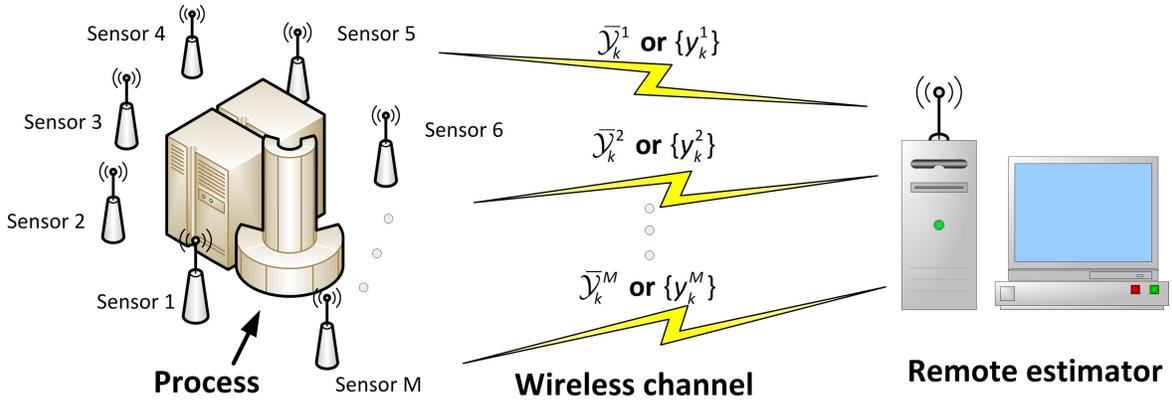


Figure 4.2: Multiple-sensor event-based remote estimation architecture.

2. For the event-based set-valued estimator, the set of estimation means is asymptotically bounded, and the outer approximations of the sets are bounded as well, which can be calculated according to arbitrary fusion sequences at each time instant.

On the other hand, it is still not clear how to design the event-triggering conditions so that the requirements on communication rate and estimation performance can be simultaneously considered, which is one of the main concerns in event-based control and estimation [43, 73]. In the following, we show how the analysis in Section 4.5 can be utilized in parameter design problems for guaranteed worst-case estimation performance and optimized communication rate.

From the analysis in the previous section, it is intuitively known that increasing the trace of \bar{Y}_k^i increases the size of the set of estimation means, which leads to the decrease of estimation performance. On the other hand, from the literature of event-based estimation [73], it is known that the increase of \bar{Y}_k^i leads to the reduction of the communication rate⁶. Therefore \bar{Y}_k^i 's can serve as tuning parameters for the tradeoff between estimation performance and communication rate. Note that the estimation performance here considers the size of the set of estimation means only, since the covariance is independent of \bar{Y}_k^i in the set-valued filtering framework.

For convenience of design and implementation, we consider the parameters \bar{Y}_k^i 's to be time invariant, namely, $\bar{Y}_k^i = \bar{Y}^i$. First we introduce the constraints on the esti-

⁶Considering the scope of this work, we omit the analysis of the exact relationship between the communication rates and \bar{Y}^i , although, in fact, this analysis can be done following the approach in [73] with the difference that no Gaussian assumptions are required under the framework of set-valued filtering in this work.

mation performance. Observing that the measurement set \mathcal{Y}_k^i is time varying (which can be $\bar{\mathcal{Y}}_k^i$ or $\{y_k^i\}$ depending on the value of γ_k^i), we consider the worst-case transient performance, namely, $\gamma_k^i = 0$ for a large number of consecutive k 's such that the measurement set is always given by $\bar{\mathcal{Y}}_k^i$ during this period. From Corollary 4.2, the upper bound on the size of the set of estimation means will evolve towards an equilibrium, which thus quantifies the worst-case performance. We still assume $\|\bar{A}\|_2 < 1$; in case that $\|\bar{A}\|_2 \geq 1$, the results developed in this section can be applied by introducing the linear transformation T defined in the proof of Theorem 4.2 to the system. From (4.44), the worst-case performance bound is given by

$$\bar{x} = \frac{\sum_{i=1}^M \sqrt{\text{Tr} AP^M C_i^\top R_i^{-1} \bar{Y}^i R_i^{-1} C_i P^M A^\top}}{1 - \|AP^M (AP^M A^\top + Q)^{-1}\|_2}, \quad (4.56)$$

where P^M , C_i and R_i are used instead of P^{s_M} , C_{s_i} and R_{s_i} for notational brevity, since P^{s_M} is independent of the fusion sequence. To guarantee the worst-case performance, we specify an upper bound \bar{x}^* and enforce the constraint $\bar{x} \leq \bar{x}^*$. From (4.56), direct verification of this constraint is not computationally efficient for design purposes. Alternatively, using the Cauchy-Schwarz inequality,

$$\begin{aligned} \bar{x} &= \frac{\sum_{i=1}^M \sqrt{\text{Tr} \bar{Y}^i R_i^{-1} C_i P^M A^\top AP^M C_i^\top R_i^{-1}}}{1 - \|AP^M (AP^M A^\top + Q)^{-1}\|_2} \\ &\leq \frac{\sum_{i=1}^M \sqrt{\text{Tr} \bar{Y}^i} \sqrt{\text{Tr} R_i^{-1} C_i P^M A^\top AP^M C_i^\top R_i^{-1}}}{1 - \|AP^M (AP^M A^\top + Q)^{-1}\|_2}, \end{aligned} \quad (4.57)$$

thus the performance inequality can be indirectly enforced by requiring

$$\frac{\sum_{i=1}^M \sqrt{\text{Tr} \bar{Y}^i} \sqrt{\text{Tr} R_i^{-1} C_i P^M A^\top AP^M C_i^\top R_i^{-1}}}{1 - \|AP^M (AP^M A^\top + Q)^{-1}\|_2} \leq \bar{x}^*, \quad (4.58)$$

which is a linear constraint of $\sqrt{\text{Tr} \bar{Y}^i}$. On the other hand, we also include requirements on the upper bounds of the communication rates of each sensor by considering $\text{Tr} \bar{Y}_i \geq \underline{\eta}^i \geq 0$, which is equivalent to $\sqrt{\text{Tr} \bar{Y}_i} \geq \sqrt{\underline{\eta}^i}$. The objective of the parameter design is to minimize the communication rate, which is done by maximizing $\sum_{i=1}^M \text{Tr} \bar{Y}_i$. To summarize, the parameter design problem is formulated as the following optimization problem:

$$\begin{aligned} \max_{a_1, a_2, \dots, a_M} \quad & \sum_{i=1}^M a_i^2 \\ \text{s.t.} \quad & \sum_{i=1}^M b_i a_i \leq \bar{x}^*, \\ & a_i \geq \sqrt{\underline{\eta}_i}, \quad i = 1, 2, \dots, M, \end{aligned} \quad (4.59)$$

where $a_i = \sqrt{\text{Tr } \bar{Y}^i}$ and $b_i = \frac{\sqrt{\text{Tr } R_i^{-1} C_i P^M A^\top A P^M C_i^\top R_i^{-1}}}{1 - \|AP^M(AP^M A^\top + Q)^{-1}\|_2} > 0$ are used for notational brevity. Note that we safely ignored the case $b_i = 0$, since from Proposition 4.1 and the definition of matrix spectral norm, $b_i = 0$ if and only if the steady-state Kalman filter gain \bar{K}^i corresponding to sensor i is zero, which implies that the consideration of sensor i will affect neither the estimation error covariance nor the size of the set of estimation means. To solve this problem, we further consider the following equivalent representation:

$$\begin{aligned} \max_{p_1, p_2, \dots, p_M} \quad & \sum_{i=1}^M (p_i + \sqrt{\underline{\eta}^i})^2 \\ \text{s.t.} \quad & \sum_{i=1}^M b_i p_i \leq q, \\ & p_i \geq 0, \quad i = 1, 2, \dots, M, \end{aligned} \tag{4.60}$$

where $q = \bar{x}^* - \sum_{i=1}^M b_i \sqrt{\underline{\eta}^i}$. Notice that this problem is feasible if and only if $q \geq 0$, which should be taken as the guideline in choosing the specifications of $\underline{\eta}^i$ and \bar{x}^* in problem (4.59). Since this problem is a maximization problem of a positive semidefinite quadratic function over a polytope, the optimal solution is at one of the vertices, which are composed by the origin $p_i = 0$ and points of the form $p_i = q/b_i$, $p_j = 0$ for $j \neq i$ and $i, j \in \mathbb{N}_{1:M}$ for this case. Let $i^* = \arg \max_{i \in \mathbb{N}_{1:M}} q/b_i + \sqrt{\underline{\eta}^i}$. Since $b_i > 0$ and $\eta_i > 0$, the optimal value function of this problem equals $(q/b_{i^*} + \sqrt{\underline{\eta}^{i^*}})^2 + \sum_{j=1, j \neq i^*}^M \underline{\eta}^j$ with optimizer $p_{i^*} = q/b_{i^*}$, $p_i = 0$ for $i \neq i^*$. This implies that the set of optimal parameters should be chosen as

$$\text{Tr } \bar{Y}^i = \begin{cases} \underline{\eta}^i, & \text{if } i \neq i^*; \\ \left(\sqrt{\underline{\eta}^i} + \left(\bar{x}^* - \sum_{j=1, j \neq i^*}^M b_j \sqrt{\underline{\eta}^j} \right) / b_i \right)^2, & \text{if } i = i^*. \end{cases} \tag{4.61}$$

Based on the value of $\text{Tr } \bar{Y}^i$, \bar{Y}^i can be chosen to satisfy further requirements, e.g., relative importance of different sensor channels. For the case of $m = 1$, \bar{Y}^i reduces to a positive scalar, then the analysis here provides a complete parameter design procedure.

4.6.2 Examples

Example 4.2. In this example, we apply the set-valued estimation approach to the scenario of event-based state estimation with one sensor and interpret the difference of

the obtained results from the existing results applicable to the same scenario [64, 73]. Consider a second-order system with parameter matrices

$$A = \begin{bmatrix} 0.5 & 0.3 \\ -0.1 & 0.8 \end{bmatrix}, \quad Q = \begin{bmatrix} 0.202 & 0.053 \\ 0.053 & 0.136 \end{bmatrix}, \quad C_1 = [0 \ 1]$$

and $R_1 = 0.2$. To make a comparison with the existing results, we consider the following specific parameterization of $\bar{\mathcal{Y}}_k^1$ in (4.55), which is equivalent to the event-triggering conditions considered in [73] for the case $m = 1$:

$$\bar{Y}_k^1 = \delta^2(C_1\tilde{P}_k^1C_1^\top + R_1), \quad \bar{c}_k^1 = C_1\tilde{x}_k^0,$$

where δ denotes the event-triggering level, \tilde{P}_k^1 and \tilde{x}_k^0 denote the estimation error covariance and the optimal prediction of the event-based MMSE estimator (see [73] for further details), respectively. In this scenario, the time-varying parameters \bar{c}_k^1 and \bar{Y}_k^1 have to be transmitted to the estimator to update the set of estimation means at non-event instants. For this system, $\|\bar{A}\|_2 = 0.51$, and thus the boundedness of the size of the set of estimation means can be guaranteed by Theorem 4.2 without introducing the linear transformations. Two other approaches are also applied according to this triggering condition, including the event-based MMSE estimator [73] and the Kalman filter with intermittent observations [64]. To consider the performance of the estimators under different average communication rates, the estimators are implemented for δ equal to 0.5 and 1.5, the resultant average communication rates of which equal 0.636 and 0.163, respectively. The estimation error plots are shown in Fig. 4.3 and Fig. 4.4.

It is observed when the average communication rate is relatively high ($\delta = 0.5$), the size of the set of estimation means of the set-valued estimator is small, and the performance in terms of estimation error of the set-valued estimator can be characterized by the centre of the set of the estimation means. Under a lower average communication rate, however, the effect of separate parameterization of stochastic and non-stochastic uncertainty becomes more apparent. The exploration of set-valued information as non-stochastic uncertainty leads to a set of estimates with the same filtering gain that contains the estimates corresponding to all point-valued measurements lying in the event-triggering sets during non-event instants, including the MMSE estimate obtained by using the exact point-valued sensor measurements for all time instants (namely, the Kalman filter with periodic observations). In this case, it is not possible to tell which one in the set is associated with the smallest estimation error (without knowing the real state). The alternative answer, however, is that the centre of the set-valued estimator always serves as a point-valued estimate with the best

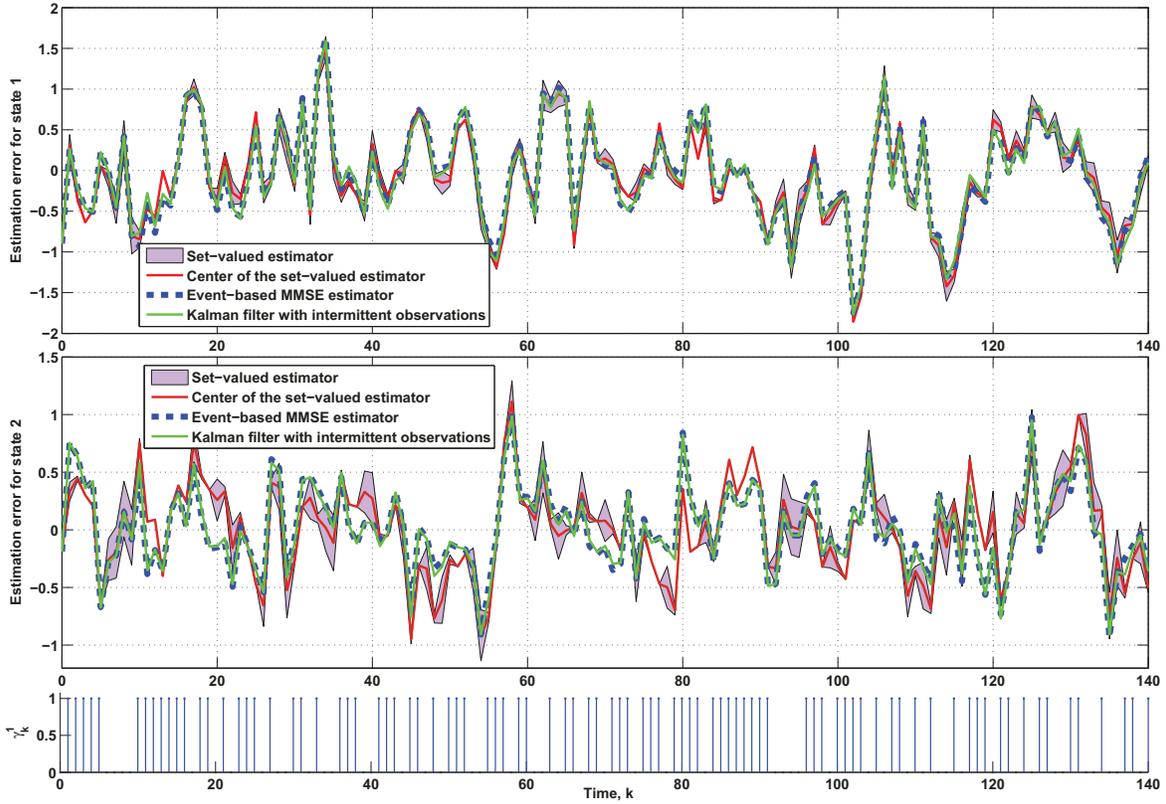


Figure 4.3: Performance comparison of the different estimation strategy for $\delta = 0.5$ (the sets of estimation means are calculated by projecting the two-dimensional ellipsoids on one dimension).

robustness performance, in the sense that it has the smallest worst-case distance to the Kalman filter with periodic observations. This worst-case distance is known to be bounded (Theorem 4.2) and based on Remark 4.2, the asymptotic upper bounds are calculated as 0.4526 and 1.3577 for δ being 0.5 and 1.5, respectively. On the other hand, the precision of the Gaussian assumptions of the non-Gaussian distributions are normally not possible to be verified, which is the basic motivation and theoretic benefit of utilizing the set-valued estimation approach. Finally, it is interesting to note that in both cases, the centre of the set-valued estimator, which can be viewed as a point-valued estimator, achieves similar performance in terms of estimation error as that of the event-based MMSE estimator and Kalman filter with intermittent observations.

Example 4.3. In this example, we apply the developed event-trigger parameter design procedure in Section 4.6.1 to a third-order system, which is obtained by discretizing the benchmark model for a three-blade horizontal-axis turbine with a full converter

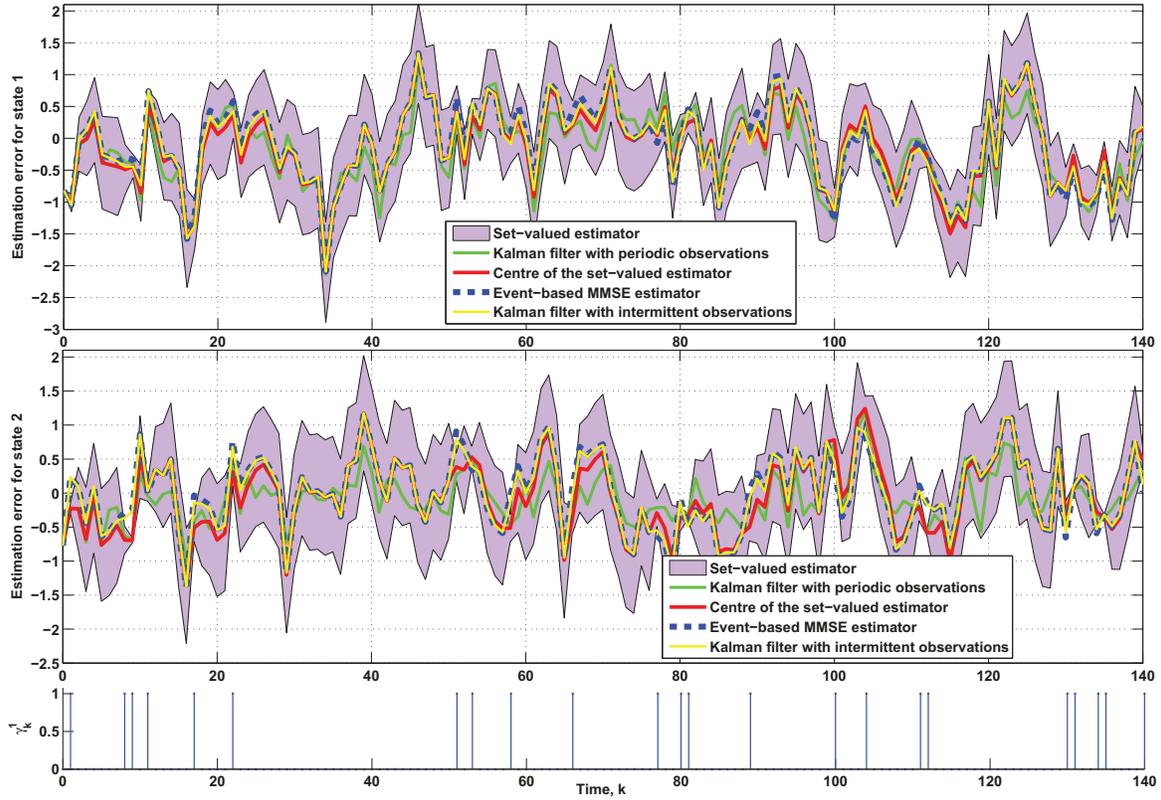


Figure 4.4: Performance comparison of the different estimation strategy for $\delta = 1.5$ (the sets of estimation means are calculated by projecting the two-dimensional ellipsoids on one dimension).

coupling [53] with sampling time $T_s = 2.5\text{s}$ and including a system noise term:

$$x_{k+1} = \begin{bmatrix} 0.9 & 0 & -1.5 \\ 66.1 & 0.3 & 2103.6 \\ 0 & 0 & 0.2 \end{bmatrix} x_k + \begin{bmatrix} 0 & -4.1 \\ 4 & -464.7 \\ 0 & 0 \end{bmatrix} u_k + w_k.$$

The input signal is generated according to the data provided in [53]. Four sensors are used to measure the state information:

$$\begin{aligned} y_k^1 &= [1 \ 0 \ 0] x_k + v_k^1 \\ y_k^2 &= [1 \ 0 \ 0] x_k + v_k^2 \\ y_k^3 &= [0 \ 0.1 \ 0] x_k + v_k^3 \\ y_k^4 &= [0 \ 0.1 \ 0] x_k + v_k^4 \end{aligned}$$

with measurement noise covariances $R_1 = 0.03$, $R_2 = 0.05$, $R_3 = 0.17$ and $R_4 = 0.18$,

respectively, and the system noise covariance is

$$Q = \begin{bmatrix} 0.2023 & 0.0530 & 0 \\ 0.0530 & 0.1360 & 0 \\ 0 & 0 & 0.1000 \end{bmatrix}.$$

We consider the “send-on-delta” triggering conditions [44], namely,

$$\gamma_k^i = \begin{cases} 0, & \text{if } y_k^i \in \bar{\mathcal{Y}}_k^i \\ 1, & \text{if } y_k^i \notin \bar{\mathcal{Y}}_k^i \end{cases} \quad (4.62)$$

where $\bar{\mathcal{Y}}_k^i = \left\{ y \in \mathbb{R}^m \mid (y - y_{\tau_k^i}^i)^\top (\bar{Y}^i)^{-1} (y - y_{\tau_k^i}^i) \leq 1 \right\}$, τ_k^i denoting the last time instant when the measurement of sensor i is transmitted. In this case, no communication is needed during the non-event instants, since \bar{Y}^i 's are constant and $y_{\tau_k^i}^i$'s are always known to the estimator. For this system, $\|\bar{A}\|_2 = 2103.6$. To guarantee the boundedness of the set of estimation means, we calculate the linear transformation

$$T = \begin{bmatrix} 80 & -0.1 & 974 \\ 0 & 1 & -380 \\ 0 & 0 & 1823.3 \end{bmatrix}$$

according to the proof of Theorem 4.2 and apply the estimation procedure to the transformed system. Furthermore, for problem (4.59), the b_i coefficients are calculated as $b_1 = 3.6049 \times 10^8$, $b_2 = 3.1543 \times 10^8$, $b_3 = 4.0788 \times 10^7$ and $b_4 = 3.852 \times 10^7$, respectively. The values for η_i 's are specified as $\eta_1 = 0.5$, $\eta_2 = 0.4$, $\eta_3 = 30$, $\eta_4 = 28$, and $\bar{x}^* = 9 \times 10^8$. The event-triggering conditions are calculated according to equation (4.61) as $\bar{Y}_1 = 0.5$, $\bar{Y}_2 = 0.4$, $\bar{Y}_3 = 35.1317$ and $\bar{Y}_4 = 28$. The set-valued event-based estimator is then implemented and the estimation performance is shown in Fig. 4.5, which is obtained by applying inverse transformation T^{-1} to the estimates. The plot of sensor transmissions are shown in Fig. 4.6, where the average communication rates for the four sensors equal 0.577, 0.632, 0.950, and 0.955, respectively. From Fig. 4.5, it is observed that bounded envelopes for the estimates are always obtained, and the centers of the ellipsoids also serve as efficient point-valued estimates for the state variables. Notice that although the constraint in (4.58) guarantees worst-case performance, it also implicitly helps to control the transient performance. Another way to quantify the transient performance is to consider probabilistic performance constraints (combined with the average communication rates), which is the topic of our future work.

4.7 Summary

In this work, the properties of set-valued Kalman filters with multiple sensor measurements are explored, which help provide further insights on event-based state esti-

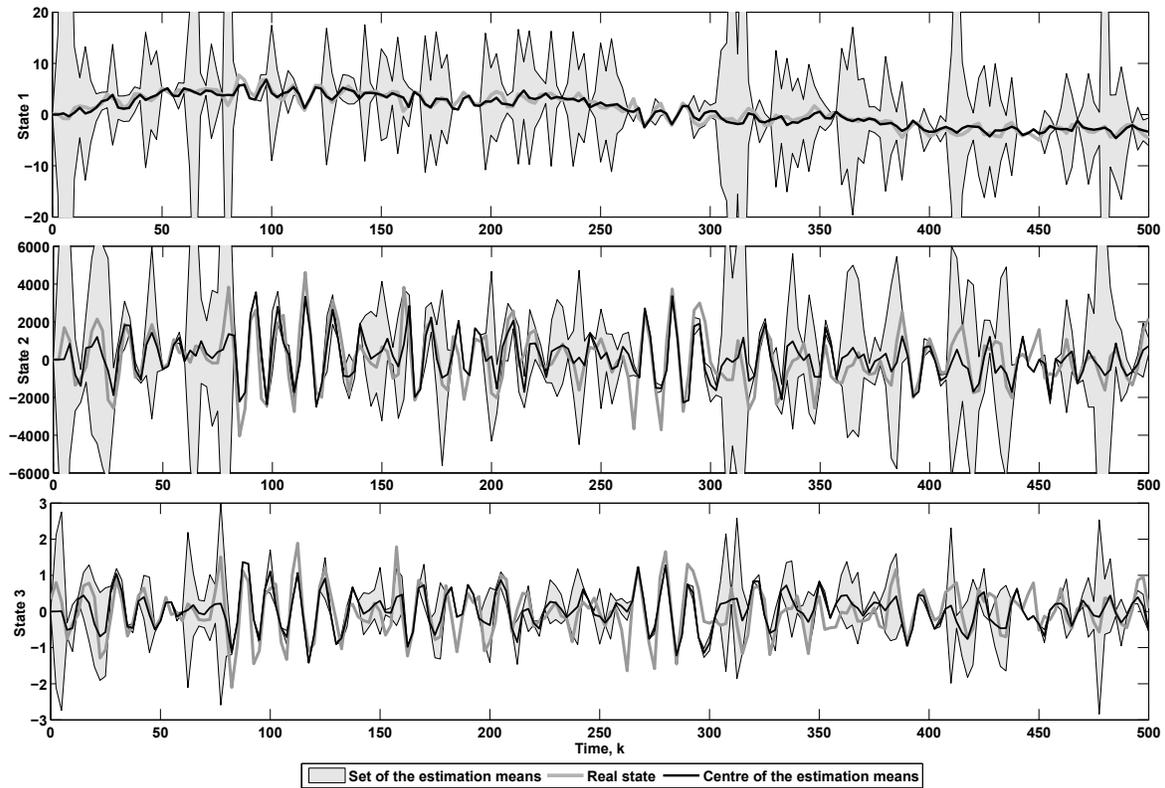


Figure 4.5: Performance of the set-valued state estimation strategy (the envelopes are calculated by projecting the three-dimensional ellipsoids on one dimension).

mation. Despite the distinct nature of the filter, it is shown that the important features of the classic Kalman filter, namely, the invariance of the estimation performance with respect to fusion sequences, the asymptotic boundedness of the performance measures (under certain assumptions, e.g., detectability and stabilizability), are maintained by both the exact set-valued filter and the proposed approximate set-valued filter. On the other hand, we show that the inclusion of more sensors does not necessarily reduce the size of the set of estimation means, and certain conditions need to be satisfied to guarantee performance improvement, which is utilized to formulate design problems in event-based estimation.

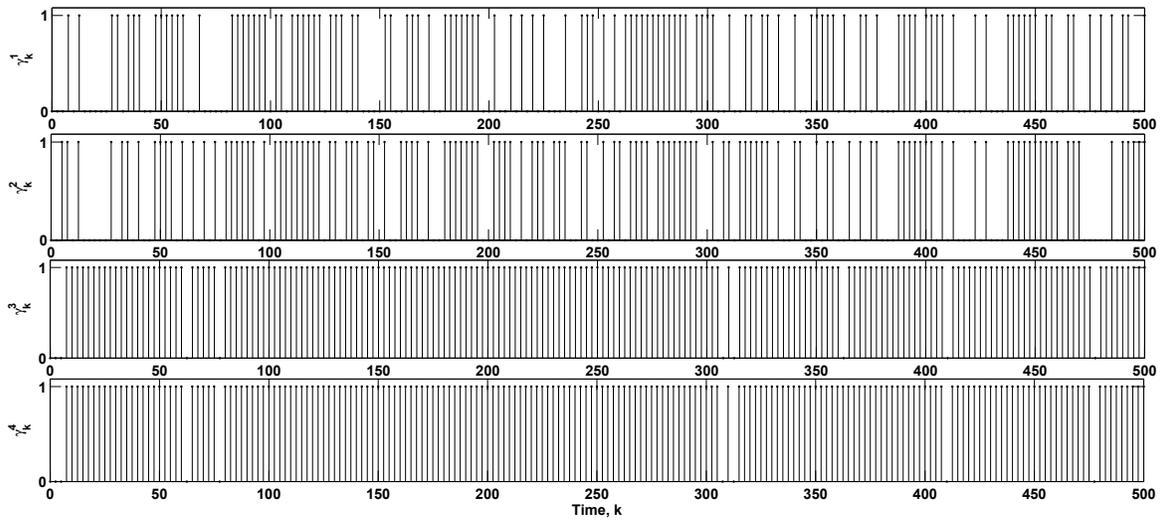


Figure 4.6: Plot of sensor transmissions.

Chapter 5

Event-based state estimation of discrete-state hidden Markov models*

5.1 Introduction

In this chapter, we consider the event-based remote estimation problem for discrete-state HMMs. The motivation stems from the fact that HMMs represent the individual component states of a dynamic system in a natural way [2]. This type of model has been extensively used in speech signal processing (see [23, 56] and references therein), and has also found applications in various areas of engineering, e.g., teleoperation systems [27], smart grid [76], machinery system monitoring [24], and motor systems [74]. We assume that the measurement updates are transmitted to the remote estimator through wired or wireless communication networks, and the estimation problem is considered for both the reliable communication channel scenario and the scenario that the channel is associated with a packet dropout process. Differently from existing results in the literature of event-based estimation, an alternative approach, the change of probability measure approach introduced in [17], is utilized to solve the event-based state estimation problems. In this approach, the problems are not directly solved under the original “real-world” probability measure; alternatively, the idea utilized is to consider the estimation problems under new probability measures, construct maps

*Parts of the results in this chapter were submitted to IEEE Transactions on Automatic Control, 2014.

from the new measures to the “real-world” measure such that the problems considered under different measures are equivalent, and finally map the results from the new measures to the “real-world” probability measure, which solves the original problems. The main results are summarized as follows.

1. For both the reliable communication channel and unreliable channel scenarios, closed-form expressions for the evolution of the probability distribution of the states conditioned on the available hybrid measurement information of the estimator have been developed. The unnormalized probability distributions introduced under the new probability measures are shown to evolve recursively according to linear maps. Also, we show that the estimation problem for a lossy communication channel, but without the event trigger, can be treated as a special case of the reliable channel results.
2. Based on the expressions for the conditional probability distributions, analytical results for communication rate analysis are obtained for both the reliable and unreliable communication channel scenarios. The expected communication rates are expressed in terms of the ratio of the weighted 1-norms of the conditional probability distributions of the states versus their 1-norms.
3. Implementation issues of the developed results are discussed. The relationship of the results and the real-valued finite-state HMMs is analyzed, and the expressions for MMSE estimates are presented. The effectiveness of results are illustrated by a numerical example and comparative simulations, and it is shown that improved estimation performance is always obtained by the proposed results through the exploitation of the information contained in the event-triggering sets.

Notation: Let \mathbb{N} denote the set of nonnegative integers. Write $\mathbb{N}_{1:M} := \{1, 2, \dots, M\}$ and $\mathbb{Z}_{-M,N} := \{-M, -M + 1, \dots, N\}$. Let $\mathbb{I} := \{i_1, i_2, \dots, i_N\} \subset \mathbb{N}$ be a set of indices. We use $[x_i]_{i \in \mathbb{I}}$ to denote $[x_{i_1}^\top, \dots, x_{i_N}^\top]^\top$. For a set \mathbb{M} , let $|\mathbb{M}|$ be its cardinality. For a probability measure \mathbf{P} (or $\hat{\mathbf{P}}$ and $\check{\mathbf{P}}$), we use \mathbf{E} (or $\hat{\mathbf{E}}$ and $\check{\mathbf{E}}$, respectively) to represent the expectation operator. For a vector $v = [v_i]_{i \in \mathbb{N}_{1:n}} \in \mathbb{R}^n$, we denote $\|v\|_1$ as its 1-norm, which is defined as $\|v\|_1 = \sum_{i=1}^n \|v_i\|$, where $\|v_i\|$ is the absolute value of v_i .

5.2 Problem Description

Firstly, we introduce a hidden Markov model on the probability space $(\Omega, \mathcal{F}, \mathbf{P})$. The hidden process considered is a finite-state, homogeneous, discrete-time Markov

chain X . Assume the initial state X_0 is given. Suppose the cardinality of the state space of X_k is N , then the state space S_X can be identified with

$$S_X = \{e_1, e_2, \dots, e_N\},$$

where e_i is the unit vector in \mathbb{R}^N with the i th element equal to 1. Let $\mathcal{F}_k^0 := \sigma\{X_0, \dots, X_k\}$, and let $\{\mathcal{F}_k\}$ be the complete filtration generated by \mathcal{F}_k^0 . By the Markov property,

$$\mathbf{P}(X_{k+1} = e_j | \mathcal{F}_k) = \mathbf{P}(X_{k+1} = e_j | X_k). \quad (5.1)$$

Let

$$a_{i,j} = \mathbf{P}(X_{k+1} = e_i | X_k = e_j), \quad A = (a_{i,j}) \in \mathbb{R}^{N \times N}, \quad (5.2)$$

then

$$\mathbf{E}(X_{k+1} | \mathcal{F}_k) = \mathbf{E}(X_{k+1} | X_k) = AX_k. \quad (5.3)$$

Define $V_{k+1} := X_{k+1} - AX_k$, so we have the state space equation

$$X_{k+1} = AX_k + V_{k+1}. \quad (5.4)$$

Note that V_{k+1} is a vector $(\mathbf{P}, \mathcal{F}_k)$ martingale increment. Let Y_k be a sensor measurement process of X_k , which takes values in a finite-state space. Let the cardinality of the state space S_Y of Y be M , then S_Y can be identified with

$$\{f_1, f_2, \dots, f_M\},$$

with f_i the unit vector in \mathbb{R}^M with the i th element equal to 1. Write

$$C = [c_i]_{i \in \mathbb{N}_{1:M}}, \quad c_i = [c_{i,j}]_{j \in \mathbb{N}_{1:N}}^\top, \quad (5.5)$$

where

$$c_{i,j} = \mathbf{P}(Y_{k+1} = f_i | X_k = e_j), \quad (5.6)$$

so that $\sum_{i=1}^M c_{i,j} = 1$ and $c_{i,j} \geq 0$. Therefore

$$\mathbf{E}(Y_{k+1} | X_k) = CX_k. \quad (5.7)$$

Define $W_{k+1} := Y_{k+1} - CX_k$, so we have the measurement equation

$$Y_{k+1} = CX_k + W_{k+1}. \quad (5.8)$$

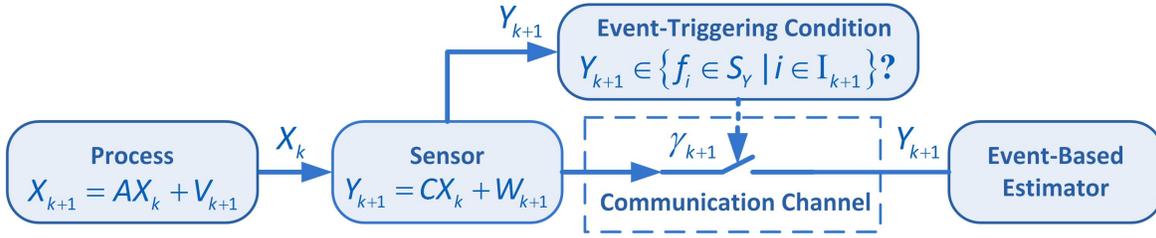


Figure 5.1: Block diagram of the event-based remote estimation system with a reliable communication channel.

Note that the discrete HMM described in (5.4) and (5.8) is similar to the one considered in Chapter 2 of [17]. Let \mathcal{G}_k be the completion of the σ -field on Ω generated by X_0, X_1, \dots, X_k and Y_1, Y_2, \dots, Y_k , and let \mathcal{Y}_k be the completion of the σ -field on Ω generated by Y_1, Y_2, \dots, Y_k . Note that

$$\mathbf{E}(Y_{k+1}|\mathcal{G}_k) = \mathbf{E}(Y_{k+1}|X_k) = CX_k, \quad (5.9)$$

so that

$$\mathbf{E}(W_{k+1}|\mathcal{G}_k) = \mathbf{E}(W_{k+1}|X_k) = CX_k - CX_k = 0, \quad (5.10)$$

implying W_{k+1} is a $(\mathbf{P}, \mathcal{G}_k)$ martingale increment.

Now we introduce the remote estimation problem (see Fig. 5.1), where the state of the HMM is estimated by a remote estimator, based on the measurement information from the sensor through a wired/wireless communication channel. Currently, we assume the communication channel is reliable. We consider the scenario that the sensor measures Y_k at every time instant k , but does not always send the value of Y_k to the estimator. At each time instant k , a set of indices $\mathbb{I}_k := \{i_1, \dots, i_{\ell_k}\} \subseteq \mathbb{N}_{1:M}$ is specified with $|\mathbb{I}_k| = \ell_k$. If $Y_k \in \{f_i \in S_Y \mid i \in \mathbb{I}_k\}$, then the value of Y_k is not sent; otherwise the value of Y_k is sent via the communication channel and the remote estimator knows the value of Y_k . Note that due to the event-triggering condition, the estimator still has the information $Y_k \in \{f_i \in S_Y \mid i \in \mathbb{I}_k\}$ even when the exact value of Y_k is not received. To model explicitly the event-triggering condition, we define a new random variable γ_k as

$$\gamma_k = \begin{cases} 1, & \text{if } Y_k \in \{f_i \mid i \notin \mathbb{I}_k\} \\ 0, & \text{otherwise} \end{cases} \quad (5.11)$$

which indicates whether a transmission is performed by the sensor at time instant k .

Let g_i be the unit vector in \mathbb{R}^{M+1} with the i th element equal to 1. To model the event-based measurement information of the estimator, we introduce a new variable

\hat{Y}_k taking values in $\{g_1, g_2, \dots, g_{M+1}\}$. This is defined as $\hat{Y}_k := D_k Y_k$, where $D_k = [d_{1,k}^\top, \dots, d_{M+1,k}^\top]^\top \in \mathbb{R}^{(M+1) \times M}$, and $d_{i,k}$ satisfies

$$d_{i,k} = \begin{cases} f_i^\top, & \text{for } i \notin \mathbb{I}_k, \\ 0, & \text{for } i \in \mathbb{I}_k, \\ \sum_{j \in \mathbb{I}_k} f_j^\top, & \text{for } i = M+1. \end{cases} \quad (5.12)$$

Let $\hat{Y}_k = [\hat{Y}_k^1, \dots, \hat{Y}_k^{M+1}]^\top$. From the definition of D_k , the state space of \hat{Y}_k is composed of the standard unit vectors

$$\{g_i \in \mathbb{R}^{(M+1) \times 1} \mid i \in \mathbb{N}_{1:M+1} \setminus \mathbb{I}_k\}, \quad (5.13)$$

and

$$\hat{Y}_k = \begin{cases} g_{M+1}, & \text{for } Y_k \in \{f_i \in S_Y \mid i \in \mathbb{I}_k\}; \\ g_i, & \text{for } Y_k = f_i, i \in \mathbb{N}_{1:M} \setminus \mathbb{I}_k. \end{cases} \quad (5.14)$$

Now we have

$$\hat{Y}_{k+1} = D_{k+1} C X_k + D_{k+1} W_{k+1} \quad (5.15)$$

$$:= \hat{C}_{k+1} X_k + \hat{W}_{k+1}, \quad (5.16)$$

where $\hat{C}_{k+1} := D_{k+1} C \in \mathbb{R}^{(M+1) \times N}$ and $\hat{W}_{k+1} := D_{k+1} W_{k+1}$. Note that since \hat{C}_{k+1} is the transition matrix from X_k to \hat{Y}_{k+1} ,

$$\mathbf{E}(\hat{W}_{k+1} | \hat{\mathcal{G}}_k) = \mathbf{E}(\hat{Y}_{k+1} - \hat{C}_{k+1} X_k | \hat{\mathcal{G}}_k) = 0, \quad (5.17)$$

which means that \hat{W}_{k+1} is a $(\mathbf{P}, \hat{\mathcal{G}}_k)$ martingale increment. Let $\hat{\mathcal{G}}_k$ be the completion of the σ -field on Ω generated by X_0, X_1, \dots, X_k and $\hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_k$, and let $\hat{\mathcal{Y}}_k$ be the completion of the σ -field on Ω generated by $\hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_k$. Let $\mathbb{I}_k := \mathbb{N}_{1:M+1} \setminus \mathbb{I}_k$. Write

$$\hat{C}_{k+1} = [\hat{c}_{i,k+1}]_{i \in \mathbb{N}_{1:M+1}}, \text{ and } \hat{c}_{i,k+1} = [\hat{c}_{i,j,k+1}]_{j \in \mathbb{N}_{1:N}}^\top. \quad (5.18)$$

From the definition of \hat{C}_{k+1} , we have

$$\hat{c}_{i,j,k+1} = \begin{cases} c_{i,j}, & \text{for } i \notin \mathbb{I}_{k+1} \\ 0, & \text{for } i \in \mathbb{I}_{k+1} \\ \sum_{i \in \mathbb{I}_k} c_{i,j}, & \text{for } i = M+1. \end{cases} \quad (5.19)$$

Note that γ_k now satisfies

$$\gamma_k = \begin{cases} 1, & \text{if } \hat{Y}_k \in \{g_i \mid i \in \mathbb{N}_{1:M} \setminus \mathbb{I}_k\}; \\ 0, & \text{otherwise.} \end{cases} \quad (5.20)$$

Before continuing, we summarize the actual model we need to consider under \mathbf{P} obtained through the above discussion:

$$\begin{aligned} X_{k+1} &= AX_k + V_{k+1}, \\ \hat{Y}_{k+1} &= \hat{C}_{k+1}X_k + \hat{W}_{k+1}, \\ \mathbf{E}(V_{k+1}|\mathcal{F}_k) &= 0, \\ \mathbf{E}(\hat{W}_{k+1}|\hat{\mathcal{G}}_k) &= 0. \end{aligned} \tag{5.21}$$

The objective is to estimate the distribution of X_k based on the measurement information $\hat{\mathcal{Y}}_k$.

5.3 Results on reliable communication channel

Now we consider a new probability measure $\hat{\mathbf{P}}$, under which we still have

$$X_{k+1} = AX_k + V_{k+1}, \tag{5.22}$$

with V_{k+1} a martingale ($\hat{\mathbf{P}}, \mathcal{F}_k$) martingale increment, namely,

$$\hat{\mathbf{E}}(V_{k+1}|\mathcal{F}_k) = 0, \tag{5.23}$$

but the \hat{Y}_{k+1} are uniformly distributed independent random variables satisfying

$$\hat{\mathbf{P}}(\hat{Y}_{k+1} = g_i|\hat{\mathcal{G}}_k) = \hat{\mathbf{P}}(\hat{Y}_{k+1} = g_i) = \frac{1}{M+1}. \tag{5.24}$$

Now we transform the probability measure $\hat{\mathbf{P}}$ to \mathbf{P} by defining the following map:

$$\left. \frac{d\mathbf{P}}{d\hat{\mathbf{P}}} \right|_{\hat{\mathcal{G}}_{k+1}} = \hat{\Lambda}_{k+1}, \tag{5.25}$$

where

$$\hat{\Lambda}_{k+1} := \prod_{l=1}^{k+1} \hat{\lambda}_l, \tag{5.26}$$

$$\hat{\lambda}_{k+1} := (M+1) \sum_{i=1}^{M+1} \langle \hat{C}_{k+1}X_k, g_i \rangle \langle \hat{Y}_{k+1}, g_i \rangle. \tag{5.27}$$

To map the model in (5.22)-(5.24) from probability measure $\hat{\mathbf{P}}$ to probability measure \mathbf{P} , we recall the following result in [17].

Lemma 5.1 (Theorem 3.2 in Chapter 2 of [17]). *Suppose $(\Omega, \mathcal{F}, \mathbf{P})$ is a probability space and $\mathcal{G} \subset \mathcal{F}$ is a sub- σ -field. Suppose $\hat{\mathbf{P}}$ is another probability measure absolutely continuous with respect to \mathbf{P} and with Radon-Nikodym derivative $\frac{d\hat{\mathbf{P}}}{d\mathbf{P}} = \hat{\Lambda}$. If ϕ is any \mathbf{P} integrable random variable, then*

$$\mathbf{E}(\phi|\mathcal{G}) = \begin{cases} \frac{\hat{\mathbf{E}}(\hat{\Lambda}\phi|\mathcal{G})}{\hat{\mathbf{E}}(\hat{\Lambda}|\mathcal{G})}, & \text{if } \hat{\mathbf{E}}(\hat{\Lambda}|\mathcal{G}) > 0; \\ 0, & \text{otherwise.} \end{cases} \quad (5.28)$$

Recall that a sequence $\{\phi_k\}$ is said to be $\hat{\mathcal{G}}$ -adapted if ϕ_k is $\hat{\mathcal{G}}_k$ -measurable for all k [17]. Since $\hat{\mathcal{Y}}_k$ is a sub- σ -field of $\hat{\mathcal{G}}_k$, we have the following result.

Lemma 5.2. *If $\{\phi_k\}$ is a $\hat{\mathcal{G}}$ -adapted integrable sequence of random variables, then*

$$\mathbf{E}(\phi_k|\hat{\mathcal{Y}}_k) = \frac{\hat{\mathbf{E}}(\hat{\Lambda}_k\phi_k|\hat{\mathcal{Y}}_k)}{\hat{\mathbf{E}}(\hat{\Lambda}_k|\hat{\mathcal{Y}}_k)}. \quad (5.29)$$

Based on the above lemmas, we have the following result.

Lemma 5.3. *If the model in (5.22) – (5.24) is mapped from probability measure $\hat{\mathbf{P}}$ to probability measure \mathbf{P} via (5.25), then the obtained model has the following properties under measure \mathbf{P} :*

$$\mathbf{E}(X_{k+1}|\mathcal{F}_k) = AX_k, \quad (5.30)$$

$$\mathbf{E}(\hat{Y}_{k+1}|\hat{\mathcal{G}}_k) = \hat{C}_{k+1}X_k. \quad (5.31)$$

Proof. From the definition of $\hat{\lambda}_{k+1}$,

$$\hat{\mathbf{E}}(\hat{\lambda}_{k+1}|\hat{\mathcal{G}}_k) = \hat{\mathbf{E}} \left[(M+1) \sum_{i=1}^{M+1} \langle \hat{C}_{k+1}X_k, g_i \rangle \langle \hat{Y}_{k+1}, g_i \rangle | \hat{\mathcal{G}}_k \right] \quad (5.32)$$

$$= (M+1) \sum_{i=1}^{M+1} \langle \hat{C}_{k+1}X_k, g_i \rangle \hat{\mathbf{E}}(\langle \hat{Y}_{k+1}, g_i \rangle | \hat{\mathcal{G}}_k) \quad (5.33)$$

$$= \sum_{i=1}^{M+1} \langle \hat{C}_{k+1}X_k, g_i \rangle = 1. \quad (5.34)$$

Now we have

$$\mathbf{E}(X_{k+1}|\hat{\mathcal{G}}_k) = \frac{\hat{\mathbf{E}}(\hat{\Lambda}_{k+1}X_{k+1}|\hat{\mathcal{G}}_k)}{\hat{\mathbf{E}}(\hat{\Lambda}_{k+1}|\hat{\mathcal{G}}_k)} \quad (5.35)$$

$$= \frac{\hat{\Lambda}_k \hat{\mathbf{E}}(\hat{\lambda}_{k+1}X_{k+1}|\hat{\mathcal{G}}_k)}{\hat{\Lambda}_k \hat{\mathbf{E}}(\hat{\lambda}_{k+1}|\hat{\mathcal{G}}_k)} \quad (5.36)$$

$$= \hat{\mathbf{E}}(\hat{\lambda}_{k+1}X_{k+1}|\hat{\mathcal{G}}_k) \quad (5.37)$$

$$= \hat{\mathbf{E}}(\hat{\lambda}_{k+1}|\hat{\mathcal{G}}_k) \hat{\mathbf{E}}(AX_k + V_{k+1}|\hat{\mathcal{G}}_k) \quad (5.38)$$

$$= AX_k, \quad (5.39)$$

where the last equality follows from $\hat{\mathbf{E}}(\hat{\lambda}_{k+1}|\hat{\mathcal{G}}_k) = 1$ and $\hat{\mathbf{E}}(V_{k+1}|\hat{\mathcal{G}}_k) = 0$, as $\hat{\mathbf{E}}(V_{k+1}|\mathcal{F}_k) = 0$ and V_{k+1} and \hat{Y}_k are mutually independent. By repeated conditioning (Lemma 1.11 in [16]), this further implies

$$\mathbf{E}(X_{k+1}|\mathcal{F}_k) = \mathbf{E}[\mathbf{E}(X_{k+1}|\hat{\mathcal{G}}_k)|\mathcal{F}_k] = AX_k, \quad (5.40)$$

as $\mathcal{F}_k \subset \hat{\mathcal{G}}_k$. Similarly,

$$\mathbf{E}(\hat{Y}_{k+1}|\hat{\mathcal{G}}_k) = \frac{\hat{\mathbf{E}}(\hat{\Lambda}_{k+1}\hat{Y}_{k+1}|\hat{\mathcal{G}}_k)}{\hat{\mathbf{E}}(\hat{\Lambda}_{k+1}|\hat{\mathcal{G}}_k)} \quad (5.41)$$

$$= \hat{\mathbf{E}}(\hat{\lambda}_{k+1}\hat{Y}_{k+1}|\hat{\mathcal{G}}_k) \quad (5.42)$$

$$= \hat{\mathbf{E}}[\hat{Y}_{k+1}(M+1) \sum_{i=1}^{M+1} \langle \hat{C}_{k+1}X_k, g_i \rangle \langle \hat{Y}_{k+1}, g_i \rangle | \hat{\mathcal{G}}_k] \quad (5.43)$$

$$= \left[\hat{\mathbf{E}}(\langle \hat{Y}_{k+1}, g_j \rangle (M+1) \sum_{i=1}^{M+1} \langle \hat{C}_{k+1}X_k, g_i \rangle \langle \hat{Y}_{k+1}, g_i \rangle | \hat{\mathcal{G}}_k) \right]_{j \in \mathbb{N}_{1:M+1}} \quad (5.44)$$

$$= \hat{C}_{k+1}X_k, \quad (5.45)$$

which completes the proof. \square

Note that the above result implies that under probability measure \mathbf{P} , the relationships $X_{k+1} = AX_k + V_{k+1}$ with V_{k+1} being a $(\mathbf{P}, \mathcal{F}_k)$ martingale, and $\hat{Y}_{k+1} = \hat{C}_{k+1}X_k + \hat{W}_{k+1}$ with \hat{W}_{k+1} being a $(\mathbf{P}, \hat{\mathcal{G}}_k)$ martingale are recovered. Therefore, we can solve the original estimation problem for (5.21) under \mathbf{P} by solving the problem for (5.22) – (5.24) under $\hat{\mathbf{P}}$ and mapping the result back to \mathbf{P} , which allows us to obtain a simple expression on \mathbf{P} for the estimator, as will be shown in the following.

5.3.1 Results on recursive estimation

For $r \in \mathbb{N}_{1:N}$, define \hat{q}_k^r and \hat{p}_k^r as

$$\hat{q}_k^r := \hat{\mathbf{E}} \left[\hat{\Lambda}_k \langle X_k, e_r \rangle \middle| \hat{\mathcal{Y}}_k \right], \quad (5.46)$$

$$\hat{p}_k^r := \mathbf{E} \left[\langle X_k, e_r \rangle \middle| \hat{\mathcal{Y}}_k \right], \quad (5.47)$$

respectively. Write $\hat{q}_k = [\hat{q}_k^1, \dots, \hat{q}_k^N]^\top$ and $\hat{p}_k = [\hat{p}_k^1, \dots, \hat{p}_k^N]^\top$, so that from Lemma 5.2

$$\hat{p}_k^r = \hat{q}_k^r / \|\hat{q}_k\|_1. \quad (5.48)$$

In this way, \hat{p}_k represents the probability distribution of X_k conditioned on $\hat{\mathcal{Y}}_k$. For the recursive estimation of X_k conditioned on $\hat{\mathcal{Y}}_k$, we shall derive the recursive evolution of \hat{q}_k^r , which is presented in the following result.

Theorem 5.1. For $k \in \mathbb{N}$, $r \in \mathbb{N}_{1:M}$,

$$\hat{q}_{k+1}^r = (M+1) \sum_{j=1}^N a_{r,j} \hat{q}_k^j \sum_{i \in \mathbb{L}_{k+1}} \hat{c}_{i,j,k+1} \hat{Y}_{k+1}^i. \quad (5.49)$$

Proof. In the following, we first present the main derivations, and then provide the detailed explanations. For $r \in \mathbb{N}_{1:N}$, we have

$$\hat{q}_{k+1}^r = \hat{\mathbf{E}}[\langle X_{k+1}, e_r \rangle \hat{\Lambda}_{k+1} | \hat{\mathcal{Y}}_{k+1}] \quad (5.50)$$

$$= \hat{\mathbf{E}} \left[\langle AX_k + V_{k+1}, e_r \rangle \hat{\Lambda}_k \sum_{i=1}^{M+1} (M+1) \langle \hat{C}_{k+1} X_k, g_i \rangle \langle \hat{Y}_{k+1}, g_i \rangle \middle| \hat{\mathcal{Y}}_{k+1} \right] \quad (5.51)$$

$$= (M+1) \hat{\mathbf{E}} \left[\langle AX_k, e_r \rangle \hat{\Lambda}_k \sum_{i=1}^{M+1} \langle \hat{C}_{k+1} X_k, g_i \rangle \langle \hat{Y}_{k+1}, g_i \rangle \middle| \hat{\mathcal{Y}}_{k+1} \right] \quad (5.52)$$

$$= (M+1) \sum_{j=1}^N a_{r,j} \hat{\mathbf{E}} \left[\langle X_k, e_j \rangle \hat{\Lambda}_k \sum_{i=1}^{M+1} \hat{c}_{i,j,k+1} \langle \hat{Y}_{k+1}, g_i \rangle \middle| \hat{\mathcal{Y}}_{k+1} \right] \quad (5.53)$$

$$= (M+1) \sum_{j=1}^N a_{r,j} \hat{\mathbf{E}} \left[\langle X_k, e_j \rangle \hat{\Lambda}_k \middle| \hat{\mathcal{Y}}_{k+1} \right] \hat{\mathbf{E}} \left[\sum_{i=1}^{M+1} \hat{c}_{i,j,k+1} \langle \hat{Y}_{k+1}, g_i \rangle \middle| \hat{\mathcal{Y}}_{k+1} \right] \quad (5.54)$$

$$= (M+1) \sum_{j=1}^N a_{r,j} \hat{\mathbf{E}} \left[\langle X_k, e_j \rangle \hat{\Lambda}_k \middle| \hat{\mathcal{Y}}_k \right] \hat{\mathbf{E}} \left[\sum_{i=1}^{M+1} \hat{c}_{i,j,k+1} \langle \hat{Y}_{k+1}, g_i \rangle \middle| \hat{Y}_{k+1} \right] \quad (5.55)$$

$$= (M+1) \sum_{j=1}^N a_{r,j} \hat{\mathbf{E}} \left[\langle X_k, e_j \rangle \hat{\Lambda}_k \middle| \hat{\mathcal{Y}}_k \right] \sum_{i \in \mathbb{L}_{k+1}} \hat{c}_{i,j,k+1} \langle \hat{Y}_{k+1}, g_i \rangle \quad (5.56)$$

$$= (M+1) \sum_{j=1}^N a_{r,j} \hat{q}_k^j \sum_{i \in \mathbb{L}_{k+1}} \hat{c}_{i,j,k+1} \langle \hat{Y}_{k+1}, g_i \rangle \quad (5.57)$$

Now we present the necessary explanations for the above derivations.

1. Equation (5.51) \rightarrow equation (5.52). It suffices to show

$$\hat{\mathbf{E}} \left[\langle V_{k+1}, e_r \rangle \hat{\Lambda}_k \sum_{i=1}^{M+1} \langle \hat{C}_{k+1} X_k, g_i \rangle \langle \hat{Y}_{k+1}, g_i \rangle \middle| \hat{\mathcal{Y}}_{k+1} \right] = 0. \quad (5.58)$$

To see this, we have

$$\hat{\mathbf{E}} \left[\langle V_{k+1}, e_r \rangle \hat{\Lambda}_k \sum_{i=1}^{M+1} \langle \hat{C}_{k+1} X_k, g_i \rangle \langle \hat{Y}_{k+1}, g_i \rangle \middle| \hat{\mathcal{Y}}_{k+1} \right] \quad (5.59)$$

$$= \hat{\mathbf{E}} \left[\hat{\mathbf{E}} \left[\langle V_{k+1}, e_r \rangle \hat{\Lambda}_k \sum_{i=1}^{M+1} \langle \hat{C}_{k+1} X_k, g_i \rangle \langle \hat{Y}_{k+1}, g_i \rangle \middle| \hat{\mathcal{G}}_k, \hat{Y}_{k+1} \right] \middle| \hat{\mathcal{Y}}_{k+1} \right] \quad (5.60)$$

$$= \hat{\mathbf{E}} \left[\hat{\mathbf{E}} \left[\langle V_{k+1}, e_r \rangle \middle| \hat{\mathcal{G}}_k, \hat{Y}_{k+1} \right] \hat{\mathbf{E}} \left[\hat{\Lambda}_k \sum_{i=1}^M \langle \hat{C}_{k+1} X_k, g_i \rangle \langle \hat{Y}_{k+1}, g_i \rangle \middle| \hat{\mathcal{G}}_k, \hat{Y}_{k+1} \right] \middle| \hat{\mathcal{Y}}_{k+1} \right]. \quad (5.61)$$

Now we show $\hat{\mathbf{E}}[\langle V_{k+1}, e_r \rangle \middle| \hat{\mathcal{G}}_k, \hat{Y}_{k+1}] = 0$. Since X_k and \hat{Y}_k are mutually independent and $V_{k+1} = X_{k+1} - X_k$, V_k and \hat{Y}_k are mutually independent as well. Thus we have

$$\hat{\mathbf{E}} \left[\langle V_{k+1}, e_r \rangle \middle| \hat{\mathcal{G}}_k, \hat{Y}_{k+1} \right] = \hat{\mathbf{E}} \left[\langle V_{k+1}, e_r \rangle \middle| \mathcal{F}_k \right] = 0. \quad (5.62)$$

2. Equation (5.52) \rightarrow equation (5.53). This follows from the fact that $X_k \in \{e_1, \dots, e_N\}$, and that for $X_k \in \{e_1, \dots, e_N\}$

$$\langle AX_k, e_r \rangle \langle X_k, e_j \rangle = a_{r,j} \langle X_k, e_j \rangle. \quad (5.63)$$

3. Equation (5.53) \rightarrow equation (5.54). Since \hat{Y}_k are i.i.d. uniformly distributed under $\hat{\mathbf{P}}$, X is independent of \hat{Y} . Thus conditioned on $\hat{\mathcal{Y}}_{k+1}$, $\langle X_k, e_j \rangle \hat{\Lambda}_k$ and $\sum_{i=1}^{M+1} \hat{c}_{i,j,k+1} \langle \hat{Y}_{k+1}, g_i \rangle$ are independent.

4. Equation (5.54) \rightarrow equation (5.55). This follows from the definition of $\hat{\Lambda}_k$ and the fact that \hat{Y}_k are i.i.d. random variables.

5. Equation (5.55) \rightarrow equation (5.56). This follows from the fact that $\hat{c}_{i,j,k+1} = 0$ for $i \in \mathbb{I}_{k+1}$ (see equation (5.19)).

□

Let $\hat{o}_k := \left[(M+1) \sum_{i \in \mathbb{L}_k} \hat{c}_{i,j,k} \hat{Y}_k^i \right]_{j \in \mathbb{N}_{1:N}} \in \mathbb{R}^N$, then equation (5.49) is equivalent to

$$\hat{q}_{k+1} = A \text{diag}(\hat{o}_{k+1}) \hat{q}_k. \quad (5.64)$$

5.3.2 Communication rate analysis

In this section, the target is to analyze the probability of sensor transmission at time instant $k + 1$, based on the hybrid measurement information $\hat{\mathcal{Y}}_k$. According to the definition of γ_{k+1} , the average communication rate at time instant $k + 1$ is given by $\mathbf{P}[\gamma_{k+1} = 1 | \hat{\mathcal{Y}}_k]$. Note that the value of γ_{k+1} is determined by \hat{Y}_{k+1} , and that $\sum_{j=1}^{M+1} \langle \hat{Y}_{k+1}, g_j \rangle = 1$. We need to evaluate $\mathbf{E}[\langle \hat{Y}_{k+1}, g_j \rangle | \hat{\mathcal{Y}}_k]$, for which we have the following result.

Lemma 5.4.

$$\mathbf{E}[\langle \hat{Y}_{k+1}, g_i \rangle | \hat{\mathcal{Y}}_k] = \sum_{j=1}^N [\hat{c}_{i,j,k+1} \hat{q}_k^j / \sum_{l=1}^N \hat{q}_k^l].$$

Proof.

$$\begin{aligned} & \mathbf{E} [\langle \hat{Y}_{k+1}, g_i \rangle | \hat{\mathcal{Y}}_k] \\ &= \mathbf{E} [\langle \hat{C}_{k+1} X_k + \hat{W}_{k+1}, g_i \rangle | \hat{\mathcal{Y}}_k] \end{aligned} \quad (5.65)$$

$$= \mathbf{E} [\langle \hat{C}_{k+1} X_k, g_i \rangle | \hat{\mathcal{Y}}_k] + \mathbf{E} [\mathbf{E} [\langle \hat{W}_{k+1}, g_i \rangle | \hat{\mathcal{G}}_k] | \hat{\mathcal{Y}}_k] \quad (5.66)$$

$$= \mathbf{E} [\langle \hat{C}_{k+1} X_k, g_i \rangle | \hat{\mathcal{Y}}_k] \quad (5.67)$$

$$= \mathbf{E} [\sum_{j=1}^N \langle \hat{C}_{k+1} e_j, g_i \rangle \langle X_k, e_j \rangle | \hat{\mathcal{Y}}_k] \quad (5.68)$$

$$= \sum_{j=1}^N \hat{c}_{i,j,k+1} \mathbf{E}[\langle X_k, e_j \rangle | \hat{\mathcal{Y}}_k] \quad (5.69)$$

$$= \sum_{j=1}^N [\hat{c}_{i,j,k+1} \hat{q}_k^j] / \sum_{l=1}^N \hat{q}_k^l, \quad (5.70)$$

where equation (5.66) is due to $\hat{\mathcal{Y}}_k \subset \hat{\mathcal{G}}_k$, equation (5.67) is due to

$$\mathbf{E} [\hat{W}_{k+1} | \hat{\mathcal{G}}_k] = 0,$$

and equation (5.70) follows from Theorem 5.1. \square

Let $\hat{O}_k = [c_i]_{i \in \mathbb{N}_{1:M} \setminus \mathbb{I}_k} \in \mathbb{R}^{(M-\ell_k) \times N}$. Based on the above expression for $\mathbf{E}[\langle Y_{k+1}, f_j \rangle | \hat{\mathcal{Y}}_k]$, we have the following result on communication rate analysis.

Theorem 5.2. *For the state estimation scheme in Fig. 5.1 and the event-triggering condition in (5.20), the expected sensor to estimator communication rate $\mathbf{P}(\gamma_{k+1} = 1 | \hat{\mathcal{Y}}_k)$ is given by*

$$\mathbf{P}(\gamma_{k+1} = 1 | \hat{\mathcal{Y}}_k) = \|\hat{O}_{k+1} \hat{q}_k\|_1 / \|\hat{q}_k\|_1.$$

Proof. From the definition of γ_{k+1} , we have

$$\begin{aligned} & \mathbf{P}(\gamma_{k+1} = 1 | \hat{\mathcal{Y}}_k) \\ &= \sum_{i \in \mathbb{N}_{1:M} \setminus \mathbb{I}_{k+1}} \mathbf{P}(\hat{Y}_{k+1} = g_i | \hat{\mathcal{Y}}_k) \end{aligned} \quad (5.71)$$

$$= \sum_{i \in \mathbb{N}_{1:M} \setminus \mathbb{I}_{k+1}} \mathbf{E} \left[\langle \hat{Y}_{k+1}, g_i \rangle | \hat{\mathcal{Y}}_k \right] \quad (5.72)$$

$$= \sum_{i \in \mathbb{N}_{1:M} \setminus \mathbb{I}_{k+1}} \sum_{j=1}^N \left[\hat{c}_{i,j,k+1} \hat{q}_k^j / \sum_{l=1}^N \hat{q}_k^l \right] \quad (5.73)$$

$$= \sum_{i \in \mathbb{N}_{1:M} \setminus \mathbb{I}_{k+1}} \sum_{j=1}^N \left[c_{i,j} \hat{q}_k^j / \sum_{l=1}^N \hat{q}_k^l \right], \quad (5.74)$$

where the last equality is due to equation (5.19). The conclusion follows from the definition of \hat{O}_k and the non-negativeness of \hat{q}_k^i and $c_{i,j}$. \square

5.3.3 Estimation with a lossy communication channel: a special case

In this subsection, we show that the results obtained in Section 5.3.1 are rich enough to cover the scenario of remote estimation with an unreliable communication channel but no event-trigger (see Fig. 5.2). At this stage, we only require that the

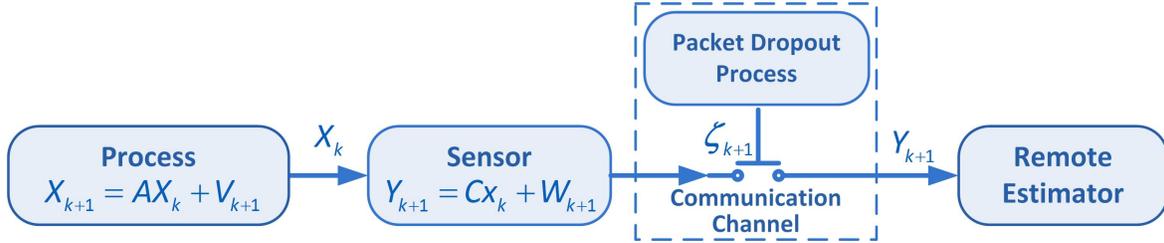


Figure 5.2: Block diagram of the periodic estimation problem with a lossy communication channel.

packet dropout process ζ_k is independent of X_k and Y_k , but do not need the specific model of it. In this case, if no information is received by the remote estimator at time k , it would only know $Y_k \in S_Y$; otherwise it knows the exact value of Y_k . In this case, $\mathbb{I}_k = \mathbb{N}_{1:M}$ when the value of Y_k is not received, and $\mathbb{I}_k = \{i \in \mathbb{N}_{1:M} | Y_k = f_i\}$ otherwise. Based on this specific definition of \mathbb{I}_k , the index set \mathbb{L}_k is generated, and thus the updating equations for the probability distribution of X_k , conditioned on the received measurement information, can be obtained by applying Theorem 5.1 for the obtained \mathbb{L}_k . We shall use this result for comparative simulation to illustrate how the estimation performance is improved by exploiting the information contained in the event-triggering set in the numerical verification section.

5.4 Results on unreliable communication channel

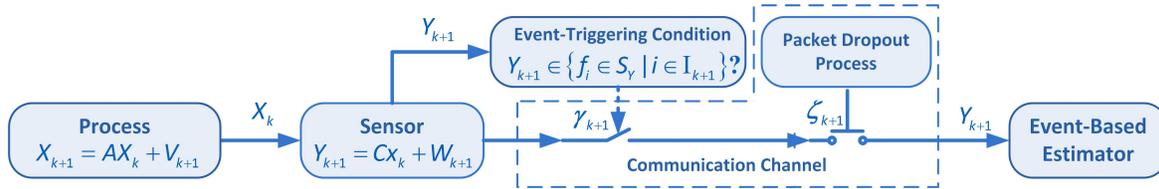


Figure 5.3: Block diagram of the event-based remote estimation system with an unreliable communication channel.

Now we consider the scenario that a packet dropout process is associated with the communication channel. We start from the model for Y_k and X_k that we have under \mathbf{P} for the reliable communication channel case:

$$X_{k+1} = AX_k + V_{k+1}, \quad (5.75)$$

$$Y_{k+1} = CX_k + W_{k+1}, \quad (5.76)$$

$$\mathbf{E}(V_{k+1} | \mathcal{F}_k) = 0, \quad (5.77)$$

$$\mathbf{E}(W_{k+1} | \mathcal{G}_k) = 0. \quad (5.78)$$

To analyze the effect of the unreliable channel, we follow the standard technique in handling packet loss [60, 64] and introduce a packet dropout process ζ_k , which is an i.i.d. Bernoulli process under \mathbf{P} satisfying

$$\mathbf{P}(\zeta_k = 1) = \mathbf{P}(\zeta_k = 1 | \mathcal{G}_k, \Gamma_{k-1}) = \lambda, \quad (5.79)$$

$$\mathbf{P}(\zeta_k = 0) = \mathbf{P}(\zeta_k = 0 | \mathcal{G}_k, \Gamma_{k-1}) = 1 - \lambda. \quad (5.80)$$

Here Γ_{k-1} denotes the completion of the σ -field on Ω generated by $\zeta_1, \zeta_2, \dots, \zeta_{k-1}$. If $\zeta_k = 1$, then the value of Y_k sent by the sensor is received by the estimator; otherwise a packet dropout occurs at time instant k and thus the value of Y_k sent by the sensor is not received by the estimator. Note that this packet dropout process only takes effect during the event instants (namely, $\gamma_k = 1$), because the sensor would not send the measurement to the remote estimator otherwise. In this way, the measurement information at the estimator side when Y_k is not received becomes more ambiguous due to the packet drop process: if $\zeta_k = 1$, then the estimator knows $Y_k \in \{f_i \in S_Y | i \in \mathbb{I}_k\}$; otherwise (i.e., when $\zeta_k = 0$) it would infer that $Y_k \in \{f_i \in S_Y | i \notin \mathbb{I}_k\}$.

We shall use $\check{\mathcal{G}}_k$ to denote the completion of the σ -field on Ω generated by Γ_k and \mathcal{G}_k , namely,

$$\check{\mathcal{G}}_k = \Gamma_k \cup \mathcal{G}_k. \quad (5.81)$$

Note that what the remote estimator can observe is (1) Y_k is received; (2) Y_k is not received. Both these two types of observations are measurable in $\check{\mathcal{G}}_k$, as they can be represented by the combination of the events of Y_k and ζ_k . To be specific, “ Y_k is received” implies $\{Y_k = f_i\} \cap \{\zeta_k = 1\}$ for some $i \notin \mathbb{I}_k$; “ Y_k is not received” implies that

$$\left\{ \left\{ \bigcup_{i \notin \mathbb{I}_k} \{Y_k = f_i\} \right\} \cap \{\zeta_k = 0\} \right\} \cup \left\{ \bigcup_{i \in \mathbb{I}_k} \{Y_k = f_i\} \right\}. \quad (5.82)$$

For notational simplicity, we introduce a new variable \check{Y}_k to denote the received information at the sensor:

$$\check{Y}_k = \begin{cases} g_i, & \text{if } \{Y_k = f_i\} \cap \{\zeta_k = 1\}, \quad i \in \mathbb{N}_{1:M}; \\ g_{M+1}, & \text{if } \left\{ \left\{ \bigcup_{i \notin \mathbb{I}_k} \{Y_k = f_i\} \right\} \cap \{\zeta_k = 0\} \right\} \cup \left\{ \bigcup_{i \in \mathbb{I}_k} \{Y_k = f_i\} \right\}. \end{cases} \quad (5.83)$$

We shall use $\check{\mathcal{Y}}_k$ to denote the completion of the σ -field on Ω generated by $\check{Y}_1, \dots, \check{Y}_k$. From the above discussion, obviously we have $\check{\mathcal{Y}}_k \subset \check{\mathcal{G}}_k$.

Now we consider a new probability measure $\check{\mathbf{P}}$, under which we still have

$$X_{k+1} = AX_k + V_{k+1}, \quad (5.84)$$

with V_{k+1} a $(\check{\mathbf{P}}, \mathcal{F}_k)$ martingale increment, namely,

$$\check{\mathbf{E}}(V_{k+1} | \mathcal{F}_k) = 0. \quad (5.85)$$

but the Y_{k+1} are uniformly distributed independent random variables satisfying

$$\check{\mathbf{P}}(Y_{k+1} = f_i | \check{\mathcal{G}}_k) = \check{\mathbf{P}}(Y_{k+1} = f_i) = \frac{1}{M}. \quad (5.86)$$

In addition, we have a model for ζ_k :

$$\check{\mathbf{P}}(\zeta_{k+1} = 1 | \check{\mathcal{G}}_k) = \check{\mathbf{P}}(\zeta_{k+1} = 1) = \lambda, \quad (5.87)$$

$$\check{\mathbf{P}}(\zeta_{k+1} = 0 | \check{\mathcal{G}}_k) = \check{\mathbf{P}}(\zeta_{k+1} = 0) = 1 - \lambda, \quad (5.88)$$

which implies that ζ_k is an i.i.d. Bernoulli process that is independent of X_k and Y_k .

Again we establish a one-way map from $\check{\mathbf{P}}$ to \mathbf{P} over $\check{\mathcal{G}}_k$:

$$\left. \frac{d\mathbf{P}}{d\check{\mathbf{P}}} \right|_{\check{\mathcal{G}}_k} = \check{\Lambda}_k, \quad (5.89)$$

where

$$\check{\lambda}_{k+1} = \sum_{i=1}^M M \langle CX_k, f_i \rangle \langle Y_{k+1}, f_i \rangle, \quad \check{\Lambda}_k = \prod_{l=1}^k \check{\lambda}_l. \quad (5.90)$$

Note that the $\check{\lambda}_k$ defined here is not related to ζ_k . This trick allows us to keep its properties when mapping it from $\check{\mathbf{P}}$ to \mathbf{P} .

Recall that a sequence $\{\phi_k\}$ is said to be $\check{\mathcal{G}}$ -adapted if ϕ_k is $\check{\mathcal{G}}_k$ -measurable for all k , [17]. Since $\check{\mathcal{Y}}_k$ is a sub- σ -field of $\check{\mathcal{G}}_k$, we have the following result.

Lemma 5.5. *If $\{\phi_k\}$ is a $\check{\mathcal{G}}$ -adapted integrable sequence of random variables, then*

$$\mathbf{E}(\phi_k | \check{\mathcal{Y}}_k) = \frac{\check{\mathbf{E}}(\check{\Lambda}_k \phi_k | \check{\mathcal{Y}}_k)}{\check{\mathbf{E}}(\check{\Lambda}_k | \check{\mathcal{Y}}_k)}. \quad (5.91)$$

Based on the above lemma, we have the following result.

Lemma 5.6. *If the model in (5.84) – (5.88) is mapped from probability measure $\check{\mathbf{P}}$ to probability measure \mathbf{P} via (5.90), then the obtained model has the following properties under measure \mathbf{P} :*

$$\mathbf{E}(X_{k+1} | \mathcal{F}_k) = AX_k, \quad (5.92)$$

$$\mathbf{E}(Y_{k+1} | \mathcal{G}_k) = CX_k. \quad (5.93)$$

$$\mathbf{E}(\zeta_{k+1} | \check{\mathcal{G}}_k) = \mathbf{E}(\zeta_{k+1}) = \lambda. \quad (5.94)$$

Proof. From the definition of $\check{\lambda}_{k+1}$,

$$\check{\mathbf{E}}(\check{\lambda}_{k+1} | \check{\mathcal{G}}_k) = \check{\mathbf{E}} \left[M \sum_{i=1}^M \langle CX_k, f_i \rangle \langle Y_{k+1}, f_i \rangle | \check{\mathcal{G}}_k \right] \quad (5.95)$$

$$= M \sum_{i=1}^M \langle CX_k, f_i \rangle \check{\mathbf{E}}(\langle Y_{k+1}, f_i \rangle | \check{\mathcal{G}}_k) \quad (5.96)$$

$$= \sum_{i=1}^M \langle CX_k, f_i \rangle = 1. \quad (5.97)$$

Now we have

$$\mathbf{E}(X_{k+1} | \check{\mathcal{G}}_k) = \frac{\check{\mathbf{E}}(\check{\Lambda}_{k+1} X_{k+1} | \check{\mathcal{G}}_k)}{\check{\mathbf{E}}(\check{\Lambda}_{k+1} | \check{\mathcal{G}}_k)} \quad (5.98)$$

$$= \frac{\check{\Lambda}_k \check{\mathbf{E}}(\check{\lambda}_{k+1} X_{k+1} | \check{\mathcal{G}}_k)}{\check{\Lambda}_k \check{\mathbf{E}}(\check{\lambda}_{k+1} | \check{\mathcal{G}}_k)} \quad (5.99)$$

$$= \check{\mathbf{E}}(\check{\lambda}_{k+1} X_{k+1} | \check{\mathcal{G}}_k) \quad (5.100)$$

$$= \check{\mathbf{E}}(\check{\lambda}_{k+1} | \check{\mathcal{G}}_k) \check{\mathbf{E}}(AX_k + V_{k+1} | \check{\mathcal{G}}_k) \quad (5.101)$$

$$= AX_k, \quad (5.102)$$

where the last equality follows from $\check{\mathbf{E}}(\check{\lambda}_{k+1}|\check{\mathcal{G}}_k) = 1$ and $\check{\mathbf{E}}(V_{k+1}|\check{\mathcal{G}}_k) = 0$, as $\check{\mathbf{E}}(V_{k+1}|\mathcal{F}_k) = 0$ and V_{k+1} and \check{Y}_k are mutually independent. By repeated conditioning (Lemma 1.11 in [16]), this further implies

$$\mathbf{E}(X_{k+1}|\mathcal{F}_k) = \mathbf{E}[\mathbf{E}(X_{k+1}|\check{\mathcal{G}}_k)|\mathcal{F}_k] = AX_k, \quad (5.103)$$

as $\mathcal{F}_k \subset \check{\mathcal{G}}_k$. Similarly,

$$\mathbf{E}(Y_{k+1}|\check{\mathcal{G}}_k) = \frac{\check{\mathbf{E}}(\check{\Lambda}_{k+1}Y_{k+1}|\check{\mathcal{G}}_k)}{\check{\mathbf{E}}(\check{\Lambda}_{k+1}|\check{\mathcal{G}}_k)} \quad (5.104)$$

$$= \check{\mathbf{E}}(\check{\lambda}_{k+1}Y_{k+1}|\check{\mathcal{G}}_k) \quad (5.105)$$

$$= \check{\mathbf{E}}[Y_{k+1}M \sum_{i=1}^M \langle CX_k, f_i \rangle \langle Y_{k+1}, f_i \rangle | \check{\mathcal{G}}_k] \quad (5.106)$$

$$= \left[\check{\mathbf{E}}(\langle Y_{k+1}, f_j \rangle M \sum_{i=1}^M \langle CX_k, f_i \rangle \langle Y_{k+1}, f_i \rangle | \check{\mathcal{G}}_k) \right]_{j \in \mathbb{N}_{1:M}} \quad (5.107)$$

$$= CX_k. \quad (5.108)$$

Finally,

$$\mathbf{E}(\zeta_{k+1}|\check{\mathcal{G}}_k) = \frac{\check{\mathbf{E}}(\check{\Lambda}_{k+1}\zeta_{k+1}|\check{\mathcal{G}}_k)}{\check{\mathbf{E}}(\check{\Lambda}_{k+1}|\check{\mathcal{G}}_k)} \quad (5.109)$$

$$= \check{\mathbf{E}}(\check{\lambda}_{k+1}\zeta_{k+1}|\check{\mathcal{G}}_k) \quad (5.110)$$

$$= \check{\mathbf{E}}[\zeta_{k+1}M \sum_{i=1}^M \langle CX_k, f_i \rangle \langle Y_{k+1}, f_i \rangle | \check{\mathcal{G}}_k] \quad (5.111)$$

$$= \check{\mathbf{E}}[\zeta_{k+1}|\check{\mathcal{G}}_k] \check{\mathbf{E}}[M \sum_{i=1}^M \langle CX_k, f_i \rangle \langle Y_{k+1}, f_i \rangle | \check{\mathcal{G}}_k] \quad (5.112)$$

$$= \lambda, \quad (5.113)$$

where the fact that ζ_k is independent of Y_k and X_k is used to obtain (5.112), and (5.113) is due to that X_k takes values in $\{e_1, \dots, e_N\}$, that $\sum_{i=1}^M c_{i,j} = 1$ and that Y_{k+1} and X_k are mutually independent. \square

The above results indicate that we are able to recover exactly the model defined under \mathbf{P} with the newly defined reverse map $\check{\Lambda}_k$. With this new probability measure and change of measure, we obtain the results for recursive estimation and communication rate analysis, as is shown below.

5.4.1 Results on recursive estimation

Write $\check{q}_k = [\check{q}_k^1, \dots, \check{q}_k^N]$ and $\check{p}_k = [\check{p}_k^1, \dots, \check{p}_k^N]$, where for $r \in \mathbb{N}_{1:N}$, \check{q}_k^r and \check{p}_k^r are defined as

$$\check{q}_k^r := \check{\mathbf{E}}[\check{\Lambda}_k \langle X_k, e_r \rangle | \check{\mathcal{Y}}_k], \quad (5.114)$$

$$\check{p}_k^r := \mathbf{E}[\langle X_k, e_r \rangle | \check{\mathcal{Y}}_k], \quad (5.115)$$

respectively, so that we still have

$$\check{p}_k^r = \check{q}_k^r / \|\check{q}_k\|_1. \quad (5.116)$$

Also, for notational simplicity, write

$$\check{C}_k = [\check{c}_{i,k}]_{i \in \mathbb{N}_{1:M+1}}, \quad \check{c}_{i,k} = [\check{c}_{i,j,k}]_{j \in \mathbb{N}_{1:N}}^\top, \quad (5.117)$$

where

$$\check{c}_{i,j,k} = \begin{cases} c_{i,j}, & \text{for } i \in \mathbb{N}_{1:M}; \\ \frac{1}{M - \lambda(M - |\mathbb{L}_k|)} [\sum_{i \in \mathbb{L}_k} c_{i,j} + \sum_{i \notin \mathbb{L}_k} (1 - \lambda)c_{i,j}], & \text{for } i = M + 1. \end{cases} \quad (5.118)$$

For the recursive estimation of X_k conditioned on $\check{\mathcal{Y}}_k$, we shall derive the recursive evolution of \check{q}_k , which is presented in the following result.

Theorem 5.3. *For $k \in \mathbb{N}$, \check{q}_k evolves according to the following recursive form:*

$$\check{q}_{k+1} = A \text{diag}(\check{c}_{k+1}) \check{q}_k, \quad (5.119)$$

where $\check{c}_k := [M \sum_{i \in \mathbb{L}_k} \check{c}_{i,j,k} \langle \check{Y}_k, g_i \rangle]_{j \in \mathbb{N}_{1:N}} \in \mathbb{R}^N$.

Proof. In the following, we first present the main derivations, and then provide the detailed explanations. For $r \in \mathbb{N}_{1:N}$, we have

$$\check{q}_{k+1}^r = \check{\mathbf{E}}[\langle X_{k+1}, e_r \rangle \check{\Lambda}_{k+1} | \check{\mathcal{Y}}_{k+1}] \quad (5.120)$$

$$= M \sum_{j=1}^N a_{r,j} \check{\mathbf{E}}[\langle X_k, e_j \rangle \check{\Lambda}_k | \check{\mathcal{Y}}_k] \check{\mathbf{E}}\left[\sum_{i=1}^M c_{i,j} \langle Y_{k+1}, f_i \rangle \middle| \check{Y}_{k+1}\right] \quad (5.121)$$

$$= M \sum_{j=1}^N a_{r,j} \check{\mathbf{E}}[\langle X_k, e_j \rangle \check{\Lambda}_k | \check{\mathcal{Y}}_k] \sum_{i \in \mathbb{L}_{k+1}} \check{c}_{i,j,k} \langle \check{Y}_{k+1}, g_i \rangle \quad (5.122)$$

$$= M \sum_{j=1}^N a_{r,j} \check{q}_k^j \sum_{i \in \mathbb{L}_{k+1}} \check{c}_{i,j,k} \langle \check{Y}_{k+1}, g_i \rangle. \quad (5.123)$$

Now we present the necessary explanations for the above derivations.

1. The derivation from equation (5.120) to equation (5.121) follows a similar argument as that for equation (5.50) \rightarrow equation (5.55) and thus is not discussed considering the space limitation.

2. Equation (5.121) \rightarrow equation (5.122). To see this, we have for $\check{Y}_{k+1} = g_{M+1}$,

$$\begin{aligned} & \check{\mathbf{E}} \left[\sum_{i=1}^M c_{i,j} \langle Y_{k+1}, f_i \rangle \middle| \check{Y}_{k+1} = g_{M+1} \right] \\ &= \sum_{i=1}^M c_{i,j} \check{\mathbf{E}} [\langle Y_{k+1}, f_i \rangle | \check{Y}_{k+1} = g_{M+1}] \end{aligned} \quad (5.124)$$

$$= \sum_{i=1}^M c_{i,j} \frac{1 - \lambda[1 - \mathbf{1}_{\mathbb{I}_{k+1}}(i)]}{M - \lambda(M - |\mathbb{I}_{k+1}|)} \quad (5.125)$$

$$= \frac{1}{M - \lambda(M - |\mathbb{I}_{k+1}|)} \left[\sum_{i \in \mathbb{I}_{k+1}} c_{i,j} + \sum_{i \notin \mathbb{I}_{k+1}} (1 - \lambda)c_{i,j} \right], \quad (5.126)$$

where equation (5.125) is due to

$$\begin{aligned} & \check{\mathbf{E}} (\langle Y_{k+1}, f_i \rangle | \check{Y}_{k+1} = g_{M+1}) \\ &= \frac{\check{\mathbf{P}}(\{Y_{k+1} = f_i\} \cap \{\check{Y}_{k+1} = g_{M+1}\})}{\check{\mathbf{P}}(\check{Y}_{k+1} = g_{M+1})} \\ &= \left[\check{\mathbf{P}} \left(\{Y_{k+1} = f_i\} \cap \left\{ \left\{ \bigcup_{j \notin \mathbb{I}_{k+1}} \{Y_{k+1} = f_j\} \cap \{\zeta_{k+1} = 0\} \right\} \right. \right. \right. \\ & \quad \left. \left. \cup \left\{ \bigcup_{l \in \mathbb{I}_{k+1}} \{Y_{k+1} = f_l\} \right\} \right) \right] \\ & \quad / \left[\check{\mathbf{P}} \left(\left\{ \bigcup_{j \notin \mathbb{I}_{k+1}} \{Y_{k+1} = f_j\} \cap \{\zeta_{k+1} = 0\} \right\} \cup \left\{ \bigcup_{l \in \mathbb{I}_{k+1}} \{Y_{k+1} = f_l\} \right\} \right) \right] \\ &= \left[\check{\mathbf{P}}(\zeta_{k+1} = 0) \check{\mathbf{P}} \left(\{Y_{k+1} = f_i\} \cap \left\{ \bigcup_{j \notin \mathbb{I}_{k+1}} \{Y_{k+1} = f_j\} \right\} \right) \right. \\ & \quad \left. + \check{\mathbf{P}} \left(\{Y_{k+1} = f_i\} \cap \left\{ \bigcup_{l \in \mathbb{I}_{k+1}} \{Y_{k+1} = f_l\} \right\} \right) \right] \\ & \quad / \left[\check{\mathbf{P}}(\zeta_{k+1} = 0) \check{\mathbf{P}} \left(\bigcup_{j \notin \mathbb{I}_{k+1}} \{Y_{k+1} = f_j\} \right) + \check{\mathbf{P}} \left(\bigcup_{l \in \mathbb{I}_{k+1}} \{Y_{k+1} = f_l\} \right) \right] \\ &= \frac{(1 - \mathbf{1}_{\mathbb{I}_{k+1}}(i)) \check{\mathbf{P}}(\zeta_{k+1} = 0) \check{\mathbf{P}}(Y_{k+1} = f_i) + \mathbf{1}_{\mathbb{I}_{k+1}}(i) \check{\mathbf{P}}(Y_{k+1} = f_i)}{\check{\mathbf{P}}(\zeta_{k+1} = 0) \check{\mathbf{P}} \left(\bigcup_{j \notin \mathbb{I}_{k+1}} \{Y_{k+1} = f_j\} \right) + \check{\mathbf{P}} \left(\bigcup_{l \in \mathbb{I}_{k+1}} \{Y_{k+1} = f_l\} \right)} \\ &= \frac{(1 - \lambda) (1 - \mathbf{1}_{\mathbb{I}_{k+1}}(i)) \frac{1}{M} + \mathbf{1}_{\mathbb{I}_{k+1}}(i) \frac{1}{M}}{(1 - \lambda)(M - |\mathbb{I}_{k+1}|) \frac{1}{M} + |\mathbb{I}_{k+1}| \frac{1}{M}} \\ &= \frac{1 - \lambda[1 - \mathbf{1}_{\mathbb{I}_{k+1}}(i)]}{M - \lambda(M - |\mathbb{I}_{k+1}|)}. \end{aligned}$$

For the case of $\check{Y}_{k+1} = g_r$, $r \in \mathbb{N}_{1:M}$, the derivation is much simpler, for which we

have

$$\begin{aligned}
& \check{\mathbf{E}} \left[\sum_{i=1}^M c_{i,j} \langle Y_{k+1}, f_i \rangle \middle| \check{Y}_{k+1} = g_r \right] \\
&= \sum_{i=1}^M c_{i,j} \check{\mathbf{E}} [\langle Y_{k+1}, f_i \rangle | \check{Y}_{k+1} = g_r] \\
&= c_{r,j} \check{\mathbf{P}} (Y_{k+1} = f_r | \check{Y}_{k+1} = g_r) \\
&= c_{r,j} \frac{\check{\mathbf{P}} (\{Y_{k+1} = f_r\} \cap \{\check{Y}_{k+1} = g_r\})}{\check{\mathbf{P}} (\check{Y}_{k+1} = g_r)} \\
&= c_{r,j} \frac{\check{\mathbf{P}} (\{Y_{k+1} = f_r\} \cap \{\check{Y}_{k+1} = f_r\} \cap \{\zeta_{k+1} = 1\})}{\check{\mathbf{P}} (\{Y_{k+1} = f_r\} \cap \{\zeta_{k+1} = 1\})} = c_{r,j}.
\end{aligned}$$

□

5.4.2 Communication rate analysis

The communication rate analysis for the scenario of an unreliable communication channel still focuses on the probability of sensor transmission at time instant $k + 1$, based on the measurement information $\check{\mathcal{Y}}_k$, as the target of this analysis is to evaluate the impact of the triggering condition and packet dropout on the energy consumption of the sensor. For this unreliable communication channel case, the average communication rate at time instant $k + 1$ is given by $\mathbf{P}[\gamma_{k+1} = 1 | \check{\mathcal{Y}}_k]$. Noticing that the value of γ_{k+1} is determined by Y_{k+1} , and that $\sum_{j=1}^M \langle Y_{k+1}, f_j \rangle = 1$, we need to evaluate $\mathbf{E}[\langle Y_{k+1}, f_j \rangle | \check{\mathcal{Y}}_k]$, for which we have the following result.

Lemma 5.7.

$$\mathbf{E}[\langle Y_{k+1}, f_i \rangle | \check{\mathcal{Y}}_k] = \sum_{j=1}^N [c_{i,j} \check{q}_k^j / \sum_{l=1}^N \check{q}_k^l].$$

Proof.

$$\begin{aligned}
& \mathbf{E} [\langle Y_{k+1}, f_i \rangle | \check{\mathcal{Y}}_k] \\
&= \mathbf{E} [\langle CX_k, f_i \rangle | \check{\mathcal{Y}}_k] + \mathbf{E} [\langle W_{k+1}, f_i \rangle | \check{\mathcal{Y}}_k] \tag{5.127}
\end{aligned}$$

$$= \mathbf{E} [\langle CX_k, f_i \rangle | \check{\mathcal{Y}}_k] + \mathbf{E} [\mathbf{E} [\langle W_{k+1}, f_i \rangle | \check{\mathcal{G}}_k] | \check{\mathcal{Y}}_k] \tag{5.128}$$

$$= \mathbf{E} [\langle CX_k, f_i \rangle | \check{\mathcal{Y}}_k] + \mathbf{E} [\mathbf{E} [\langle W_{k+1}, f_i \rangle | \mathcal{G}_k] | \check{\mathcal{Y}}_k] \tag{5.129}$$

$$= \mathbf{E} [\langle CX_k, f_i \rangle | \check{\mathcal{Y}}_k] \tag{5.130}$$

$$= \sum_{j=1}^N [c_{i,j} \check{q}_k^j] / \sum_{l=1}^N \check{q}_k^l, \tag{5.131}$$

where equation (5.128) is due to $\check{\mathcal{Y}}_k \subset \check{\mathcal{G}}_k$, equation (5.129) is due to the fact that $\check{\mathcal{G}}_k = \mathcal{G}_k \cup \Gamma_k$ and that ζ_k and W_k are mutually independent, equation (5.130) is due to $\mathbf{E}[W_{k+1}|\mathcal{G}_k] = 0$, and equation (5.131) follows from a similar argument as that in the proof of Lemma 5.4 (i.e., equation (5.67) \rightarrow equation (5.70)). \square

Based on the above expression for $\mathbf{E}[\langle Y_{k+1}, f_j \rangle | \check{\mathcal{Y}}_k]$, we have the following result on communication rate analysis.

Theorem 5.4. *For the state estimation scheme in Fig. 5.3, the deterministic event-trigger γ_k and the packet dropout process ζ_k , the expected sensor to estimator communication rate $\mathbf{P}(\gamma_{k+1} = 1 | \check{\mathcal{Y}}_k)$ is given by*

$$\mathbf{P}(\gamma_{k+1} = 1 | \check{\mathcal{Y}}_k) = \|\hat{O}_{k+1}\check{q}_k\|_1 / \|\check{q}_k\|_1 .$$

This result follows from a similar argument as that of Theorem 5.2 and so the proof is omitted.

5.5 Numerical implementation and verification

In this section, we discuss the numerical implementation of the proposed event-based estimates and illustrate the efficiency of the results by a numerical example.

5.5.1 Numerical implementation

Firstly, we discuss the numerical implementation of the event-based estimators. From the analysis in Sections 5.3 and 5.4, the estimators for both the reliable and the unreliable communication channels can be written in the form

$$q_{k+1} = \mathcal{A}_k q_k, \tag{5.132}$$

$$p_{k+1} = q_{k+1} / \|q_{k+1}\|_1. \tag{5.133}$$

Here q_k denotes an unnormalized probability distribution under the new measure, p_k denotes the probability distribution under the original measure, and $\mathcal{A}_k \in \mathbb{R}^{N \times N}$ is a time-varying transition matrix that depends on the new measure and the problem considered. Equation (5.132) evolves as a linear time-varying system, which causes numerical problems in recursive implementation. If \mathcal{A}_k is unstable, $\lim_{k \rightarrow \infty} q_k \rightarrow \infty$ and thus the value of q_k will blow up as time goes by. If \mathcal{A}_k is stable, $\lim_{k \rightarrow \infty} q_k = 0$ and in this case it will be difficult to numerically evaluate p_k as $\|q_k\|_1$ approaches 0. Note

that this numerical issue is not related with the correctness of the results developed, as what is actually needed for the state estimate is p_{k+1} , the 1-norm of which always equals 1 by equation (5.133). To overcome this issue, we observe the following result.

Proposition 5.1. *Let \tilde{q}_k and \tilde{p}_k satisfy*

$$\tilde{q}_{k+1} = \mathcal{A}_k \tilde{q}_k / \|\tilde{q}_k\|_1, \quad \tilde{p}_{k+1} = \tilde{q}_{k+1} / \|\tilde{q}_{k+1}\|_1, \quad \tilde{q}_0 = q_0. \quad (5.134)$$

Then $\tilde{p}_k = p_k$ for all $k \in \mathbb{N}$.

Proof. By definition, we have for $k \geq 1$,

$$\tilde{q}_{k+1} = \frac{\prod_{t=0}^k \mathcal{A}_t \tilde{q}_0}{\prod_{t=0}^k \left\| \prod_{l=0}^{t-1} \mathcal{A}_l \tilde{q}_0 \right\|_1}, \quad (5.135)$$

where $\prod_{t=0}^k \mathcal{A}_t$ satisfies $\prod_{t=0}^k \mathcal{A}_t = \mathcal{A}_k \prod_{t=0}^{k-1} \mathcal{A}_t$ and the convention $\prod_{l=0}^{-1} \mathcal{A}_l = I$ is assumed. Thus

$$\tilde{p}_{k+1} = \tilde{q}_{k+1} / \|\tilde{q}_{k+1}\|_1 \quad (5.136)$$

$$= \frac{\prod_{t=0}^k \mathcal{A}_t \tilde{q}_0}{\prod_{t=0}^k \left\| \prod_{l=0}^{t-1} \mathcal{A}_l \tilde{q}_0 \right\|_1} \bigg/ \left\| \frac{\prod_{t=0}^k \mathcal{A}_t \tilde{q}_0}{\prod_{t=0}^k \left\| \prod_{l=0}^{t-1} \mathcal{A}_l \tilde{q}_0 \right\|_1} \right\|_1 = \frac{\prod_{t=0}^k \mathcal{A}_t \tilde{q}_0}{\left\| \prod_{t=0}^k \mathcal{A}_t \tilde{q}_0 \right\|_1}. \quad (5.137)$$

Since $\tilde{q}_0 = q_0$, we have $q_{k+1} = \prod_{t=0}^k \mathcal{A}_t \tilde{q}_0$ from (5.132), and the conclusion follows from (5.133). \square

From this result, the numerical issue raised above can be solved by normalizing q_k using its 1-norm at each time instant. Now we continue to discuss the application to finite-state real-valued hidden Markov models. For applications to state estimation, we start with a process x_k with real-valued states $\{x^1, \dots, x^N\}$, which are identified with the unit vectors $\{e_1, \dots, e_N\}$ in our analysis. We then try to find an estimate of x_k at each time instant that is optimal in some sense. Based on the expressions for the conditional distributions, the optimal estimates are straightforward to obtain. For performance comparison, we consider the MMSE estimate, which, as is well known, is the mean of the conditional distribution of x_k . Write $\mathbf{x} = [x^1, \dots, x^N]$. For the reliable communication channel case, the MMSE estimate \hat{x}_k conditioned on the available information of the remote estimator is

$$\hat{x}_k = \mathbf{x} \hat{p}_k, \quad (5.138)$$

while for the unreliable communication channel case, the MMSE estimate \check{x}_k conditioned on the available information of the remote estimator is

$$\check{x}_k = \mathbf{x}\check{p}_k. \quad (5.139)$$

Note that the above expressions apply equally to the case that x_k has vector real-valued states. The actual values of the measurement processes are not necessarily needed, as they do not affect \hat{x}_k and \check{x}_k in this formulation.

5.5.2 A numerical example

In this section, the proposed event-based estimates are illustrated by a numerical example. Consider a scalar real-valued process x_k with state space $\{x^1, \dots, x^N\}$ satisfying $N = 41$, $x^1 = -5$, $x^N = 5$, and $x^{i+1} - x^i = \Delta x$. For $\tau \in \mathbb{Z}_{-N:N}$, define $\phi(\tau, \sigma) := \frac{1}{\sqrt{2\pi}\sigma} \exp[-\frac{(\tau/N)^2}{\sigma^2}]$. The transition matrix A is constructed as $A = [A_1, \dots, A_N]$, where

$$A_i = [\phi(\tau, 0.1) / \sum_{l \in \mathbb{Z}_{1-i:N-i}} \phi(l, 0.1)]_{\tau \in \mathbb{Z}_{1-i:N-i}}$$

so that $\|A_i\|_1 = 1$ holds. The measurement process y_k is also a scalar real-valued process with state space $\{y^1, \dots, y^M\}$ satisfying $M = 41$, $y^1 = -2$, $y^M = 2$ and $y^{i+1} - y^i = \Delta y$. Note that the choice $M = N$ here does not necessarily imply the sensor has a huge number of measurement channels, but only means that the number of states of the measurement process is equal to the number of states of the hidden process. The measurement matrix C is constructed as $C = [C_1, \dots, C_N]$, where

$$C_i = [\phi(\tau, 0.05) / \sum_{l \in \mathbb{Z}_{1-i:M-i}} \phi(l, 0.05)]_{\tau \in \mathbb{Z}_{1-i:M-i}}.$$

First we focus on the reliable communication channel case. For this case, we consider the ‘‘send-on-delta’’ triggering condition, which is given as [44]

$$\gamma_k = \begin{cases} 1, & \text{if } \|y_k - y_{\tau_k}\|_1 > \delta \\ 0, & \text{otherwise,} \end{cases} \quad (5.140)$$

where τ_k is the last time instance when the measurement value is transmitted. For performance comparison, the MMSE state estimate obtained by ignoring the information contained in event-triggering conditions (which we refer to as the ‘p-MMSE estimate’) is considered. To indicate the best estimate possible, the MMSE estimate obtained using all past measurements at each time instant, which we refer to as the ‘n-MMSE estimate’, is also presented for comparison. The performance comparison of

the proposed event-based MMSE estimate (which we refer to as ‘e-MMSE estimate’) obtained for $\delta = 1.5$, the n-MMSE estimate and the p-MMSE estimate obtained using the same communication sequence γ_k is shown in Fig. 5.4. The resulting average estimation errors for p-MMSE, n-MMSE and e-MMSE estimates are 1.2169, 0.833 and 0.9138, respectively. To illustrate the tradeoff between the average communication rate and estimation performance, the average estimation error comparison for different average communication rates (obtained by using different δ 's) for the e-MMSE estimates and the p-MMSE estimates is further shown in Fig. 5.5. From all these comparisons, we observe that the exploration of the set-valued information contained in the event-triggering conditions leads to improved estimation performance in terms of average estimation error. In addition, from Fig. 5.4, it is shown that compared with the n-MMSE estimate, the performance can be maintained by the proposed e-MMSE estimate even when communication rate is much decreased by the triggering conditions. Also, note that when the triggering set is chosen to be sufficiently large, the performance of the e-MMSE coincides with that of the p-MMSE, as the information provided by the event-triggering set is the same as the information of a packet dropout when δ is chosen too large, (e.g., $\delta = 4.1$, which would imply a zero communication rate as $y_k \in [-2, 2]$).

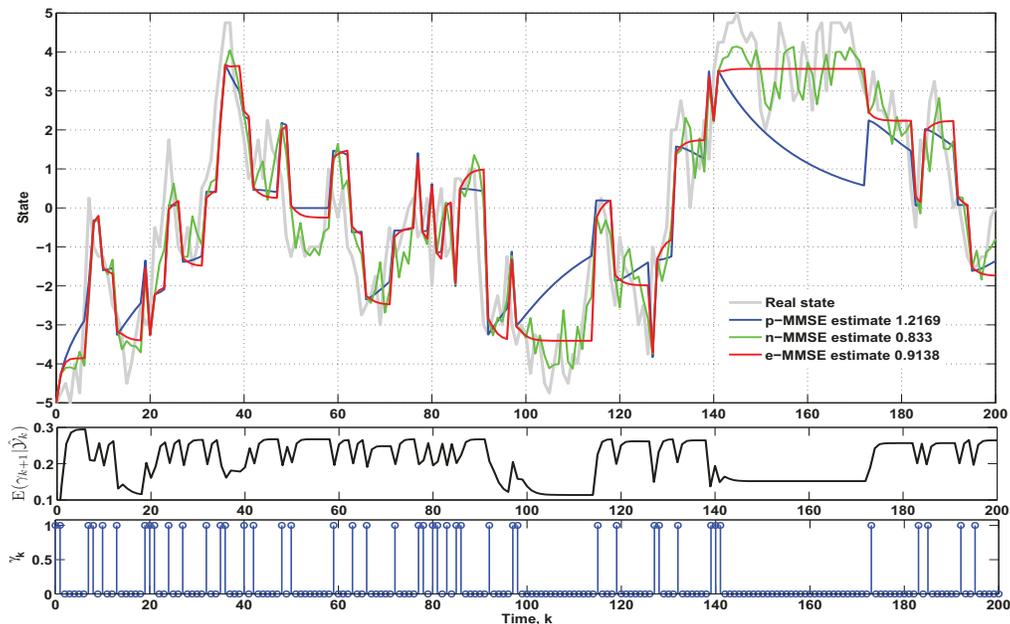


Figure 5.4: Estimation performance comparison for the reliable communication channel case.

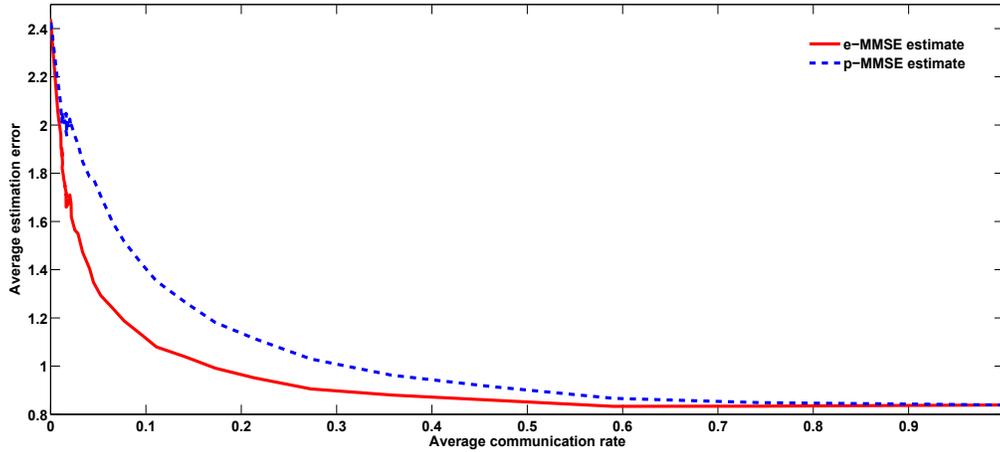


Figure 5.5: Tradeoff between average communication rate and estimation performance for the reliable communication channel case.

We now continue to illustrate the unreliable communication channel case. Due to the packet drop process ζ_k , the sensor cannot tell whether the estimator receives the measurement update or not during the event instants, and thus the ‘send-on-delta’ conditions can no longer be implemented¹. In this regard, we consider an alternative event-triggering condition:

$$\gamma_k = \begin{cases} 1, & \text{if } \|y_k\|_1 > \delta \\ 0, & \text{otherwise} \end{cases} \quad (5.141)$$

Basically, this condition implies that the measurement update is not sent to the remote estimator if y_k has a relatively small magnitude. For this scenario, the performance comparison of the proposed e-MMSE estimate for $\delta = 2.5$ with packet dropout rate $1 - \lambda = 0.2$, the n-MMSE estimate and the p-MMSE estimate obtained using the actual communication sequence $\gamma_k \zeta_k$ is shown in Fig. 5.6. The resultant average estimation errors for p-MMSE, n-MMSE and e-MMSE estimates are 1.1684, 0.8168 and 0.9585, respectively. The tradeoff between the average estimation error comparison and the average communication rates for the e-MMSE estimates and the p-MMSE estimates are further shown in Fig. 5.7 for the cases $\lambda = 0.4, 0.6$ and 0.8 , respectively. Again, from these results we observe that the exploration of the set-valued information contained in the event-triggering conditions leads to improved estimation performance in terms of average estimation error.

¹Note that feedback communication from the estimator is not helpful in this case as well, so long as the same lossy communication channel is used.

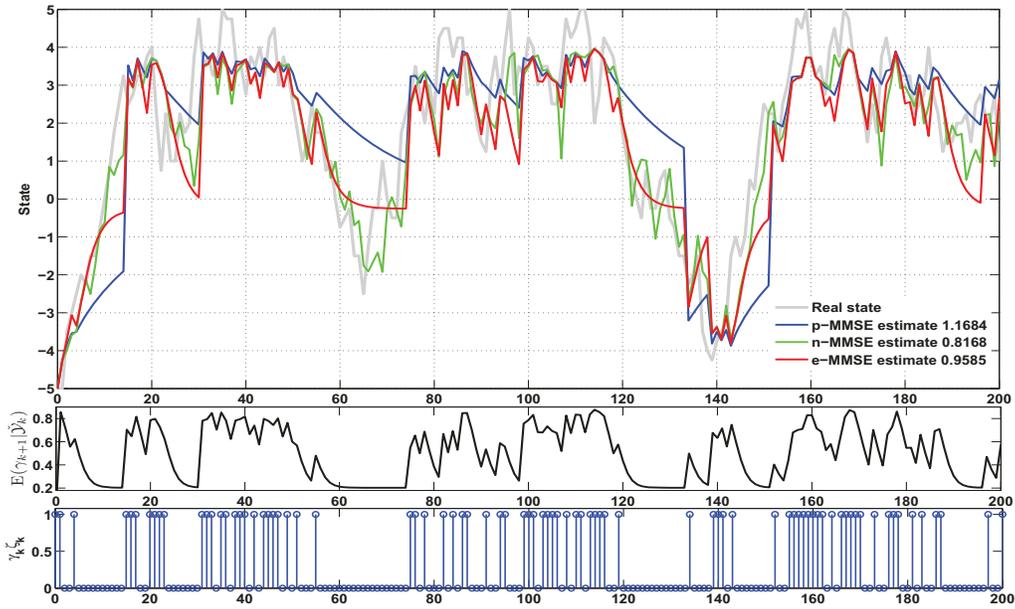


Figure 5.6: Estimation performance comparison for the unreliable communication channel case.

5.6 Summary

In this chapter, the change of probability measure approach is utilized to solve the event-based state estimation problem for discrete-state HMMs. For both the reliable and unreliable communication channel cases, closed-form expressions are obtained for the conditional probability distributions of the states on the hybrid measurement information available at the remote estimator, and the expected communication rates are found to be expressed as the ratio between the weighted 1-norm and 1-norm of the obtained conditional probability distribution of the states. So far, our attempts at handling the ambiguous information contained in the event-triggering sets with the change of probability measure approach seem encouraging. The results not only solve the event-based estimation problem for the finite state HMMs, but reveals a systematic approach to handling the event-based measurement information.

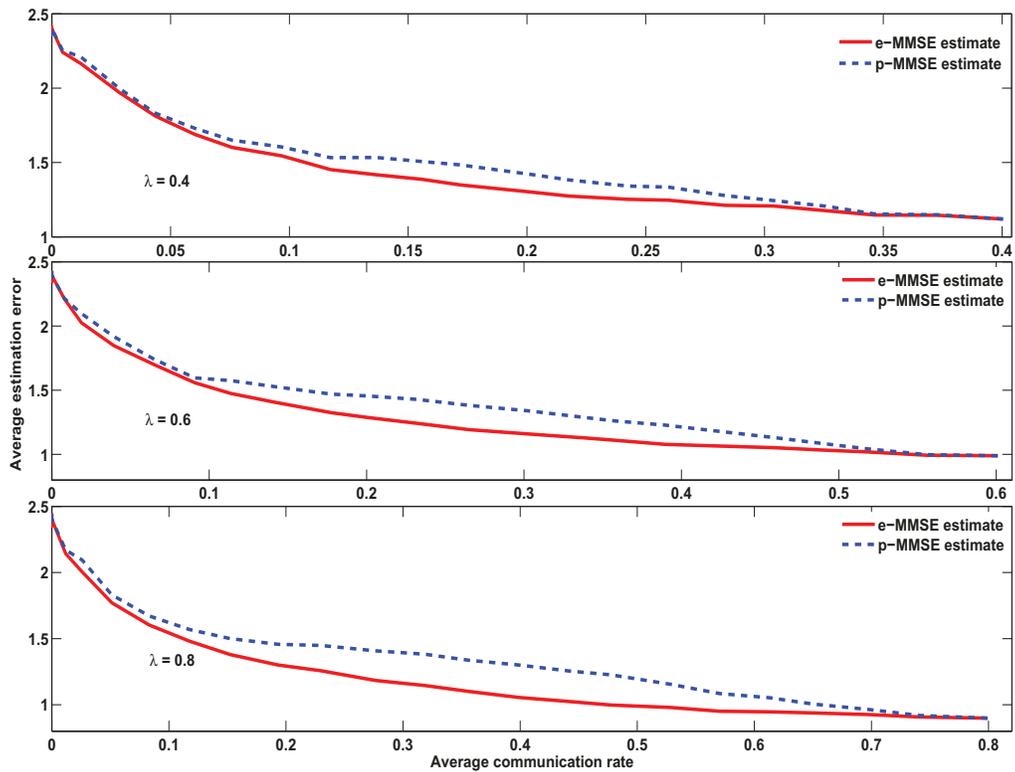


Figure 5.7: Tradeoff between average communication rate and estimation performance for the unreliable communication channel case.

Chapter 6

Conclusions and Future Work

6.1 Concluding remarks

In this thesis, the event-based state estimation problem is considered from different perspectives. The target is to find event-based estimators that is both easy to implement (in the sense of computational complexity) and optimal in a certain sense (e.g., ML, MMSE). The outcomes of our research attempts are further summarized as follows:

1. The structure of the event-based ML estimates is provided; for the one-step case, we show that it has a similar form as the Kalman filter with intermittent observations. A general approach for estimating upper and lower bounds of the communication rates has been proposed.
2. Under a Gaussian assumption on the conditional distribution of the state on the past hybrid measurement information, the event-based MMSE estimate is proposed for general event-triggering conditions and multiple sensor measurements; for the case of single-channel sensors, performance improvement in terms of smaller estimation error covariance is rigorously proved.
3. The properties of the exact and approximate set-valued filters with multiple sensor measurements are investigated. The results are applied to event-based estimation, and an event-triggering condition design procedure is proposed by simultaneously considering the requirements on estimation performance and communication rates.
4. With the change of probability measure approach, the event-based state estimation problems of discrete state HMMs for both the reliable and unreliable

communication channel cases are formulated and solved. Closed-form expressions for both the estimates and average communication rates are derived.

In addition, viewed from the extensive examples considered and the obtained estimation performance, the results developed are potentially attractive to a variety of industrial/commercial applications.

6.2 Future work

So far, the results of our exploration on simple-structured optimal event-based estimators are encouraging. These results, however, only form the tip of the iceberg in event-based estimation and much remains to be done. The following problems need to be further considered and addressed in the future research work, which are potentially the building blocks of a systematic approach to event-based estimation.

6.2.1 Application of the discrete state HMM results to general dynamical systems

The value of the discrete state HMM results would be enormously increased if the applicability to general dynamical systems can be further addressed. This seems possible, as the states of dynamical systems can be naturally represented by the states of HMMs. Another fact that increases the feasibility is that in sampled-data control systems, the measurements and control signals are received/implemented in a quantized fashion using analog to digital and digital to analog converters.

6.2.2 Event-based estimation problems for HMMs with continuous state spaces

To solve the event-based estimation problem from a more fundamental perspective, HMMs with continuous state spaces need to be considered. The difficulty in solving this problem mainly lies in the construction of a new probability measures and the structure of the event-triggering conditions under the new measure.

6.2.3 The effect of time delays in communication protocols

Although the effect of packet dropout is considered in Chapter 5, the effect of time delays have not yet been taken into account. It seems that the consideration of time

delays may further add to the computation complexity of the algorithms; however, alternative approaches might exist to overcome this difficulty.

6.2.4 Periodic local estimation vs. event-based remote estimation

Event-based remote estimation helps maintain the estimation performance when the communication and computation resources are limited; however, there also exist many applications with enough computation power so that the estimates can be generated periodically by the local estimators. In this regard, it is interesting to investigate the relative advantages of the two, and determine when to adopt event-based remote estimation or periodic local estimation.

Bibliography

- [1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. Wireless sensor networks: a survey. *Computer Networks*, 38:393–422, 2002.
- [2] T. Al-ani. Hidden markov models in dynamic system modelling and diagnosis. In P. Dymarski, editor, *Hidden Markov Models, Theory and Applications*, pages 27–50. InTech, 2011.
- [3] T. Alamo, J.M. Bravo, and E.F. Camacho. Guaranteed state estimation by zonotopes. *Automatica*, 41(6):1035–1043, 2005.
- [4] A. Alessandri, M. Baglietto, and G. Battistelli. Receding-horizon estimation for discrete-time linear systems. *IEEE Transactions on Automatic Control*, 48(3):473 – 478, Mar 2003.
- [5] F. A. Aliev and L. Ozbek. Evaluation of convergence rate in the central limit theorem for the kalman filter. *IEEE Transactions on Automatic Control*, 44(10):1905–1909, 1999.
- [6] ANALOG DEVICES. AD7988 datasheet. http://www.analog.com/static/imported-files/data_sheets/AD7988-1_7988-5.pdf, 2012.
- [7] B. D. O. Anderson and J. B. Moore. *Optimal Filtering*. Prentice-Hall, Englewood Cliffs, NJ, 1979.
- [8] I. Arasaratnam and S. Haykin. Cubature kalman filters. *IEEE Transactions on Automatic Control*, 54(6):1254–1269, 2009.
- [9] K.J. Åström and B.M. Bernhardsson. Comparison of Riemann and Lebesgue sampling for first order stochastic systems. In *Proceedings of the 41st IEEE Conference on Decision and Control*, volume 2, pages 2011 – 2016 vol.2, Dec. 2002.

- [10] M.S. Bazaraa, H.D. Sherali, and C.M. Shetty. *Nonlinear Programming: Theory and Algorithms*. John Wiley & Sons, 3 edition, 2006.
- [11] S. Bittanti, P. Colaneri, and G. De Nicolao. The difference periodic riccati equation for the periodic prediction problem. *IEEE Transactions on Automatic Control*, 33(8):706 – 712, 1988.
- [12] J. Chen. A partial order on uncertainty and information. *Journal of Theoretical Probability*, 2011.
- [13] P. Chevrel and S. Siala. Robust DC-motor speed control without any mechanical sensor. In *Proceedings of the 1997 IEEE International Conference on Control Applications*, pages 244–246, 1997.
- [14] L. Chisci, A. Garulli, and G. Zappa. Recursive state bounding by parallelotopes. *Automatica*, 32(7):1049–1055, 1996.
- [15] G. De Nicolao. On the time-varying riccati difference equation of optimal filtering. *SIAM Journal on Control and Optimization*, 30(6):1251–1269, 1992.
- [16] R. J. Elliott. *Stochastic Calculus and Applications*. Springer-Verlag, 1982.
- [17] R. J. Elliott, L. Aggoun, and J. B. Moore. *Hidden Markov Models: Estimation and Control*. Springer, 1995.
- [18] Executive Office of the President and President’s Council of Advisors on Science and Technology. Designing a digital future: Federally funded research and development in networking and information technology. *Report to the President and Congress*, Dec. 2010.
- [19] US National Science Foundation.
<http://www.nsf.gov/pubs/2014/nsf14542/nsf14542.htm>.
- [20] G.F. Franklin, J.D. Powell, and A. Emami-Naeini. *Feedback control of dynamic systems*. Pearson Prentice Hall, 5 edition, 2006.
- [21] G. Freiling and V. Inoescu. Time-varying discrete Riccati equation: some monotonicity results. *Linear Algebra and its Applications*, 286:135–148, 1999.
- [22] M. Fu and L. Xie. The sector bound approach to quantized feedback control. *IEEE Transactions on Automatic Control*, 50(11):1698–1711, Nov 2005.

- [23] M. Gales and S. Young. The application of hidden markov models in speech recognition. *Foundations and Trends in Signal Processing*, 1(3):195–304, 2008.
- [24] O. Geramifard, J. Xu, J. Zhou, and X. Li. Multimodal hidden markov model-based approach for tool wear monitoring. *IEEE Transactions on Industrial Electronics*, 61(6):2900–2911, June 2014.
- [25] G.C. Goodwin, M.M. Seron, and J. De Dona. *Constrained Control and Estimation: An Optimisation Approach*. Springer, 2005.
- [26] D. Han, Y. Mo, J. Wu, B. Sinopoli, and L. Shi. Stochastic event-triggered sensor scheduling for remote state estimation. In *Proceedings of 52nd IEEE conference on Decision and Control*, pages 6079–6084, 2013.
- [27] B. Hannaford and P. Lee. Hidden markov model analysis of force/torque information in telemanipulation. *The International Journal of Robotics Research*, 10(5):528–539, 1991.
- [28] L. Hetel and E. Fridman. Robust sampled-data control of switched affine systems. *IEEE Transactions on Automatic Control*, 58(11):2922–2928, Nov 2013.
- [29] J. Holtz. Sensorless control of induction motor drives. *Proceedings of the IEEE*, 90(8):1359–1394, 2002.
- [30] J. Holtz. Sensorless control of induction machines – with or without signal injection? *IEEE Transactions on Industrial Electronics*, 53(1):7–30, 2005.
- [31] IEEE 802.15.4. Wireless medium access control and physical layer specifications for low rate wireless personal area networks. 2006.
- [32] O.C. Imer and T. Başar. Optimal estimation with limited measurements. In *Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference*, pages 1029 – 1034, Dec. 2005.
- [33] K. Ito and K. Xiong. Gaussian filters for nonlinear filtering problems. *IEEE Transactions on Automatic Control*, 45(5):910–927, 2000.
- [34] N.L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions*, volume 1. Wiley, 2 edition, 1994.

- [35] S. Julier, J. Uhlmann, and H.F. Durrant-Whyte. A new method for the non-linear transformation of means and covariances in filters and estimators. *IEEE Transactions on Automatic Control*, 45(3):477–482, 2000.
- [36] R.L. Kosut, M.K. Lau, and S.P. Boyd. Set-membership identification of systems with parametric and nonparametric uncertainty. *IEEE Transactions on Automatic Control*, 37(7):929–941, Jul 1992.
- [37] A. Kurzhanski and I. Vályi. *Ellipsoidal Calculus for Estimation and Control*. Birkhauser, 1996.
- [38] L. Li and M. Lemmon. Performance and average sampling period of sub-optimal triggering event in event triggered state estimation. In *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference*, pages 1656–1661, Dec. 2011.
- [39] L. Li, M. Lemmon, and X. Wang. Event-triggered state estimation in vector linear processes. In *Proceedings of the 2010 American Control Conference*, pages 2138–2143, July 2010.
- [40] D. Liberzon and D. Nesic. Input-to-state stabilization of linear systems with quantized state measurements. *IEEE Transactions on Automatic Control*, 52(5):767–781, May 2007.
- [41] E. Lutwak. Intersection bodies and dual mixed volumes. *Advances in Mathematics*, 71:232–261, 1988.
- [42] B.G. Manjunath and S. Wilhelm. Moments calculation for the doubly truncated multivariate normal density. arXiv:1206.5387, Jun 2012.
- [43] X. Meng and T. Chen. Optimal sampling and performance comparison of periodic and event based impulse control. *IEEE Transactions on Automatic Control*, 57(12):3252–3259, 2012.
- [44] M. Miskowicz. Send-on-delta concept: An event-based data reporting strategy. *Sensors*, 6(1):49–63, 2006.
- [45] Y. Mo and B. Sinopoli. Kalman filtering with intermittent observations: Tail distribution and critical value. *IEEE Transactions on Automatic Control*, 57(3):677–689, 2012.

- [46] A. Molin and S. Hirche. An iterative algorithm for optimal event-triggered estimation. In *4th IFAC Conference on Analysis and Design of Hybrid Systems*, pages 64–69, Eindhoven, Netherlands, 2012.
- [47] D.R. Morrell and W.C. Stirling. Set-valued filtering and smoothing. *IEEE Transactions on Systems, Man and Cybernetics*, 21(1):184–193, 1991.
- [48] K.R. Muske, J.B. Rawlings, and J.H. Lee. Receding horizon recursive state estimation. In *Proceedings of the 1993 American Control Conference*, pages 900 – 904, June 1993.
- [49] Ghulam Mustafa and Tongwen Chen. h_∞ filtering for nonuniformly sampled systems: A markovian jump systems approach. *Systems & Control Letters*, 60(10):871 – 876, 2011.
- [50] A.G.O. Mutambara. *Decentralized Estimation and Control for Multisensor Systems*. CRC Press, Inc., Boca Raton, Florida, USA, 1998.
- [51] B. Noack, V. Klumpp, and U.D. Hanebeck. State estimation with sets of densities considering stochastic and systematic errors. In *Proceedings of the 12th International Conference on Information Fusion (Fusion 2009)*, pages 1751–1758, Seattle, Washington, USA, July 2009.
- [52] B. Noack, F. Pfaff, and U.D. Hanebeck. Combined stochastic and set-membership information filtering in multisensor systems. In *Proceedings of the 15th International Conference on Information Fusion (Fusion 2012)*, pages 1218–1224, Singapore, July 2012.
- [53] P.F. Odgaard, J. Stoustrup, and M. Kinnaert. Fault-tolerant control of wind turbines: A benchmark model. *IEEE Transactions on Control Systems Technology*, 21(4):1168–1182, 2013.
- [54] M. Rabi, G. Moustakides, and J. Baras. Adaptive sampling for linear state estimation. *SIAM Journal on Control and Optimization*, 50(2):672–702, 2012.
- [55] M. Rabi, G.V. Moustakides, and J.S. Baras. Multiple sampling for estimation on a finite horizon. In *Proceedings of the 45th IEEE Conference on Decision and Control*, pages 1351–1357, Dec. 2006.

- [56] L. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, Feb 1989.
- [57] C. V. Rao, J. B. Rawlings, and J. H. Lee. Constrained linear state estimation—a moving horizon approach. *Automatica*, 37(10):1619–1628, 2001.
- [58] R. Schneider. *Convex Bodies: the Brunn-Minkowski Theory*. Cambridge University Press, 1996.
- [59] D. Shi and T. Chen. Optimal periodic scheduling of sensor networks: A branch and bound approach. *Systems & Control Letters*, 62(9):732–738, 2013.
- [60] L. Shi, M. Epstein, and R.M. Murray. Kalman filtering over a packet-dropping network: A probabilistic perspective. *IEEE Transactions on Automatic Control*, 55(3):594–604, March 2010.
- [61] L. Shi, K.H. Johansson, and L. Qiu. Time and Event-Based Sensor Scheduling for Networks with Limited Communication Resources. In *Proceedings of the 18th IFAC World Congress*, pages 13263–13268, 2011.
- [62] J. Sijs and M. Lazar. Event based state estimation with time synchronous updates. *IEEE Transactions on Automatic Control*, 57(10):2650–2655, Oct. 2012.
- [63] J. Sijs, B. Noack, and U.D. Hanebeck. Event-based state estimation with negative information. In *Proceedings of the 16th International Conference on Information Fusion*, pages 2192–2199, 2013.
- [64] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M.I. Jordan, and S.S. Sastry. Kalman filtering with intermittent observations. *IEEE Transactions on Automatic Control*, 49(9):1453 – 1464, Sept. 2004.
- [65] J. C. Spall and K. D. Wall. Asymptotic distribution theory for the kalman filter state estimator. *Communications in Statistics - Theory and Methods*, 13(16):1981–2003, 1984.
- [66] W.C. Stirling and D.R. Morrell. Convex Bayes decision theory. *IEEE Transactions on Systems, Man and Cybernetics*, 21(1):173–183, 1991.
- [67] G. Su and J.W. McKeever. Low-cost sensorless control of brushless DC motors with improved speed range. *IEEE Transactions on Power Electronics*, 19(2):296–302, 2004.

- [68] G.M. Tallis. The moment generating function of the truncated multinormal distribution. *Journal of the Royal Statistical Society, Series B (Methodological)*, 23(1):223–229, 1961.
- [69] G.M. Tallis. Elliptical and radial truncation in normal populations. *The Annals of Mathematical Statistics*, 34(3):940–944, 1963.
- [70] R. Tempo, G. Calafiore, and F. Dabbene. *Randomized Algorithms for Analysis and Control of Uncertain Systems*. Springer-Verlag, London, 2005.
- [71] Texas Instruments. A True System-on-Chip Solution for 2.4-GHz IEEE 802.15.4 and ZigBee Applications. 2011.
- [72] J. Weimer, J. Araujo, and K. H. Johansson. Distributed event-triggered estimation in networked systems. In *Proceedings of the 4th IFAC conference on Analysis and Design of Hybrid Systems*, pages 178–185, Eindhoven, Netherlands, June 2012.
- [73] J. Wu, Q. Jia, K. Johansson, and L. Shi. Event-based sensor data scheduling: Trade-off between communication rate and estimation quality. *IEEE Transactions on Automatic Control*, 58(4):1041–1045, 2013.
- [74] S.S.H. Zaidi, S. Aviyente, M. Salman, Kwang-Kuen Shin, and E.G. Strangas. Prognosis of gear failures in DC starter motors using hidden markov models. *IEEE Transactions on Industrial Electronics*, 58(5):1695–1706, May 2011.
- [75] ZigBee RF4CE. ZigBee RF4CE specification. *ZigBee Alliance document 094945r00ZB*, 2010.
- [76] S. Zonouz, K.M. Rogers, R. Berthier, R.B. Bobba, W.H. Sanders, and T.J. Overbye. Scpse: Security-oriented cyber-physical state estimation for power grid critical infrastructures. *IEEE Transactions on Smart Grid*, 3(4):1790–1799, Dec 2012.