Optimal State Estimation of Cyber-Physical Systems

by

Jiarao Huang

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

Abstract

This thesis focuses on state estimation problems of cyber-physical systems to overcome the challenges brought by their features, e.g, resource limitations in sensor networks and model uncertainties due to the changes in interconnections of components.

To achieve smart allocation of the scarce resources of cyber-physical systems, two event-based state estimation problems are formulated and solved for systems described by hidden Markov models utilizing a new reference measure approach with the change of probability measure. For a linear Gaussian system with an energy harvesting sensor, the joint conditional probability distribution of the state and energy is obtained based on the event-triggered information received at the remote estimator under the energy-dependent measurement transmission policy.

The robust state estimation problems are investigated for linear Gaussian systems with event-triggered scheduling and systems with unknown exogenous inputs utilizing the risk-sensitive approach, where closed-form risk-sensitive state estimates are derived. A fully distributed robust consensus-based filtering algorithm for systems measured by a sensor network is proposed with stability analysis on the local estimators.

Based on the proposed results, state estimates that are optimal in a certain sense can be calculated in a simple recursive way, which are potentially applicable to industrial processes. The effectiveness of the proposed methods is validated by simulation examples, showing performance improvements in certain scenarios in the sense of mean estimation errors.

Preface

The introduction and literature review in Chapter 1 are my original work. Chapters 2-5 of this thesis consist parts of the international collaborations with Dr. Dawei Shi at Beijing Institute of Technology. The ideas, problem formulations, mathematical derivations and simulation examples presented in this thesis are my original work.

Parts of Chapter 2 of this thesis have been published as

- J. Huang, D. Shi, and T. Chen, "Energy-based event-triggered state estimation for hidden Markov models," *Automatica*, 79:256-264, 2017.

- J. Huang, D. Shi, and T. Chen, "Dynamically event-triggered state estimation of hidden Markov models through a lossy communication channel," in *Proceedings of the IEEE 55th Conference on Decision and Control*, pp. 5122-5127, Las Vegas, USA, December 12-14, 2016.

The results in Chapter 3 have been published as

- J. Huang, D. Shi, and T. Chen, "Event-triggered state estimation with an energy harvesting sensor," *IEEE Transactions on Automatic Control*, 62(9):4768-4775, 2017.

Part of the results in Chapter 4 has been accepted for publication as

- J. Huang, D. Shi, and T. Chen, "Robust event-triggered state estimation: A risk-sensitive approach," accepted by *Automatica* as a regular paper, 2018.

Another portion of the results in Chapter 4 has been published as

- J. Huang, D. Shi, and T. Chen, "Robust state estimator design for systems with unknown exogenous inputs: A Risk-Sensitive Approach," in *Proceedings of the 14th IEEE International Conference on Control and Automation*, pp. 136-141, Anchorage, Alaska, USA, June 12-15, 2018.

Part of Chapter 5 has been accepted for publication as

- J. Huang, D. Shi, and T. Chen, "Distributed robust state estimation for sensor networks: A risk-sensitive approach," accepted by *the 57th IEEE Conference on Decision and Control*, Miami Beach, FL, USA, December 17-19, 2018.

In all the papers, the problems were formulated largely on my own, and the theoretical results and simulation examples were developed by myself. Dr. Chen was the supervisory author and provided insightful suggestions on the results and the manuscript organization. Dr. Shi provided valuable comments on the technical derivations and contributed to manuscript edits.

Acknowledgements

I would like to express my deepest gratitude to those who have helped and supported me during my PhD program over the past four years. First of all, I would like to thank my supervisor, Dr. Tongwen Chen, for all his efforts, guidance, academic support and help through my entire PhD research. Dr. Chen gave me the chance to step into the control community and be a theoretical researcher. With his full expertise, it was Dr. Chen who patiently taught me how to formulate a theoretical problem, persevere with it, conduct research, write academic articles, revise the manuscript, write a response to reviewers' comments and give an oral presentation. The most important thing I learned from Dr. Chen was how to think critically and always be sensitive to the cutting edge research topics. Without his guidance and persistent help, this dissertation would not have been possible.

Secondly, I would like to thank my academic collaborator, Dr. Dawei Shi, from the Harvard John A. Paulson School of Engineering and Applied Sciences, Harvard University for his altruistic comments and suggestions on the results. I am especially grateful to Dr. Shi for his enthusiastic encouragement and advice on my career path.

I would like to express my special gratitude for my previous supervisor, Prof. Huijun Gao, at the Research Institute of Intelligent Control and Systems, Harbin Institute of Technology, for his persistent support and help during my PhD program. It was Prof. Gao who encouraged me to pursue my PhD abroad, which led me to the University of Alberta.

Further, I would like to thank all the members of Dr. Chen's group and staff from the Department of Electrical and Computer Engineering of the University of Alberta for their continuous help and support in the past four years. I would also like to thank

Prof. Qing Zhao and Prof. Mahdi Tavakoli for their comments and suggestions in my thesis and final defense. Thanks to their valuable suggestions, the final dissertation was completed in a timely manner. A special thank you is given to Linan Xu, for sharing his insights with me in our discussions and also his encouragement during my PhD program.

Lastly, I would like to acknowledge the support and love of my parents. Their unselfish and endless love gave me strength and inspired me to do better. They are source of my happiness, and I feel so lucky to have them in my life. Moreover, I would like to thank my friends who have accompanied, supported and encouraged me during the past four years.

Contents

List of Figures

List of Acronyms

CPS	Cyber-Physical Systems
WSN	Wireless Sensor Networks
MMSE	Minimum Mean Squared Error
LMI	Linear Matrix Inequalities
ETC	Event-Triggering Conditions
i.i.d.	independent and identically distributed
MAP	Maximum A Posteriori
MAE	Mean Absolute Error
ADC	Analog-to-Digital Converter
MSE	Mean Square Error
KF with IO	Kalman Filter with Intermittent Observations
RSET	Risk-Sensitive Event-Triggered
MMSE ET	Minimum Mean Squared Error Event-Triggered
UMV	Unified Minimum Variance
LTI	Linear Time-Invariant
RS	Risk-Sensitive
RMSE	Root Mean Square Errors

Chapter 1

Introduction

In this chapter, we introduce the research scope of this thesis and provide the background information for the state estimation problem of cyber-physical systems (CPS). We present a survey of the related investigations in the literature, which is then followed by the summary of the contributions of this thesis.

1.1   Background

This section begins with a description of CPS and their applications. The main scope of thesis belongs to the optimal state estimation of CPS. Several optimal filtering methods are reviewed and event-triggered scheduling is introduced to deal with the resource constraints in CPS.

1.1.1   Overview

Cyber-physical systems (CPS) represent a new generation of systems with integrated computational and physical capabilities that can interact with humans, e.g., through embedded computers, networked monitors and physical processes. The sensors, actuators, processors, and almost everything on the planet, can be inter-connected and become parts of CPS, which makes the Internet of Things, using technologies like wireless sensor networks (WSN) [78], Zigbee [6], radio frequency identification [54], Wi-Fi [37] and 4G [10]/5G [4]. Examples of CPS cover a wide range of applications, e.g., environmental monitoring [9], autonomous automobile systems [35], smart grids [44], medical monitoring [36] and robotics systems [41]. In the past few decades, the Internet transformed how humans interact and communicate with one another, revolutionized how and where information is accessed, and even changed how people buy and sell prod-

ucts. Similarly, CPS will transform how humans interact with and control the physical world around us, by which enormous societal and economic benefits will be created [52]. In many applications of CPS, the sensors and estimators are separated and connected by wired/wireless communication channels, which results in the increasing demand for investigating remote state estimation problems. The main scope of this thesis belongs to the optimal state estimation of CPS.

Due to the configuration of CPS, the systems usually operate under constrained resources. For example, in WSN, the communication channel is commonly shared by many sensors, which results in the limited communication bandwidth for a single sensor. On the other hand, many sensor networks in applications are powered by batteries and involve a large number of sensor nodes, which makes the maintenance and replacement of sensors' batteries time-consuming and expensive, especially when the networks are located in hazardous environments. Thus, the problem of extending the lifetime of wireless sensors with limited available battery energy while maintaining their functionality becomes an imperative issue. As a possible solution, event-triggered scheduling has received considerable attention in the control community by transmitting data packets only when the so-called event-triggering conditions are violated. Event-triggered state estimation forms the first research topic of this thesis.

Another challenge in CPS is the performance degradation caused by model uncertainties, which can result from system identification, nonlinear effects and changes in the interconnections of components or external environment. This has motivated the pursuit of robust filtering for CPS in the past decade. Inspired by the existing results on the robustness of risk-sensitive filters in the literature, we solve a robust estimation problem for CPS by utilizing the risk-sensitive filtering approach, which forms the second line of inquiry of this thesis.

1.1.2 Optimal state estimation: H_2, H_∞ and risk-sensitive filtering

Normally, state estimators are designed in an optimal sense with respect to certain performance measures. For optimal state estimation, the most widely known estimator is the H_2 estimator (the Kalman filter [32]), which minimizes the quadratic criterion of the estimation errors and operates by propagating the mean and covariance of the state through time. The Kalman filter is established as a fundamental tool for analyzing a broad class of estimation problems. Various books and papers have derived and presented the filter equations, so we skip the detailed part. For linear systems, if the

noises are uncorrelated zero-mean Gaussian, the Kalman filter is the optimal estimator that minimizes the weighted two-norm of the expected value of the estimation error; if the noises are uncorrelated and zero-mean, but not Gaussian, then the Kalman filter is the best linear estimator that minimizes the two-norm criterion [66]. In other words, nonlinear filters may give a smaller estimation error, but the Kalman filter is the best linear filter for non-Gaussian noises. Even for colored or correlated noises, the optimal estimator can be obtained by modifying the Kalman filter [66]. For nonlinear systems, there are various formulations of nonlinear Kalman filters to give the approximate solutions to the estimation problem, for example, the extended Kalman filter [29] and the unscented Kalman filter [30].

The Kalman filter is an effective tool for scenarios that the process has a known dynamics and that the noises have known statistical properties. Unfortunately, the exact model and/or the noise descriptions are unknown or inaccurate in many industrial applications. At this time, a robust filter which can tolerate such model uncertainty is in need. The H_∞ (also called minimax estimation) estimator is specifically designed for robustness and makes no assumptions on the statistics of the process and measurement noise. The H_∞ estimator minimizes the H_∞ norm of the transfer function from the disturbances to the estimation errors [66]. We may interpret the H_∞ norm as the maximum energy gain from the noises to the estimation errors. Hence, the H_∞ estimator minimizes the worst-case estimation error. However, closed-form solutions to the optimal H_∞ estimation are available only in some special cases and it is common in the literature to settle for a suboptimal solution. In other words, instead of direct minimization of the H_∞ norm, one usually selects a performance bound and seeks an estimation strategy that satisfies the threshold. In this way, the H_∞ estimator gives hard upper bounds on the estimation errors, no matter what the disturbances are (as long as they are of finite energy). For detailed tutorials on the H_∞ estimation, see [66].

Different from the aforementioned H_2 (also called risk-neutral) filtering and the H_∞ filtering, the risk-sensitive filtering [70] minimizes the expectation of the exponential of the squared estimation error multiplied by a risk-sensitive parameter, which penalizes all the higher order moments of the estimation error energy. Such filtering is more robust to system and noise uncertainties compared with the H_2 filtering [7] and closely related to the H_∞ filtering in the limit of small noise [14]. The risk-sensitive parameter, which can be tuned in the cost function, allows a trade-off between optimal filtering for the nominal model and robustness to model uncertainty. The risk-sensitive filter approaches the Kalman filter when the risk-sensitive parameter of the exponential cost

approaches zero. Furthermore, in the small noise limit, the risk-sensitive estimation problem approaches a worst-case estimation problem in a deterministic noise scenario given from a differential game. In this thesis, the risk-sensitive filtering approach is utilized to solve the robust state estimation problems. For further results on risk-sensitive filtering, see [14, 70] and references therein.

1.1.3 Event-triggered scheduling

Due to the increasing demand of maintaining system performance with limited communication and energy resources, event-triggered sampling and signal processing have received considerable attention in the control community. Event-triggered scheduling limits the sensor communication to instances when the system needs attention. Different from the classical time-triggered (periodical) scheduling, event-triggered scheduling is reactive and performs the transmissions from the sensor to estimator only when, for instance, the plant state deviates more than a certain threshold from a desired value. In this way, even when the measurements are not sent to the estimator at some of the time instants, the estimator is still able to infer some information about the unsent measurements according to the fact that the event-triggering conditions are satisfied at these time instants. Thus, an event-triggered scheduling scheme provides an attractive way of handling communication and energy constraints: more measurement information is available to the estimator, which potentially implies improved estimation performance.

Before continuing, we introduce how an event-triggered scheduling scheme is implemented at the sensor side. Basically, each sensor is equipped with an event-triggered data scheduler. At each time instant, the sensor samples the process and produces a measurement, and the scheduler of the sensor decides the value of a logic variable γ_k, by testing some event-triggering condition, where γ_k determines whether a data transmission is allowed or not. The event-trigger γ_k usually takes value in $\{0, 1\}$; only if $\gamma_k = 1$, the sensor sends the measurement to the estimator. According to the nature of the event-triggering conditions, they generally fall into two categories, deterministic event-triggering conditions and stochastic event-triggering conditions. For each category, there are four different types, namely, open-loop conditions, send-on-delta conditions, innovation-based conditions and variance-based conditions. Generally speaking, each deterministic event-triggering condition is a special case of another stochastic event-triggering condition with specific parameterization. In this thesis, we focus on the more general category, namely, stochastic event-triggering conditions.

The main scope of this thesis is the state estimation for CPS. The first problem

when we include event-triggered scheduling in CPS is the design of the event-triggering condition. Improved estimation performance can be achieved by properly designing the event-triggering strategy if restrictions on average communication rates exist. As soon as the scheduling strategy is chosen, the primary problem is estimator design. The main challenge for the event-triggered estimator design is how to explore the information that is implicitly provided by the event-triggering conditions. Due to the existence of the event-triggering conditions, the obtained estimators do not have a simple recursive form unless the event-triggered estimator design problem is carefully formulated. Theoretically, the joint optimal design of the event-triggering strategy and estimator could achieve even better estimation performance compared with the separate design. However, the interaction between the scheduling strategy and estimator makes this joint design problem very difficult to solve, which is out of the scope of this thesis.

Once the estimator is obtained, it is critical to evaluate the performance of the estimator. Stability of the event-triggered estimator is a crucial property to be verified. For some of the results, the stability property is very difficult to be analyzed. Another important aspect in performance assessment is to verify that the exploitation of the information contained in the event-triggering conditions during the non-event instants indeed helps maintain the performance at much reduced communication rates. In other words, it is necessary to verify that the performance of an event-triggered estimator outperforms the performance of an estimator obtained under the same communication rate but without exploiting the information provided by the event-triggering conditions. Also, analysis of the average communication rate from the sensor to the estimator is an important and nontrivial problem in performance evaluation.

1.2 Literature survey

In this section, we present a detailed literature survey on recent advances in event-triggered state estimation and risk-sensitive filtering.

1.2.1 Event-triggered state estimation

To achieve smart allocation of scarce resources, event-triggering strategies (event-triggering conditions) need to be carefully designed. In general, the optimal design of event-triggering strategies is theoretically difficult to analyze and solve, which is beyond the scope of this thesis. Instead of searching for optimal strategies, existing investigations normally searched for a suboptimal strategy and the problem formula-

tions generally fell into two categories: (1) solve a constrained optimization problem [50, 51], e.g., to optimize estimation error covariance under the constraints on average transmission rates; (2) solve an unconstrained optimization problem [45, 75], where the objective function is the weighted sum of different terms, e.g., estimation error covariance and communication cost.

In addition to the scheduling strategy design, another important issue is to find the optimal estimate for a specified event-triggering scheme, which is the main focus of my research work. The primary concern of event-triggered estimator design is how the information provided by the proposed event-triggering scheme can be properly exploited to maintain/improve the estimation performance. A number of investigations have been made on this topic and recent developments are introduced as follows. Sijs and Lazar [63] developed a state estimator based on a general event sampling strategy, where an asymptotically bound on the estimation error covariance matrix was obtained. Sijs et al. [64] obtained an event-based state estimator by minimizing the maximum mean squared error and treating the event-triggering conditions as non-stochastic uncertainties. Trimpe [72] considered an estimation problem for linear systems with multiple sensors, where each sensor ran a copy of the remote estimator and made the triggering decision based on this estimate; a further extension was made in [73] by removing the ideal assumption of identical estimates on all sensors. Wu et al. [77] proposed an event-based sensor data scheduler for linear Gaussian systems and a simple form of a minimum mean squared error (MMSE) estimator was derived by adopting a Gaussian approximation technique. Shi et al. [58] extended the work in [77] to linear Gaussian systems with multiple sensors and considered general event-triggering schemes. Event-triggered state estimators were designed by formulating constrained optimization problems in [59]. Lee et al. [38] considered a problem of event-based state estimation for continuous-time nonlinear systems utilizing the Markov chain approximation method. Stochastic event-triggering conditions parameterized by Gaussian kernels were proposed in Han et al. [23] for linear Gaussian systems, and the exact MMSE estimates were obtained in recursive and closed forms. In [84], an event-based state estimator for a complex network was designed such that the MMSE was ultimately bounded. To study the effect of a lossy channel on event-based estimation, the reference probability measure approach was utilized to exploit the event-triggered measurement information in [60]. The scenario of event-based state estimation for systems with unknown exogenous inputs was considered in [57].

The aforementioned investigations on event-triggered state estimation assume that

accurate system models are known and disturbances have known statistical properties. However, model mismatch commonly exists in engineering applications, which motivates the pursuit of robust event-triggered estimator design. Recently, a considerable amount of research effort has been targeted on robust event-triggered filter design to guarantee a prescribed H_∞ performance for the filtering error dynamics, e.g., [5, 16, 21, 26, 42, 82], where the design of the desired H_∞ filter gains is usually developed in the form of linear matrix inequalities (LMI). For more results on the event-triggered estimator design, see [62] and the references therein.

Aside from the developments in theoretical analysis, event-triggered scheduling and the corresponding estimators have been successfully applied in a number of engineering areas. Trimpe and D'Andrea [74] applied event-triggered estimators to the estimation and control of a balancing cube with six rotating modules. Hirche and Buss [25] used an event-triggered data reconstruction technique in haptic teleoperation to estimate the measurement data, where the transmission of the data packet was controlled by a deadband event-triggering condition. A distributed event-based estimation method was introduced in [75] and was tested on a water tank system. Chen et al. [8] applied some typical event-triggering conditions to a DC torque motor system and presented experimental and comparative estimation study on event-triggered state estimation.

1.2.2 Risk-sensitive filtering

Since the pioneering work by Speyer et al. [70], which first proposed the exponential form of the estimation criterion to be minimized, several results on risk-sensitive filtering have been reported in the literature. For example, Dey and Moore addressed the risk-sensitive filtering and smoothing problem for hidden Markov models with finite-discrete states in [14] and for discrete-time nonlinear Gauss-Markov state-space models in [13]. A risk-sensitive filtering problem was further investigated for continuous-time nonlinear systems in [12]. In [7], Boel et al. gave a precise meaning to the robustness of risk-sensitive filters for systems with model uncertainty by showing that risk-sensitive estimators enjoyed an error bound which was the sum of two terms, the first of which coincided with an upper bound on the error one would obtain if one knew exactly the underlying probability model, while the second term was a measure of the distance between the true and design probability models. A sequential filtering scheme for the risk-sensitive state estimation of partially observed Markov chains was presented in [53]. Zhang et al. investigated steady-state risk-sensitive filtering, prediction and smoothing problems for discrete-time singular systems in [80]. The authors in [49] derived a

risk-sensitive multiple-model filtering algorithm for jump Markov linear systems. A widely utilized method in the literature to obtain the risk-sensitive filter is the reference measure approach, where the filtering problem is first solved under a reference measure and then transformed back to the "real-world" measure. For those who are interested, see [17] for details on the reference measure approach.

For risk-sensitive filters, the implementation issue can sometimes be a challenge since closed-form expressions for such filters are available only for a very limited class of models including finite state-space Markov chains and linear Gaussian models. Sadhu et al. [56] presented a novel particle implementation of risk-sensitive filters for nonlinear, non-Gaussian state-space models utilizing a probabilistic reinterpretation of the risk-sensitive filter recursions.

For ergodic properties of the risk-sensitive filters, the asymptotic stability properties of discrete-time risk-sensitive filters with non-Gaussian initial conditions were studied in [11]. The authors in [39] proposed a contraction analysis of risk-sensitive Riccati equations, where two conditions were needed to guarantee that the risk-sensitive Riccati map was contractive, one for the risk-sensitive parameter and the other for the variance of the initial state of the system.

1.3   Summary of the contributions

The results presented in this thesis target at remote state estimation problems for CPS with resource limitations or model uncertainties, which results in different state estimators that are optimal in a certain sense. To simplify the problem, we usually consider one issue at a time, whereas Section 4.1 is the only part that consider both issues. The major contributions of this thesis that distinguish it from other work are summarized as follows.

In Chapter 2, we consider energy-based event-triggered state estimation problems. A problem of energy-based event-triggered remote state estimation for systems described by discrete finite-state hidden Markov models is investigated in Section 2.1. We consider energy harvesting sensors, which absorb power from the environment or other resources and convert it to electrical power. The event-triggering condition (ETC) considered depends on the sensor energy level, which evolves according to a Markov process. The reference measure approach is used to obtain the optimal estimates of the state based on event-triggered measurement information available at the remote estimator. The effectiveness of the proposed method is illustrated with simulation results for a linear

Gaussian system quantized and parameterized into a hidden Markov model. In Section 2.2, a problem of event-based state estimation for hidden Markov models is investigated. We consider the scenario that the transmission of the sensor measurement is decided by a dynamic event-trigger, the state of which depends on both the sensor measurement and the previous triggering state. An independent and identically distributed (i.i.d.) Bernoulli process is utilized to model the effect of packet dropout. Using the reference probability measure approach, expressions for the unnormalized and normalized conditional probability distributions of the states on the event-triggered measurement information are derived, based on which optimal event-based state estimates can be obtained. The effectiveness of the proposed results is illustrated through a numerical example together with comparative simulations.

In Chapter 3, an event-triggered state estimation problem for a linear Gaussian system with an energy harvesting sensor is investigated. A stochastic energy-dependent event-triggering transmission protocol is proposed to balance the communication rate and estimation performance according to the sensor's battery energy. The joint conditional probability distribution of the state and sensor's energy level on the combined set-valued and point-valued event-triggered measurements available to the remote estimator is derived. Based on this distribution, the recursive MMSE estimates of the state and sensor energy level are obtained. Also, the relationship between the average communication rate and energy harvesting rate is discussed. Finally, a numerical example is provided to evaluate the effectiveness of the proposed results.

In Chapter 4, a problem of risk-sensitive state estimation is considered. In Section 4.1, we investigate a robust event-triggered remote state estimation problem for linear Gaussian systems with a stochastic event-triggering condition. The reference measure approach is used to obtain a robust event-triggered estimate that minimizes the so-called risk-sensitive criterion, which refers to the expectation of the exponential of the sum of the squared estimation error. The closed-form expressions of the risk-sensitive event-triggered posterior and prior estimates are presented, which are shown to evolve in simple recursive Kalman-like structures. Moreover, two sufficient stability conditions for the proposed estimators are given. Comparative simulation results demonstrate the effectiveness of the proposed risk-sensitive event-triggered filter. In Section 4.2, a robust state estimation problem for discrete-time systems with unknown exogenous inputs is investigated utilizing a risk-sensitive filtering approach. By proposing a Radon-Nikodym derivative, we introduce a reference measure under which the measurement and system state become independent. Based on this independent property and by

treating the unknown inputs as a process modeled by a non-informative prior, we derive the reformulated risk-sensitive cost criterion under the reference measure and further propose a recursive algorithm for the risk-sensitive state estimate, which is further validated by a simulation example.

In Chapter 5, we investigate a distributed robust state estimation problem for linear Gaussian systems measured by a sensor network, where the sensors can communicate only with their neighbors and each sensor runs a local filter to estimate the state of the process based on the measurements from its neighbors. We present a distributed risk-sensitive filtering algorithm, where the high-gain dynamic consensus filter is utilized to compute the fused measurement data and the fused covariance-inverse matrices, based on which, the local filter is updated in a Riccati-based linear recursive form. For linear time-invariant systems, the asymptotic stability of local estimators in the proposed distributed filtering algorithm is guaranteed if the value of the risk-sensitive parameter is chosen such that the centralized risk-sensitive filter is asymptotically stable. The robustness of the proposed risk-sensitive filtering algorithm to system uncertainties is verified by simulation results.

Chapter 2

Event-triggered state estimation for hidden Markov models

In this chapter, we investigate event-triggered state estimation problems utilizing the reference measure approach, where the system considered is described by a discrete finite-state hidden Markov model, which can represent the individual component states of a dynamic system [2].

2.1 Energy-based event-triggered state estimation for hidden Markov models*

In this section, we consider systems equipped with an energy harvesting sensor, where the sensor's battery level dynamics can be modeled by a finite state Markov chain. An energy-based event-triggered scheduling protocol is proposed and a novel estimation method is developed to estimate the system state and sensor's energy level utilizing the reference measure approach.

2.1.1 Introduction

Existing investigations on event-based state estimation normally assume that the sensors are powered by non-rechargeable batteries. With the recent development of wireless electronic systems, energy harvesting technology provides a promising way to reduce maintenance costs and to achieve longer lifetime of sensors by absorbing external energy and converting it to electrical energy. Applications of energy harvesting sensors have appeared in a number of engineering systems, e.g., satellite communications [18]

---

*Parts of the results in this section appeared in *Automatica*, 79:256-264, 2017.

and animal tracking [65]. In this section, we consider energy harvesting sensors and investigate a problem of remote state estimation based on the event-triggered information provided by such sensors. The system considered is described by a discrete finite-state hidden Markov model, applications of which are found in various areas of engineering [19, 22, 24, 79, 83].

The main contributions of the work in this section are summarized in two folds. First, an energy-based event-triggered transmission protocol is proposed, which considers the amount of energy available for measurement transmission in event-trigger and estimator design. We consider an energy harvesting sensor with a rechargeable battery. The dynamics of the sensor's energy level is modeled by a discrete-time Markov chain, which is a typical way of modeling an energy harvesting process in various application scenarios, e.g., energy harvesting wireless sensing for environment monitoring [27], energy harvesting cognitive radio [71], and battery-powered wireless networks [20]. The proposed energy-dependent event-triggering condition allows better utilization of the sensor's energy (for instance, when the battery energy level is relatively low, the sensor will send the measurement only if the system has significant deviations or other predefined important changes), so that the system can compromise estimation performance with lower communication and energy costs.

Second, a state estimation method to approach the energy-dependent event-triggered measurement information is proposed using the reference probability measure approach. Different from [60], the exact triggering condition is not known to the estimator in this work as the energy level is not always transmitted, which adds to the difficulty for state estimation. The main challenge is how to construct appropriate reference measures so that simple recursive expressions of certain weighted state conditional distributions can be obtained under the reference measures by exploiting the information pattern induced by the proposed energy-based event-triggered transmission protocol. By defining a first reference measure and a map from the "real-world" measure to the reference measure, we derive a recursive form of the unnormalized conditional probability distribution of the state under the reference measure, which depends on the conditional distribution of the energy level. To estimate the energy level, we propose a second level of the reference measure, under which the recursive form of the unnormalized conditional distribution of the energy level under this measure is obtained. With the help of these two reference measures, the estimates of the system states under the "real-world" probability measure can be further derived.

Notation: Let $\mathbb{N}$ denote the set of nonnegative integers. Write $\mathbb{N}_{1:M} := \{1, 2, \ldots, M\}$
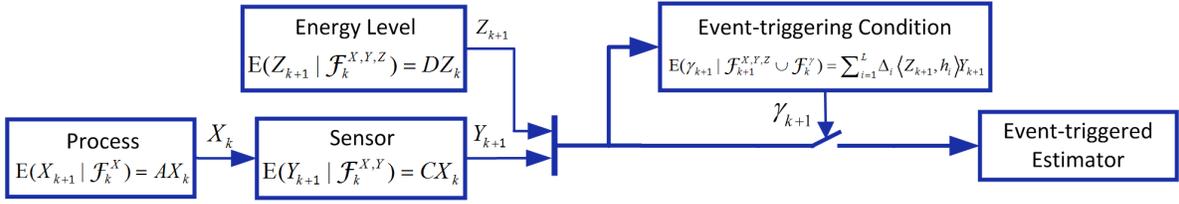
Figure 2.1: Block diagram of the event-based remote estimation system.

and $\mathbb{Z}_{-M:N} := \{-M, -M+1, \ldots, N\}$. For a set $\mathbb{M}$, let $|\mathbb{M}|$ be its cardinality. For a probability measure P ($\hat{P}$ or $\check{P}$), we use E ($\hat{E}$ or $\check{E}$, respectively) to represent the expectation operator. We use $[x_i]_{i \in \mathbb{N}_{1:n}}$ to denote $[x_1^T, \ldots, x_n^T]^T$ and use $E[x_i]_{i \in \mathbb{N}_{1:n}}$ to denote $\left[E[x_1]^T, \ldots, E[x_n]^T\right]^T$ (the same for $\hat{E}$ and $\check{E}$). 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. Let $x, y \in \mathbb{R}^m$, then $\langle x, y \rangle := x^T y$ denotes the inner product between x and y. $\mathbf{1}_{\{A\}}$ is the indicator function of set $\boldsymbol{A}$.

2.1.2 Problem description

Firstly, we introduce a hidden Markov model on the probability space $(\Omega, \mathcal{F}, P)$. 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 $S_X = \{e_1, e_2, \ldots, e_N\}$, where e_i is the unit vector in $\mathbb{R}^N$ with the ith element equal to 1. Let $\mathcal{F}_k^0 := \sigma\{X_0, \cdots, X_k\}$, and let $\mathcal{F}_k^X$ be the complete filtration generated by $\mathcal{F}_k^0$. By the Markov property, $P(X_{k+1} = e_i | \mathcal{F}_k^X) = P(X_{k+1} = e_i | X_k)$. Let $A := [a_i]_{i \in \mathbb{N}_{1:N}}$, $a_i := [a_{i,j}]_{j \in \mathbb{N}_{1:N}}^T$, where $a_{i,j} := P(X_{k+1} = e_i | X_k = e_j)$, such that $\sum_{i=1}^N a_{i,j} = 1$. Then

$$E(X_{k+1} | \mathcal{F}_k^X) = E(X_{k+1} | X_k) = A X_k. \tag{2.1}$$

Let Y_{k+1} be a sensor measurement process of X_k, which takes values in a finite-state space S_Y. Let the cardinality of the state space S_Y of Y be M, then S_Y can be identified with $S_Y = \{f_1, f_2, \ldots, f_M\}$, where f_i is the unit vector in $\mathbb{R}^M$ with the ith element equal to 1. Write $C := [c_i]_{i \in \mathbb{N}_{1:M}}$, $c_i := [c_{i,j}]_{j \in \mathbb{N}_{1:N}}^T$, where $c_{i,j} := P(Y_{k+1} = f_i | X_k = e_j)$, so that $\sum_{i=1}^M c_{i,j} = 1$ and $c_{i,j} \geq 0$. Let $\mathcal{F}_k^Y$ be the completion of the σ-field on Ω generated by $Y_0, Y_1, \ldots, Y_k$. Since Y_{k+1} measures X_k, we have

$$E(Y_{k+1} | \mathcal{F}_k^X \cup \mathcal{F}_k^Y) = E(Y_{k+1} | X_k) = C X_k. \tag{2.2}$$

We assume the measurement process is associated with an energy harvesting sensor, in which a rechargeable battery is utilized to store the energy harvested; we model the

battery level dynamics with a finite-state Markov chain, which can be efficiently utilized to capture the energy-harvesting dynamics, particularly when the amount of energy harvested by the sensor at each time instant can be modeled as an i.i.d. random variable [20, 27, 71]. The energy level of the sensor at time instant k is denoted as Z_k, which is independent of X_k and Y_k, taking value in a finite-state space S_Z. Let the cardinality of the state space S_Z be L, then S_Z can be identified with $S_Z = \{h_1, h_2, \ldots, h_L\}$, where h_i is the unit vector in $\mathbb{R}^L$ with ith element equal to 1. Let $\mathcal{F}_k^Z$ be the completion of the σ-field on Ω generated by $Z_1, Z_2, \ldots, Z_k$, where Z is independent of X and Y. Let $D := [d_i]_{i \in \mathbb{N}_{1:L}}$, $d_i := [d_{i,j}]_{j \in \mathbb{N}_{1:L}}^\top$, where $d_{i,j} := \mathrm{P}(Z_{k+1} = h_i | Z_k = h_j)$, such that $\sum_{i=1}^L d_{i,j} = 1$ and $d_{i,j} \geq 0$. Now we have

$$\mathrm{E}(Z_{k+1} | \mathcal{F}_k^X \cup \mathcal{F}_k^Y \cup \mathcal{F}_k^Z) = \mathrm{E}(Z_{k+1} | Z_k) = D Z_k. \tag{2.3}$$

Now we introduce the remote estimation problem and the setup of the problem is shown in Fig. 2.1, where the state of the hidden Markov model is estimated by a remote estimator, based on the measurement information from the sensor through a wired/wireless communication channel. In particular, we consider a scenario that the sensor measures Y_k at each time instant k and decides whether to send the measurements to the remote estimator or not according to a stochastic energy-based event-triggering condition (ETC), in which the probability of transmitting the measurements to the estimator at time instant k depends on the realizations of both the observation Y_k and energy level Z_k. Let γ_k be the decision variable taking values in $S_\gamma = \{\alpha_1, \alpha_2\}$, where $\alpha_1 = [1\ 0]^\top$ represents that the measurements will not be sent to the remote estimator, and $\alpha_2 = [0\ 1]^\top$ represents that the measurement Y_k together with the energy level Z_k will be sent to the remote estimator. In this work, the energy-based ETC is stochastic and identified by

$$\mathrm{P}\left(\gamma_k = \alpha_1 | Z_k = h_i, Y_k = f_j\right) = \delta_{ij},$$
$$\mathrm{P}\left(\gamma_k = \alpha_2 | Z_k = h_i, Y_k = f_j\right) = 1 - \delta_{ij}.$$

For notational brevity, we define a transformation matrix Δ_i as

$$\Delta_i = \begin{bmatrix} \delta_{i1} & \cdots & \delta_{iM} \\ 1 - \delta_{i1} & \cdots & 1 - \delta_{iM} \end{bmatrix}.$$

In this way, we have

$$\mathrm{E}\left(\gamma_k | \mathcal{F}_k^X \cup \mathcal{F}_k^Y \cup \mathcal{F}_k^Z \cup \mathcal{F}_{k-1}^\gamma\right) = \sum_{i=1}^L \Delta_i \langle Z_k, h_i \rangle Y_k, \tag{2.4}$$

14

where $\mathcal{F}_{k-1}^\gamma$ is the completion of the σ-field generated by $\gamma_1, \ldots, \gamma_{k-1}$. The measurement Y_k and energy level Z_k are received by the remote estimator at time instant k if and only if $\gamma_k = \alpha_2$. For $Z_k = h_i$ and $Y_k = f_j$, the ETC can be implemented by checking whether the realization of a uniformly distributed random variable ζ_k on $[0, 1]$ is greater than δ_{ij} or not. If $\zeta_k \leq \delta_{ij}$, $\gamma_k = \alpha_1$; otherwise $\gamma_k = \alpha_2$. When $\delta_{ij} \in \{0, 1\}$, the ETC can be implemented by testing whether $\delta_{ij} = 0$ holds or not for $Y_k = f_j$ and $Z_k = h_i$. Denoting $\hat{Y}_k = \{\gamma_k, \langle\gamma_k, \alpha_2\rangle Y_k, \langle\gamma_k, \alpha_2\rangle Z_k\}$ as the information received at the estimator at time instant k, we use $\hat{\mathcal{Y}}_k$ to denote the completion of the σ-field generated by $\{\hat{Y}_1, \hat{Y}_2, \ldots, \hat{Y}_k\}$, which describes all the information available at the remote estimator up to time instant k. In summary, our model now can be represented by (2.1)-(2.4).

Our main goal of this work is to obtain the conditional probability distribution of the system state on partially observed measurements. We propose a measure transformation to a new probability space under which, the state and observations are independent of each other. We derive the conditional probability distribution under the new probability measure, then transform it back to the one under the "real-world" measure by normalization.

2.1.3 Estimate of the system state

2.1.3.1 Constructions of the first reference measure $\hat{P}$

Consider a new reference measure $\hat{P}$, under which we still have

$$\hat{E}\left(X_{k+1}|\mathcal{F}_k^X\right) = \hat{E}(X_{k+1}|X_k) = AX_k, \tag{2.5}$$

while Y_{k+1} are uniformly i.i.d. satisfying

$$\hat{P}\left(Y_{k+1} = f_i\big|\mathcal{F}_k^X \cup \mathcal{F}_k^Y\right) = \hat{P}\left(Y_{k+1} = f_i\right) = 1/M. \tag{2.6}$$

According to equation (2.6), processes X and Y are independent under the first reference measure $\hat{P}$, which is well used in the derivations of the conditional distributions under $\hat{P}$. We still assume

$$\hat{E}\left(Z_{k+1} = h_i\big|\mathcal{F}_k^X \cup \mathcal{F}_k^Y \cup \mathcal{F}_k^Z\right) = DZ_k, \tag{2.7}$$

$$\hat{E}\left(\gamma_k\big|\mathcal{F}_k^X \cup \mathcal{F}_k^Y \cup \mathcal{F}_k^Z \cup \mathcal{F}_{k-1}^\gamma\right) = \sum_{i=1}^L \Delta_i \langle Z_k, h_i \rangle Y_k. \tag{2.8}$$

To link the new measure with the original "real-world" measure, we define a map from $\hat{P}$ to P:

$$\frac{dP}{d\hat{P}}\bigg|_{\mathcal{G}_{k+1}} = \hat{\Lambda}_{k+1}, \tag{2.9}$$

15

where $\mathcal{G}_{k+1}$ is $\mathcal{F}_{k+1}^X \cup \mathcal{F}_{k+1}^Y \cup \mathcal{F}_{k+1}^Z \cup \mathcal{F}_{k+1}^\gamma$, and

$$\hat{\Lambda}_{k+1} := \prod_{l=1}^{k+1} \hat{\lambda}_l, \tag{2.10}$$

$$\hat{\lambda}_{k+1} := M \sum_{i=1}^M \langle CX_k, f_i \rangle \langle Y_{k+1}, f_i \rangle. \tag{2.11}$$

Lemma 2.1. *[17, Theorem 3.2, Chapter 2]. Suppose $(\Omega, \mathcal{F}, \mathrm{P})$ is a probability space and $\mathcal{G} \subset \mathcal{F}$ is a sub-σ-field. Suppose $\hat{\mathrm{P}}$ is another probability measure absolutely continuous with respect to P and with Radon-Nikodym derivative $\frac{d\mathrm{P}}{d\hat{\mathrm{P}}} = \hat{\Lambda}$. If ϕ is any P integrable random variable, then*

$$\mathrm{E}(\phi|\mathcal{G}) = \begin{cases} \frac{\hat{\mathrm{E}}(\hat{\Lambda}\phi|\mathcal{G})}{\hat{\mathrm{E}}(\hat{\Lambda}|\mathcal{G})}, & \text{if } \hat{\mathrm{E}}(\hat{\Lambda}|\mathcal{G}) > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{2.12}$$

Since $\hat{\mathcal{Y}}_k$ is a sub-σ-field of $\mathcal{G}_k$, the following result can be obtained.

Lemma 2.2. *[17, Theorem 3.3, Chapter 2]. If ϕ_k is a $\mathcal{G}$-adapted integrable sequence of random variables, then*

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

The above lemma provides a way of mapping the probability distribution under reference measure $\hat{\mathrm{P}}$ back to the "real-world" measure P. Based on Lemmas 2.1 and 2.2, we investigate the properties of the map defined in (2.9)-(2.11) from the new reference measure to the original probability measure, which is summarized in the following result.

Lemma 2.3. *If the model in (2.5)-(2.8) is mapped from the probability measure $\hat{\mathrm{P}}$ to the probability measure P via the Radon-Nikodym derivative in (2.9)-(2.11), then the obtained model satisfies the properties under measure P as stated in (2.1)-(2.4).*

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

$$\begin{aligned} \hat{\mathrm{E}}(\hat{\lambda}_{k+1}|\mathcal{G}_k) &= \hat{\mathrm{E}}(M \sum_{i=1}^M \langle CX_k, f_i \rangle \langle Y_{k+1}, f_i \rangle |\mathcal{G}_k) \\ &= M \sum_{i=1}^M \langle CX_k, f_i \rangle \hat{\mathrm{E}}(\langle Y_{k+1}, f_i \rangle |\mathcal{G}_k) \\ &= \sum_{i=1}^M \langle CX_k, f_i \rangle = 1. \end{aligned}$$

By assumption, X and Y are mutually independent under the new probability measure $\hat{\mathrm{P}}$. According to Lemma 2.1 and definition of X_{k+1}, we have

$$
\begin{aligned}
\mathrm{E}(X_{k+1}|\mathcal{G}_k) &= \frac{\hat{\mathrm{E}}(\hat{\Lambda}_{k+1}X_{k+1}|\mathcal{G}_k)}{\hat{\mathrm{E}}(\hat{\Lambda}_{k+1}|\mathcal{G}_k)} \\
&= \frac{\hat{\Lambda}_k \hat{\mathrm{E}}(\hat{\lambda}_{k+1}X_{k+1}|\mathcal{G}_k)}{\hat{\Lambda}_k \hat{\mathrm{E}}(\hat{\lambda}_{k+1}|\mathcal{G}_k)} \\
&= \hat{\mathrm{E}}\big[M \textstyle\sum_{i=1}^{M}\langle CX_k, f_i\rangle \langle Y_{k+1}, f_i\rangle X_{k+1}\big|\mathcal{G}_k\big] \\
&= M \textstyle\sum_{i=1}^{M}\langle CX_k, f_i\rangle \frac{1}{M}\hat{\mathrm{E}}(X_{k+1}|\mathcal{G}_k) \\
&= \hat{\mathrm{E}}(X_{k+1}|\mathcal{G}_k) = AX_k,
\end{aligned}
$$

By repeated conditioning and $\mathcal{F}_k^X \subset \mathcal{G}_k$,

$$
\mathrm{E}(X_{k+1}|\mathcal{F}_k^X) = \mathrm{E}(\mathrm{E}(X_{k+1}|\mathcal{G}_k)|\mathcal{F}_k^X) = AX_k.
$$

Noticing that $[\langle Y_{k+1}, f_j\rangle]_{j\in\mathbb{N}_{1:M}} = Y_{k+1}$ and $[\langle CX_k, f_j\rangle]_{j\in\mathbb{N}_{1:M}} = CX_k$, we have

$$
\begin{aligned}
&\hat{\mathrm{E}}(\hat{\lambda}_{k+1}Y_{k+1}|\mathcal{G}_k) \\
&= \hat{\mathrm{E}}\big[Y_{k+1}M \textstyle\sum_{i=1}^{M}\langle CX_k, f_i\rangle \langle Y_{k+1}, f_i\rangle\big|\mathcal{G}_k\big] \\
&= \big[\hat{\mathrm{E}}\big[\langle Y_{k+1}, f_j\rangle M \textstyle\sum_{i=1}^{M}\langle CX_k, f_i\rangle \langle Y_{k+1}, f_i\rangle\big|\mathcal{G}_k\big]\big]_{j\in\mathbb{N}_{1:M}} \\
&= \big[\hat{\mathrm{E}}\big(M\langle Y_{k+1}, f_j\rangle \langle CX_k, f_j\rangle\big|\mathcal{G}_k\big)\big]_{j\in\mathbb{N}_{1:M}} = CX_k.
\end{aligned}
$$

From Lemma 2.1 and the definition of Y_{k+1}, we have

$$
\mathrm{E}(Y_{k+1}|\mathcal{G}_k) = \hat{\mathrm{E}}(\hat{\lambda}_{k+1}Y_{k+1}|\mathcal{G}_k) = CX_k.
$$

By repeated conditioning and $(\mathcal{F}_k^X \cup \mathcal{F}_k^Y) \subset \mathcal{G}_k$,

$$
\mathrm{E}\left(Y_{k+1}|\mathcal{F}_k^X \cup \mathcal{F}_k^Y\right) = \mathrm{E}\left(\mathrm{E}(Y_{k+1}|\mathcal{G}_k)|\mathcal{F}_k^X \cup \mathcal{F}_k^Y\right) = CX_k.
$$

By following a similar procedure as above, we obtain

$$
\mathrm{E}(Z_{k+1}|\mathcal{F}_k^Z) = AZ_k,
$$
$$
\mathrm{E}(\gamma_k|\mathcal{F}_k^X \cup \mathcal{F}_k^Y \cup \mathcal{F}_k^Z \cup \mathcal{F}_{k-1}^\gamma) = \textstyle\sum_{i=1}^{L}\Delta\langle Z_k, h_i\rangle Y_k.
$$

The proof is completed. $\qquad\square$

From Lemma 2.3, we can see that the map in (2.9) builds a link from the reference measure to the original "real-world" measure. By mapping the new model satisfying

17

(2.5)-(2.8) under the new reference measure $\hat{P}$ to the "real-world" probability measure, the system model as stated in (2.1)-(2.4) under P can be obtained. As shown in (2.6), Y_k is independent of all other variables, which provides convenience during the derivation of the optimal state estimator. The event-trigger γ_k and energy level Z_k are not involved in the map in (2.9), so that their properties are preserved under $\hat{P}$. Next, we will derive the probability distribution of the state conditioned on the information at the remote estimator under $\hat{P}$ using the properties described in (2.5)-(2.8); then map the estimate back to P to obtain the conditional probability distribution of the state in the "real-world" measure based on Lemma 2.2.

2.1.3.2 Recursive form of the state estimation

Define $\hat{q}_k^r := \hat{E}[\hat{\Lambda}_k \langle X_k, e_r \rangle | \hat{\mathcal{Y}}_k]$ and $\hat{p}_k^r := E[\langle X_k, e_r \rangle | \hat{\mathcal{Y}}_k]$. Respectively, write $\hat{q}_k = [\hat{q}_k^1 \; \ldots \; \hat{q}_k^N]^{\mathrm{T}}$ and $\hat{p}_k = [\hat{p}_k^1 \; \ldots \; \hat{p}_k^N]^{\mathrm{T}}$. From Lemma 2.2, we have

$$\hat{p}_k^r = \hat{q}_k^r / \|\hat{q}_k\|_1. \tag{2.13}$$

With the conditional probability distribution of state $\hat{p}_k^r$, we may use different kinds of estimators, e.g., the MMSE estimator, the maximum a posteriori (MAP) estimator or the mean absolute error (MAE) estimator. In order to obtain $\hat{p}_k$, we derive the recursive form of $\hat{q}_k^r$, which is presented in the following result.

Theorem 2.1. *For the system described in (2.5)-(2.7) with event-triggering scheme in (2.8), for $k \in \mathbb{N}$ and $r \in \mathbb{N}_{1:M}$, the unnormalized conditional probability distribution of the state satisfies*

$$\hat{q}_{k+1}^r = M \sum_{j=1}^N a_{r,j} \hat{q}_k^j \sum_{i=1}^M c_{i,j} \hat{E}\big[\langle Y_{k+1}, f_i \rangle | \hat{\mathcal{Y}}_{k+1}\big],$$

where $\hat{E}\big[\langle Y_{k+1}, f_i \rangle | \hat{\mathcal{Y}}_{k+1}\big] = \langle Y_{k+1}, f_i \rangle$ *if* $\gamma_{k+1} = \alpha_2$; *and*

$$\hat{E}[\langle Y_{k+1}, f_i \rangle | \hat{\mathcal{Y}}_{k+1}] = \frac{\sum_{t_1=1}^L \delta_{t_1 i} \hat{P}(Z_{k+1} = h_{t_1} | \hat{\mathcal{Y}}_k)}{\sum_{t=1}^L \hat{P}(Z_{k+1} = h_t | \hat{\mathcal{Y}}_k) \sum_{s=1}^M \delta_{ts}},$$

with $\hat{P}(Z_{k+1} = h_t | \hat{\mathcal{Y}}_k) = \sum_{j_1=1}^L d_{t,j_1} \hat{E}\big(\langle Z_k, h_{j_1} \rangle | \hat{\mathcal{Y}}_k\big)$ *if* $\gamma_{k+1} = \alpha_1$.

Proof. By repeated conditioning, for $r \in \mathbb{N}_{1:N}$, we have

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

$$= M\hat{E}\big[\langle AX_k, e_r \rangle \hat{\Lambda}_k \sum_{i=1}^M \langle CX_k, f_i \rangle \langle Y_{k+1}, f_i \rangle \big| \hat{\mathcal{Y}}_{k+1}\big].$$

18

Recalling that $\langle AX_k, e_r \rangle = \sum_{j=1}^{N} a_{r,j} \langle X_k, e_j \rangle$ and $\langle CX_k, f_i \rangle = \sum_{j_0=1}^{N} c_{i,j} \langle X_k, e_{j_0} \rangle$, we have

$$
\begin{aligned}
\hat{q}_{k+1}^r &= M\hat{\mathrm{E}}\big[\textstyle\sum_{j=1}^{N} a_{r,j} \langle X_k, e_j \rangle \hat{\Lambda}_k \sum_{i=1}^{M} \sum_{j_0=1}^{N} c_{i,j_0} \langle X_k, e_{j_0} \rangle \langle Y_{k+1}, f_i \rangle | \hat{\mathcal{Y}}_{k+1}\big] \\
&= M\hat{\mathrm{E}}\big[\textstyle\sum_{j=1}^{N} a_{r,j} \langle X_k, e_j \rangle \hat{\Lambda}_k \sum_{i=1}^{M} c_{i,j} \langle Y_{k+1}, f_i \rangle | \hat{\mathcal{Y}}_{k+1}\big],
\end{aligned}
$$

where we use the fact that $\langle X_k, e_j \rangle \langle X_k, e_{j_0} \rangle \neq 0$ only if $j = j_0$. Since $\langle X_k, e_j \rangle \hat{\Lambda}_k$ is independent of $\hat{Y}_{k+1}$, we have

$$
\begin{aligned}
\hat{q}_{k+1}^r &= M \textstyle\sum_{j=1}^{N} a_{r,j} \hat{\mathrm{E}}\big[\langle X_k, e_j \rangle \hat{\Lambda}_k | \hat{\mathcal{Y}}_k\big] \sum_{i=1}^{M} c_{i,j} \hat{\mathrm{E}}\big[\langle Y_{k+1}, f_i \rangle | \hat{\mathcal{Y}}_{k+1}\big] \\
&= M \textstyle\sum_{j=1}^{N} a_{r,j} \hat{q}_k^j \sum_{i=1}^{M} c_{i,j} \hat{\mathrm{E}}\big[\langle Y_{k+1}, f_i \rangle | \hat{\mathcal{Y}}_{k+1}\big].
\end{aligned}
$$

Now our main focus is to obtain $\hat{\mathrm{E}}\big[\langle Y_{k+1}, f_i \rangle | \hat{\mathcal{Y}}_{k+1}\big]$ and we consider two scenarios. 1) If $\gamma_{k+1} = \alpha_2$, the event is triggered and measurement is received, so the estimator knows the exact values of Y_{k+1} and Z_{k+1}. In this way, we have $\hat{\mathrm{E}}\big[\langle Y_{k+1}, f_i \rangle | \hat{\mathcal{Y}}_{k+1}\big] = \langle Y_{k+1}, f_i \rangle$. 2) If $\gamma_{k+1} = \alpha_1$, the event is not triggered and the estimator does not know the exact values of Y_{k+1} and Z_{k+1}. The estimate becomes

$$
\begin{aligned}
\hat{\mathrm{E}}[\langle Y_{k+1}, f_i \rangle | \hat{\mathcal{Y}}_{k+1}] &= \hat{\mathrm{P}}\big(Y_{k+1} = f_i | \hat{\mathcal{Y}}_k, \gamma_{k+1} = \alpha_1\big) \\
&= \frac{\hat{\mathrm{P}}\big(Y_{k+1} = f_i, \gamma_{k+1} = \alpha_1 | \hat{\mathcal{Y}}_k\big)}{\hat{\mathrm{P}}\big(\gamma_{k+1} = \alpha_1 | \hat{\mathcal{Y}}_k\big)}.
\end{aligned}
$$

Here we focus on the numerator part first. According to the Bayes' theorem,

$$
\begin{aligned}
&\hat{\mathrm{P}}\big(Y_{k+1} = f_i, \gamma_{k+1} = \alpha_1 | \hat{\mathcal{Y}}_k\big) \\
&= \hat{\mathrm{P}}\big(\gamma_{k+1} = \alpha_1 | Y_{k+1} = f_i, \hat{\mathcal{Y}}_k\big) \hat{\mathrm{P}}(Y_{k+1} = f_i | \hat{\mathcal{Y}}_k) \\
&= \frac{1}{M} \textstyle\sum_{t=1}^{L} \hat{\mathrm{P}}\big(\gamma_{k+1} = \alpha_1 | Z_{k+1} = h_t, Y_{k+1} = f_i, \hat{\mathcal{Y}}_k\big) \hat{\mathrm{P}}\big(Z_{k+1} = h_t | Y_{k+1} = f_i, \hat{\mathcal{Y}}_k\big) \\
&= \frac{1}{M} \textstyle\sum_{t=1}^{L} \delta_{ti} \hat{\mathrm{P}}\big(Z_{k+1} = h_t | \hat{\mathcal{Y}}_k\big),
\end{aligned}
$$

where the last equality is due to that γ_{k+1} is independent of $\hat{\mathcal{Y}}_k$, and Z_{k+1} is independent of Y_{k+1}. Following a similar procedure as in evaluating the numerator, we have the denominator as follows

$$
\hat{\mathrm{P}}\big(\gamma_{k+1} = \alpha_1 | \hat{\mathcal{Y}}_k\big) = \frac{1}{M} \sum_{t=1}^{L} \hat{\mathrm{P}}\big(Z_{k+1} = h_t | \hat{\mathcal{Y}}_k\big) \sum_{s=1}^{M} \delta_{ts}.
$$

By repeated conditioning, we have

$$
\begin{aligned}
\hat{\mathrm{P}}\big(Z_{k+1} = h_t | \hat{\mathcal{Y}}_k\big) &= \hat{\mathrm{E}}\big[\langle Z_{k+1}, h_t \rangle | \hat{\mathcal{Y}}_k\big] \\
&= \textstyle\sum_{j_1=1}^{L} d_{t,j_1} \hat{\mathrm{E}}\big[\langle Z_k, h_{j_1} \rangle | \hat{\mathcal{Y}}_k\big],
\end{aligned}
$$

which completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

From the theorem above, the unnormalized conditional probability distribution under the first reference measure $\hat{P}$ is given in a recursive form and can be transformed back to the conditional probability distribution under P by equation (2.13). If we consider the simple scenario: there is only one state $Z_k = h_1$ for battery energy at the sensor and the probability of not transmitting $P(\gamma_{k+1} = \alpha_1 | Y_{k+1} = f_i) = 1$ for $i \in \mathbb{I}_{k+1}$ and $P(\gamma_{k+1} = \alpha_1 | Y_{k+1} = f_j) = 0$ for $j \in \mathbb{N}_{1:M} \backslash \mathbb{I}_{k+1}$. In this scenario, the conditional probability distribution of X_{k+1} under P derived in our work is consistent with the result considering deterministic ETC under reliable communication channels (Theorem 4 in [60]), though the notation in [60] is slightly different from this investigation. The theorem above demonstrates that the unnormalized distribution $\hat{q}_{k+1}$ of X_k conditioned on the past hybrid measurement information $\hat{\mathcal{Y}}_k$ evolves recursively according to a map based on $\hat{q}_k$. However, the estimate of the energy level is not obtained yet; for this we propose a second level of the probability measure change to obtain the conditional probability distribution of the energy level under the reference measure $\hat{P}$, namely, $\hat{E}\big[\langle Z_k, h_{j_1}\rangle | \hat{\mathcal{Y}}_k\big]_{j_1 \in \mathbb{N}_{1:M}}$.

2.1.4   Estimate of the energy level

In order to acquire the estimate of Z_k conditioned on the information available at the remote estimator under $\hat{P}$, we define a second reference measure $\check{P}$ and propose a second map from $\check{P}$ to $\hat{P}$. We may focus on Y_k, Z_k and γ_k, and do not need to consider X_k in the second map since Z_k is independent of X_k under $\hat{P}$.

2.1.4.1   Construction of a second reference measure $\check{P}$

Consider a new measure, under which we have

$$\check{P}\left(Y_{k+1} = f_i \big| \mathcal{F}_k^Y \cup \mathcal{F}_k^Z \cup \mathcal{F}_k^\gamma\right) = 1/M, \tag{2.14}$$

$$\check{P}\left(Z_{k+1} = h_i \big| \mathcal{F}_k^Y \cup \mathcal{F}_k^Z \cup \mathcal{F}_k^\gamma\right) = 1/L, \tag{2.15}$$

$$\check{P}\left(\gamma_k = \alpha_i \big| \mathcal{F}_k^Y \cup \mathcal{F}_k^Z \cup \mathcal{F}_{k-1}^\gamma\right) = 1/2. \tag{2.16}$$

The above properties show that Y_k, Z_k and γ_k are i.i.d. and pairwise independent. Denote $\mathcal{F}_k^Y \cup \mathcal{F}_k^Z \cup \mathcal{F}_k^\gamma$ as $\check{\mathcal{G}}_k$. Define a mapping from $\check{P}$ to $\hat{P}$

$$\left.\frac{\mathrm{d}\hat{P}}{\mathrm{d}\check{P}}\right|_{\check{\mathcal{G}}_k} = \check{\Lambda}_k, \tag{2.17}$$

where

$$\check{\Lambda}_k = \prod_{l=1}^k \check{\lambda}_l, \tag{2.18}$$

$$\check{\lambda}_k = \left[2\sum_{j=1}^2 \sum_{i=1}^L \langle \Delta_i \langle Z_k, h_i \rangle Y_k, \alpha_j \rangle \langle \gamma_k, \alpha_j \rangle\right]$$
$$\cdot \left[L \sum_{r=1}^L \langle DZ_{k-1}, h_r \rangle \langle Z_k, h_r \rangle\right]. \tag{2.19}$$

Based on the map in (2.17) and Lemma 2.1, the following result is obtained.

Lemma 2.4. *If the model in (2.14)-(2.16) is mapped from $\check{P}$ to $\hat{P}$ by the Radon-Nikodym derivative in (2.17)-(2.19), then the obtained model satisfies the dynamics illustrated in (2.6)-(2.8) under $\hat{P}$.*

Proof. This lemma can be proved by following similar procedure as in Lemma 2.3. The detailed proof is omitted due to length limitation. $\square$

This result establishes a map from the second reference measure $\check{P}$ to the first reference measure $\hat{P}$, based on which, the estimation problem of the energy level under the first reference measure $\hat{P}$ can be solved by solving an unnormalized estimation problem under the second reference measure $\check{P}$ and mapping the result back to $\hat{P}$.

2.1.4.2   Recursive form of the energy level estimation

Next, we derive a recursive form of the conditional probability distribution of the energy level under $\check{P}$. Define $\check{u}_k^s := \check{E}[\check{\Lambda}_k \langle Z_k, h_s \rangle | \hat{\mathcal{Y}}_k]$, and $\check{v}_k^s := \hat{E}[\langle Z_k, h_s \rangle | \hat{\mathcal{Y}}_k]$. Write $\check{u}_k = \begin{bmatrix} \check{u}_k^1 & \dots & \check{u}_k^L \end{bmatrix}^{\mathrm{T}}$ and $\check{v}_k = \begin{bmatrix} \check{v}_k^1 & \dots & \check{v}_k^L \end{bmatrix}^{\mathrm{T}}$, respectively. Recall that $\hat{\mathcal{Y}}_k \subset \check{\mathcal{G}}_k$. From Lemma 2, we have

$$\check{v}_k^s = \check{u}_k^s / \|\check{u}_k\|_1. \tag{2.20}$$

Now we focus on the recursive form of unnormalized conditional probability distribution of the energy level $\check{u}_{k+1}$, which is summarized as follows.

Theorem 2.2. *For the system described by (2.14)-(2.16), for $k \in \mathbb{N}$ and $s \in \mathbb{N}_{1:L}$, the unnormalized conditional distribution of the energy level $\check{u}_{k+1}^s$ satisfies*

$$\check{u}_{k+1}^s = 2L\langle Z_{k+1}, h_s \rangle \langle \Delta_s Y_{k+1}, \alpha_2 \rangle \sum_{j_2=1}^L d_{s,j_2} \check{u}_k^{j_2},$$

if $\gamma_{k+1} = \alpha_2$; and if $\gamma_{k+1} = \alpha_1$

$$\check{u}_{k+1}^s = \frac{2}{M} \sum_{i=1}^M \langle \Delta_s f_i, \alpha_1 \rangle \sum_{j_2=1}^L d_{s,j_2} \check{u}_k^{j_2}.$$

Proof. According to the definition of $\check{u}_{k+1}$, we have

$$
\begin{aligned}
\check{u}_{k+1}^s =&\check{\mathrm{E}}\big[\check{\Lambda}_{k+1}\langle Z_{k+1}, h_s\rangle\big|\hat{\mathcal{Y}}_{k+1}\big]\\
=&\check{\mathrm{E}}\Big[\check{\Lambda}_k\big(2\sum_{j=1}^{2}\sum_{i=1}^{L}\langle\Delta_i\langle Z_{k+1}, h_i\rangle Y_{k+1}, \alpha_j\rangle\langle\gamma_{k+1}, \alpha_j\rangle\big)\\
&\cdot\big(L\sum_{r=1}^{L}\langle DZ_k, h_r\rangle\langle Z_{k+1}, h_r\rangle\big)\langle Z_{k+1}, h_s\rangle\big|\hat{\mathcal{Y}}_{k+1}\Big].
\end{aligned}
$$

If $\gamma_{k+1} = \alpha_2$, Z_{k+1} and Y_{k+1} are received and measurable on $\hat{\mathcal{Y}}_{k+1}$, then the case is trivial and we have

$$
\begin{aligned}
\check{u}_{k+1}^s =&2L\langle Z_{k+1}, h_s\rangle\bigg(\sum_{i=1}^{L}\big\langle\Delta_i\langle Z_{k+1}, h_i\rangle Y_{k+1}, \alpha_2\big\rangle\bigg)\\
&\cdot\bigg(\sum_{r=1}^{L}\langle Z_{k+1}, h_r\rangle\sum_{j_2=1}^{L}d_{r,j_2}\check{\mathrm{E}}\Big[\check{\Lambda}_k\langle Z_k, h_{j_2}\rangle\big|\hat{\mathcal{Y}}_{k+1}\Big]\bigg)\\
=&2L\langle Z_{k+1}, h_s\rangle\langle\Delta_s Y_{k+1}, \alpha_2\rangle\sum_{j_2=1}^{L}d_{s,j_2}\check{u}_k^{j_2},
\end{aligned}
$$

where we used the fact that $\check{\Lambda}_k\langle Z_k, h_{j_2}\rangle$ is independent of Z_{k+1} and Y_{k+1} under $\check{\mathrm{P}}$ to simplify the result.

If $\gamma_{k+1} = \alpha_1$, the event it not triggered and the estimator does not know exact values of Z_{k+1} and Y_{k+1}. In this case, by using Bayes' rule, we have

$$
\begin{aligned}
\check{u}_{k+1}^s =&\check{\mathrm{E}}\big[\check{\Lambda}_k\big(2\sum_{i=1}^{L}\langle\Delta_i\langle Z_{k+1}, h_i\rangle Y_{k+1}, \alpha_1\rangle\big)\\
&\cdot\big(L\sum_{r=1}^{L}\langle DZ_k, h_r\rangle\langle Z_{k+1}, h_r\rangle\big)\langle Z_{k+1}, h_s\rangle\big|\hat{\mathcal{Y}}_{k+1}\big]\\
=&\sum_{s_1=1}^{L}\check{\mathrm{P}}\big(Z_{k+1} = h_{s_1}\big|\hat{\mathcal{Y}}_k, \gamma_{k+1} = \alpha_1\big)\check{\mathrm{E}}\big[\check{\Lambda}_k\big(2\sum_{i=1}^{L}\langle\Delta_i\langle Z_{k+1}, h_i\rangle Y_{k+1}, \alpha_1\rangle\big)\\
&\cdot\big(L\sum_{r=1}^{L}\langle DZ_k, h_r\rangle\langle Z_{k+1}, h_r\rangle\big)\langle Z_{k+1}, h_s\rangle\big|\hat{\mathcal{Y}}_k, Z_{k+1} = h_{s_1}\big].
\end{aligned}
$$

Recall that $\check{\mathrm{P}}(Z_{k+1} = h_{s_1}|\hat{\mathcal{Y}}_k, \gamma_{k+1} = \alpha_1) = 1/L$ under $\check{\mathrm{P}}$, then (2.21) becomes

$$
\begin{aligned}
\check{u}_{k+1}^s =&2\sum_{s_1=1}^{L}\check{\mathrm{E}}\big[\check{\Lambda}_k\big(\sum_{i=1}^{L}\langle\Delta_i\langle h_{s_1}, h_i\rangle Y_{k+1}, \alpha_1\rangle\big)\\
&\cdot\big(\sum_{r=1}^{L}\langle DZ_k, h_r\rangle\langle h_{s_1}, h_r\rangle\big)\langle h_{s_1}, h_s\rangle\big|\hat{\mathcal{Y}}_k\big]\\
=&2\check{\mathrm{E}}\big[\check{\Lambda}_k\langle\Delta_s Y_{k+1}, \alpha_1\rangle\langle DZ_k, h_s\rangle\big|\hat{\mathcal{Y}}_k\big].
\end{aligned}
$$

Now, we consider all possible values for Y_{k+1},

$$
\begin{aligned}
\check{u}_{k+1}^s =&\sum_{i=1}^{M}\check{\mathrm{P}}\big(Y_{k+1} = f_i\big|\hat{\mathcal{Y}}_k\big)2\check{\mathrm{E}}\big[\check{\Lambda}_k\langle\Delta_s Y_{k+1}, \alpha_1\rangle\langle DZ_k, h_s\rangle\big|\hat{\mathcal{Y}}_k, Y_{k+1} = f_i\big]\\
=&\frac{2}{M}\sum_{i=1}^{M}\langle\Delta_s f_i, \alpha_1\rangle\sum_{j_2=1}^{L}d_{s,j_2}\check{u}_k^{j_2},
\end{aligned}
$$

which proves the theorem. $\qquad\square$

By introducing the second level of probability measure change, the unnormalized conditional distribution of the energy level is derived in a recursive form, which evolves according to a linear map under the second reference measure $\check{P}$. According to Theorem 2.2, the distribution of the energy level under the first reference measure $\hat{P}$, namely, $\hat{E}[\langle Z_k, h_s \rangle | \hat{\mathcal{Y}}_k]$, is obtained by mapping $\check{u}_{k+1}$ back to $\hat{P}$ via normalization in equation (2.20).

Combining Theorem 2.1 and Theorem 2.2, the unnormalized probability distributions of the state and energy can be obtained. First, the result in Theorem 2.2 is utilized to calculate the unnormalized conditional distribution of the energy level under the second reference measure $\check{P}$, which evolves according to a linear map in a recursive form. Next, the conditional distribution of the energy level under the first reference measure $\hat{P}$ is obtained by normalization. Then $\hat{E}[\langle Z_k, h_s \rangle | \hat{\mathcal{Y}}_k]$ is substituted into the recursive form of unnormalized conditional probability distribution of the state under the first reference measure $\hat{P}$ as stated in Theorem 2.1 to obtain $\hat{q}_{k+1}$. Finally, the probability distribution of the state conditioned on the combined information at the remote estimator under the "real-world" probability measure is obtained by normalization in equation (2.13), based on which the MMSE estimator can be derived. Once we obtain the state estimate, the transformations and reference probability measures are not needed in the implementation of the estimates since the reference measures are only utilized as mathematical tools in the derivations.

2.1.5 Simulation results

In this section, the proposed results are applied to state estimation of linear Gaussian systems and a numerical example is used to illustrate the effectiveness of the proposed event-based state estimation algorithm. We consider a stable first-order linear Gaussian system which is parameterized as

$$x_{k+1} = ax_k + w_k, \tag{2.21}$$

where w_k is a Gaussian noise with covariance Q_w, x_k is the state with initial state x_0, which is zero-mean Gaussian with covariance P_0. For the measurement process, we assume

$$y_{k+1} = cx_k + v_k, \tag{2.22}$$

where v_k is a Gaussian noise with covariance Q_v. We assume w_k, v_k and x_0 are mutually independent. At steady state, x_k becomes a stationary process with distribution

23

$\mu_x(x) = \exp(-\frac{x^2}{2Q_x})/\sqrt{2\pi Q_x}$, where $Q_x = Q_w/(1-a^2)$. The utilization of digital control, where an analog-to-digital converter (ADC) is used to convert the continuous physical quantities to digital ones with finite digits, leads inevitably to quantization. ADC is necessary when the controller is implemented in a computer, or a digital network is used to transmit the measurement to the remote estimator.

In this example, we quantize the state and measurement of the scalar linear Gaussian system into finite states and obtain a corresponding hidden Markov model using the quantization approach introduced in [61]. The state x_k is quantized into a total of N regions, which are denoted by $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$. By using a unit-vector parameterization technique, a finite-state process $\{X_k\}$ taking values in $\{e_1, e_2, \ldots, e_N\}$ is used to represent the quantized state of system by associating each region $\mathbf{x}_i$ with e_i, i.e, $x_k \in \mathbf{x}_i \iff X_k = e_i$. The quantization regions $\mathbf{x}_i$ are given by $\mathbf{x}_i := \{x \in \mathbb{R} \mid \underline{x}_i \le x \le \bar{x}_i\}$, where $\underline{x}_i$ and $\bar{x}_i$ are the upper and lower bounds of $\mathbf{x}_i$. We consider a simple algorithm of choosing $\mathbf{x}_i$, where $\mathbf{x}_1 = [-\infty, \underline{x}]$, $\mathbf{x}_N = [\bar{x}, \infty]$, and the other $N - 2$ quantization regions are uniformly distributed between $[\underline{x}, \bar{x}]$. For $i \in \mathbb{N}_{2:N-1}$, we have $\underline{x}_i = \bar{x}_{i-1}$, $\bar{x}_i = \underline{x}_i + (\bar{x} - \underline{x})/(N-2)$. Noticing that a and Q_w are time-invariant, the transition probability matrix in (2.1) can be calculated offline using standard numerical integration techniques as

$$
\begin{aligned}
a_{i,j} &= \mathrm{P}(X_{k+1} = e_i | X_k = e_j) \\
&= \frac{1}{\int_{\mathbf{x}_j} \mu_x(\xi)\mathrm{d}\xi} \int_{\mathbf{x}_j} \left[Q(\frac{\underline{x}_i - ax}{\sqrt{Q_w}}) - Q(\frac{\bar{x}_i - ax}{\sqrt{Q_w}}) \right] \mu_x(x)\,\mathrm{d}x.
\end{aligned} \tag{2.23}
$$

Following a similar quantization procedure, Y_k is quantized into a total of M regions, which are denoted by $\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_M\}$, and $\{f_1, f_2, \ldots, f_M\}$ is used to represent the quantized measurement of system by associating $\mathbf{y}_i$ with f_i, i.e. $y_k \in \mathbf{y}_i \iff Y_k = f_i$, with $\mathbf{y}_i := \{y \in \mathbb{R} \mid \underline{y}_i \le x \le \bar{y}_i\}$. For $i \in \mathbb{N}_{2:M-1}$, $\mathbf{y}_1 = [-\infty, \underline{y}]$, $\mathbf{y}_N = [\bar{y}, \infty]$, and $\underline{y}_i = \bar{y}_{i-1}$, $\bar{y}_i = \underline{y}_i + (\bar{y} - \underline{y})/(M-2)$. Combining equation (2.22), we have the transition probability as

$$
\begin{aligned}
c_{i,j} &= \mathrm{P}(Y_{k+1} = e_i | X_k = e_j) \\
&= \frac{1}{\int_{\mathbf{x}_j} \mu_x(\xi)\mathrm{d}\xi} \int_{\mathbf{x}_j} \left[Q(\frac{\underline{y}_i - cx}{\sqrt{Q_v}}) - Q(\frac{\bar{y}_i - cx}{\sqrt{Q_v}}) \right] \mu_x(x)\,\mathrm{d}x.
\end{aligned} \tag{2.24}
$$

In the following, we use a numerical example to show the implementation of the estimator. Take $a = 0.9$, $c = 0.5$, $Q_w = 0.3$ and $Q_v = 0.3$. For the quantization regions, consider $N = M = 128$, which corresponds to a 7-bit ADC, with $\underline{x} = -8$, $\bar{x} = 8$, $\underline{y} = -5$, and $\bar{y} = 5$. We calculate the transition probability matrices A and C of the

hidden Markov model according to (2.23) and (2.24). We assume the energy harvested at each time instant are i.i.d. random variables and the probability of harvesting 0,1 and 2 units of energy are 0.75, 0.2 and 0.05 correspondingly. At time instant k, the sensor harvests energy from the environment first, then store the harvested energy into the sensor's battery; when charging is finished, the energy level is denoted as z_k. After that, the sensor determines if the measurements are transmitted to the estimator. We assume 1 unit of battery's energy is consumed if the transmission is performed. If the battery is out of power, the sensor turns into sleeping mode with the data packet being dropped. We assume the minimum and maximum energy storage of the sensor's battery are 0 and 15 units of energy; the initial value z_0 is randomly chosen and known by the estimator. Suppose the 16 energy levels are represented by $\{h_1, ..., h_{16}\}$, where $Z_k = h_1$ refers to the sleeping mode and $Z_k = h_i$ refers to the scenario when $i - 1$ units of energy are stored in the battery. The sensor's battery level can be modeled by a Markov chain as described in (2.3), where the parameters in D can be obtained easily.

We define $\phi(\tau, \mu, \sigma) := \exp\left[-(\tau - \mu)^2/(2\sigma^2)\right]$ for $\tau \in \mathbb{Z}_{-n,n}$. We assume the sensor will examine the energy-based event-triggered condition as described by equation (2.4) to decide if measurements are transmitted or not, where δ_{ij}'s are chosen as $\delta_{1j} = 0$ and for $i \geq 2$

$$
\begin{aligned}
\delta_{ij} =&\mathrm{P}\left(\gamma_k = \alpha_1 | Z_k = h_i, Y_k = f_j\right) \\
=&1_{\{|j-\theta|<20-i\}}\phi(j, \theta, 30),
\end{aligned}
\tag{2.25}
$$

where θ is defined implicitly by $\tilde{Y}_k = f_\theta$ with $\tilde{Y}_k$ denoting the quantized value of the previously transmitted measurement $\tilde{y}$ at time instant k. Theorem 2.1 and Theorem 2.2 are combined to obtain the conditional probability distribution of the state under the "real-world" measure, based on which, the MMSE estimate of the state can be obtained. $\gamma_k = \alpha_2$ and $\gamma_k = \alpha_1$ are presented as 1 and 0 correspondingly in figures for convenience. The mean square error (MSE) of the estimates from the Kalman filter, the Kalman filter with intermittent observations (KF with IO) in [67], and the proposed method using the same communication sequence $\{\gamma_k\}$ are 0.570, 0.757 and 0.629 correspondingly as shown in Fig. 2.2, indicating that the estimation performance of the proposed method is better than KF with IO, which does not exploit the information contained in the stochastic triggering condition.

To further show the merits of the proposed energy-based ETC, we consider the normal ETC for comparison; to do this, we take the event-triggering condition to be
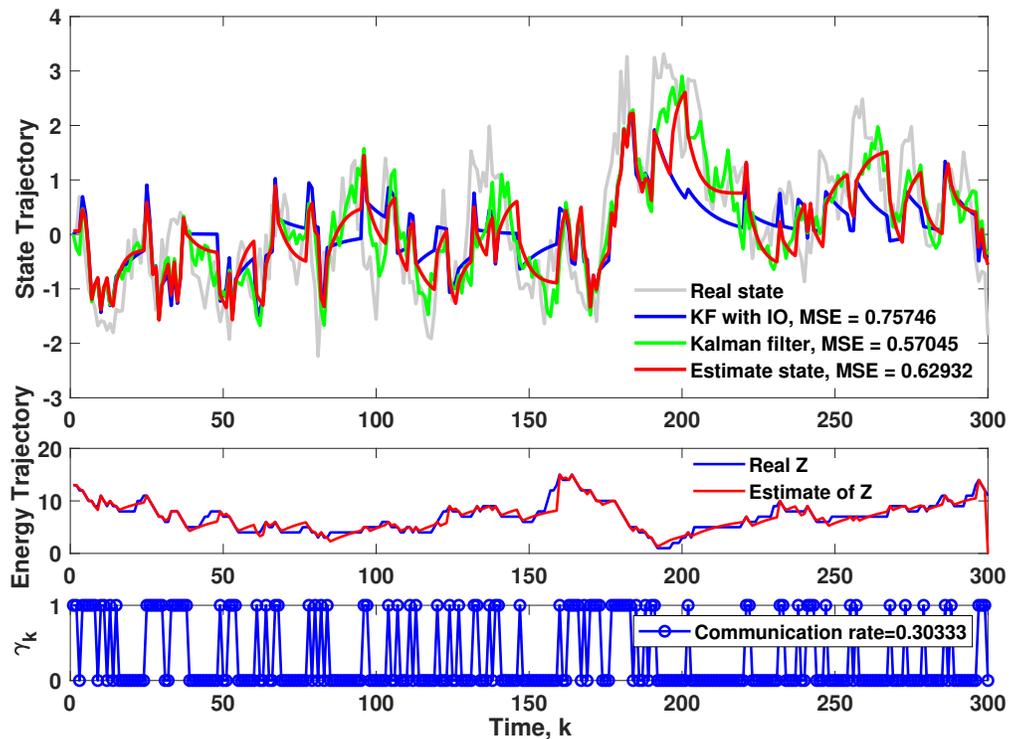
Figure 2.2: Performance comparison between three methods considering the energy-based ETC.

independent of the energy level such that

$$P\left(\gamma_k = \alpha_1 | Y_k = f_j, Z_k \neq h_1\right) = (\textstyle\sum_{l=1}^{L} \delta_{lj})/L,$$

and $P\left(\gamma_k = \alpha_1 | Y_k = f_j, Z_k = h_1\right) = 0$ (as no energy is available for transmission when $Z_k = h_1$). Since this ETC does not depend on the energy level in general, the energy level is not estimated for this case. All the other parameters of the system and realizations of x_k, y_k and the energy harvested at each time instant are exactly the same as those in Fig. 2.2. The real state, the estimates from the Kalman filter, KF with IO and the normal ETC are shown in Fig. 2.3. Although the normal ETC has a lightly higher average transmission rate, the MSE of the proposed estimate with energy-based ETC in Fig. 2.2 (0.629) is smaller than that with normal ETC in Fig. 2.3 (0.777), indicating that the energy-based ETC outperforms the normal ETC through proper energy management.

Next, the comparison of the energy level distributions while considering the energy-based ETC and normal ETC is shown in Fig. 2.4. Simulations are run for $T = 50,000$ time instants under the two different ETCs with the other parameters remaining the

26
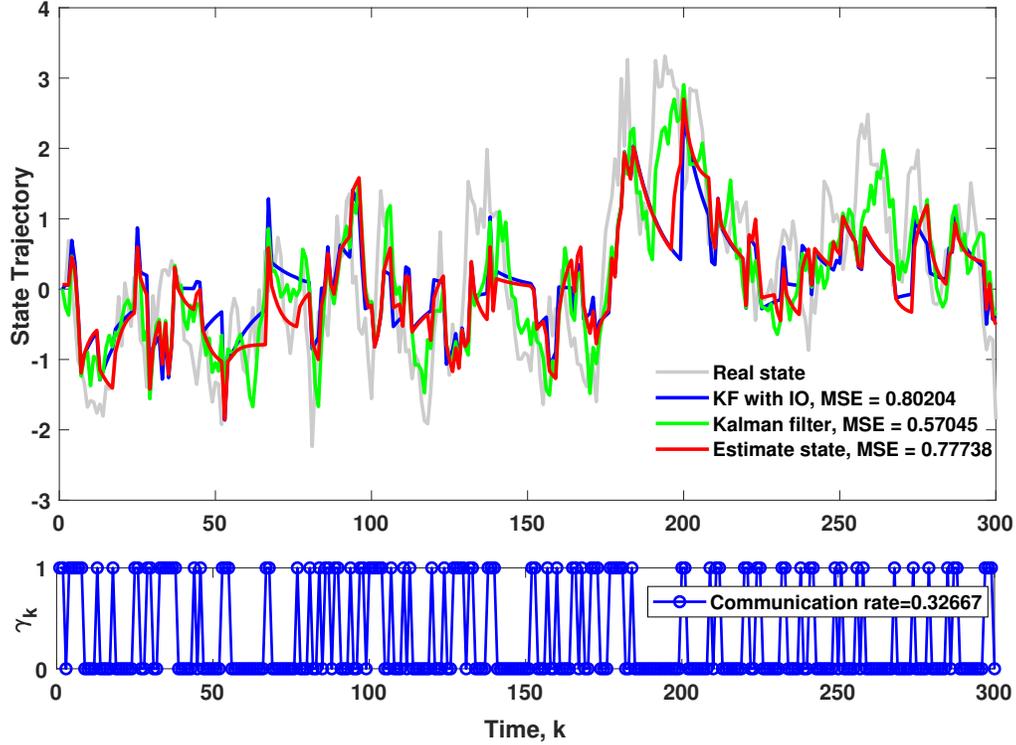
Figure 2.3: Performance comparison between three methods considering the normal ETC.

same. The average communication rates $\frac{1}{N}\sum_{k=1}^{N} 1_{\{\gamma_k=\alpha_2\}}$ are 0.30 in both scenarios. Fig. 2.4 shows that for the scenario with energy-based ETC, the energy distribution approximately follows a unimodal distribution centralized to the middle energy level and the sensor is hardly in the sleeping mode; for the scenario with normal ETC, the energy distribution is shifted to energy level 1, indicating that the utilization of the energy-based ETC leads to smarter sensor energy management in reducing the occurrence rate of the sensor's sleeping mode.

Finally, we focus on the energy-based ETC; the tradeoff between the communication rates and estimation performances under different energy levels[1] is shown in Fig. 2.5 ($T = 300,000$ time instants). The corresponding communication rate for one specific energy level i is defined as $\frac{1}{N_i}\sum_{Z_k=h_i} 1_{\{\gamma_k=\alpha_2\}}$, where N_i is the total number of time instants with energy level $Z_k = h_i$. When the energy level is relatively low, the communication rate by the left vertical axis is small; meanwhile, the estimation error of the proposed method shown by the right vertical axis is large, indicating that the estima-

---

[1] The cases when the energy level is equal to 0 and 15 are not considered, since these cases occur by very small probabilities (see Fig. 2.4).
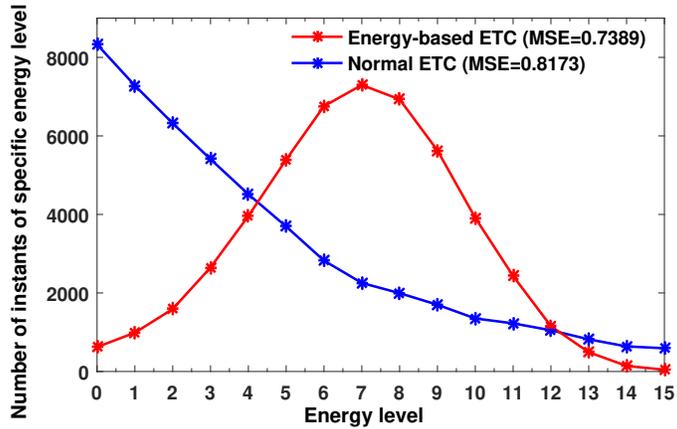
Figure 2.4: Comparison of the energy distributions obtained by using different ETC (MSE of the Kalman filter = 0.6379).
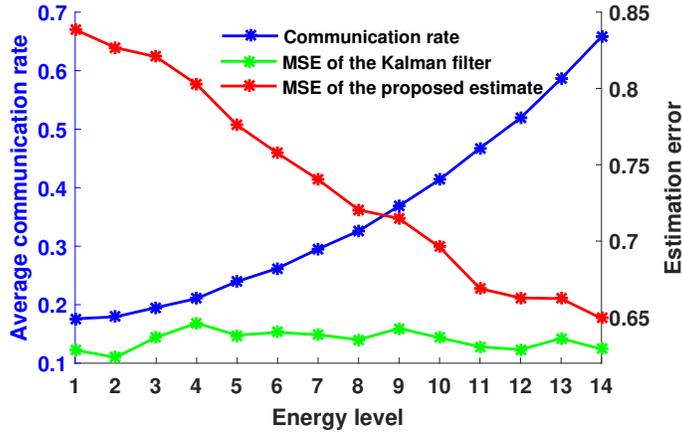
Figure 2.5: Tradeoff between communication rates and estimation performances over different energy levels.

tion performance is sacrificed to maintain low average communication rate when the battery's energy is in shortage.

2.1.6 Conclusion

In this section, an energy-based event-triggered state estimation problem for discrete hidden Markov models is investigated. The stochastic ETC proposed is controlled by the energy level of an energy harvesting sensor, which evolves according to a Markov chain itself. We propose a map using the reference measure approach and derive the recursive form of the unnormalized conditional probability distribution of the state under the first reference measure $\hat{P}$, which is based on the unnormalized conditional

28

distribution of the energy level under $\hat{P}$. To calculate the unnormalized conditional distribution of the energy level, a second reference measure is proposed. By introducing these two reference measures, the conditional probability distribution of the state in the "real-world" probability measure can be obtained, based on which, the MMSE estimator can be derived. Finally, the results are applied to a quantized linear Gaussian system and simulation results are given to demonstrate the effectiveness of the proposed method. For future work, the extension of the developed results to the scenario of event-triggered remote state estimation for discrete-state Markov models observed in Gaussian noises will be investigated.

2.2 Dynamically event-triggered state estimation of hidden Markov models through a lossy communication channel[*]

In this section, we consider an estimation problem of hidden Markov models with dynamic event-triggering conditions, where the state of the event-trigger depends not only on the measurement of the sensor, but also on its own state at the previous time instant. The event-triggered state estimator is developed by first deriving results under a reference measure and then mapping them back to the "real-world" measure.

2.2.1 Introduction

In this section, a problem of event-based state estimation with dynamic event-triggering conditions for hidden Markov models is investigated. In the existing investigations, the event-triggering conditions considered were normally static and known to the remote estimator, which limits the potential of utilizing event-triggered transmission protocols in maintaining estimation performance at reduced communication cost. One feasible way of overcoming this issue is to introduce dynamics in the event-triggering condition so that an additional degree of freedom can be provided to the event-trigger in deciding whether or not to send the measurements at each time instant. The consequence, however, is that the corresponding event-triggering condition will not be exactly known to the estimator; this adds to the difficulty to solve the event-based estimation problem, and in particular, the situation will become even more complicated when

---

[*]The results in Section 2.2 appeared in *Proceedings of the IEEE 55th Conference on Decision and Control*, pp. 5122-5127, Las Vegas, USA, December 12-14, 2016.

the effect of packet dropout is considered, which is normally inevitable when the measurements are transmitted through a wired or wireless communication channel. Based on these considerations, a remote estimation problem of this type is investigated in this section for hidden Markov models, and the main contributions are summarized as follows:

1. A dynamic event-triggering transmission protocol is proposed. The state of the event-trigger depends not only on the measurement of the sensor, but also on its own state at the previous time instant. The packet dropout effect is considered and modeled by an i.i.d. Bernoulli process.

2. To solve the problem of remote estimation, a reference probability measure is constructed, under which the sensor measurement process is i.i.d. uniformly distributed, and the state of the event-trigger is also i.i.d. uniformly distributed, independent of its previous state and the sensor measurement. A map that links the reference measure to the real-world measure is proposed.

3. Under the reference measure, the unnormalized conditional distribution of the state on the event-triggered measurement information is shown to evolve recursively according to a linear map. Based on this result, the expression for the event-based state estimate under the real-world measure can be further developed.

Notation: Let $\mathbb{N}$ denote the set of nonnegative integers. Write $\mathbb{N}_{1:M} := \{1, 2, \ldots, M\}$ and $\mathbb{Z}_{-M:N} := \{-M, -M+1, \ldots, N\}$. For a set $\mathbb{M}$, let $|\mathbb{M}|$ be its cardinality. For a probability measure P (or $\hat{\mathrm{P}}$ and $\check{\mathrm{P}}$), we use E (or $\hat{\mathrm{E}}$ and $\check{\mathrm{E}}$, respectively) to represent the expectation operator. We use $[x_i]_{i\in\mathbb{N}_{1:n}}$, to denote $[x_1^{\mathrm{T}}, \ldots, x_n^{\mathrm{T}}]^{\mathrm{T}}$ and use $\mathrm{E}[x_i]_{i\in\mathbb{N}_{1:n}}$ to denote $\left[\mathrm{E}[x_1]^{\mathrm{T}}, \ldots, \mathrm{E}[x_n]^{\mathrm{T}}\right]^{\mathrm{T}}$ (the same for $\hat{\mathrm{E}}$ and $\check{\mathrm{E}}$). 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. Let $x, y \in \mathbb{R}^m$, then $\langle x, y \rangle := x^{\mathrm{T}}y$ denotes the inner product between x and y.

2.2.2   Problem destription

Firstly, we introduce a hidden Markov model on the real-world probability measure $(\Omega, \mathcal{F}, \mathrm{P})$. 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 $S_X =$

$\{e_1, e_2, \ldots, e_N\}$, where e_i is the unit vector in $\mathbb{R}^N$ with the ith element equal to 1. Let $\mathcal{F}_k^0$ denote the σ-field on Ω generated by X_0, X_1, ..., X_k, i.e., $\mathcal{F}_k^0 := \sigma\{X_0, \cdots, X_k\}$, and let $\mathcal{F}_k^X$ be the complete filtration generated by $\mathcal{F}_k^0$. Let $A := [a_i]_{i \in \mathbb{N}_{1:N}}$, $a_i := [a_{i,j}]_{j \in \mathbb{N}_{1:N}}^\top$, where $a_{i,j} := \mathrm{P}(X_{k+1} = e_i | X_k = e_j)$, such that $\sum_{i=1}^N a_{i,j} = 1$. Then

$$\mathrm{E}(X_{k+1}|\mathcal{F}_k^X) = \mathrm{E}(X_{k+1}|X_k) = AX_k. \tag{2.26}$$

Let Y_k be a sensor measurement process of X_k, which takes values in a finite-state space S_Y. Let the cardinality of the state space S_Y of Y be M, then S_Y can be identified with $S_Y = \{f_1, f_2, \ldots, f_M\}$, where f_i is the unit vector in $\mathbb{R}^M$ with the ith element equal to 1. Write $C := [c_i]_{i \in \mathbb{N}_{1:M}}$, $c_i := [c_{i,j}]_{j \in \mathbb{N}_{1:N}}^\top$, where $c_{i,j} := \mathrm{P}(Y_k = f_i | X_k = e_j)$, so that $\sum_{i=1}^M c_{i,j} = 1$ and $c_{i,j} \geq 0$. Therefore

$$\mathrm{E}(Y_k|X_k) = CX_k.$$

Let $\mathcal{F}_k^Y$ be the completion of the σ-field on Ω generated by Y_0, Y_1, ... ,Y_k. As Y_k measures the state of the hidden process X_k, the distribution of Y_k, as long as X_k is given, does not depend on $Y_{1:k-1}$. Thus the measurement model can be fully described by

$$\mathrm{E}(Y_k|\mathcal{F}_k^X \cup \mathcal{F}_{k-1}^Y) = \mathrm{E}(Y_k|X_k) = CX_k. \tag{2.27}$$

In this work, we investigate remote state estimation of finite-state hidden Markov model based on event-triggered measurements, where an event-trigger γ_k is used to decide whether the measurement is sent or not at each time instant. Obviously the cardinality of γ_k is 2, so the state space S_γ can be identified with

$$S_\gamma = \{\alpha_1, \alpha_2\},$$

where α_i is the unit vector in $\mathbb{R}^2$ with the ith element equal to 1. Specifically, if $\gamma_k = \alpha_1$, an event is triggered and the measurement is sent to the remote estimator through a wired/wireless channel; if $\gamma_k = \alpha_2$, the event is not triggered and no measurement is transmitted. In this work, we consider a dynamic event-triggering condition, which depends not only on the measurement at each time instant, but also on the previous triggering state γ_{k-1}. The intuition of this event-triggering protocol is to allow the decision of whether or not to transmit the measurement data at each time instant to be made based on whether the data is transmitted at the previous instant, since the sensor readings may not be significantly deviated for two consecutive time instants. Mathematically, this scheduling strategy is described by

$$\mathrm{E}(\gamma_k|\mathcal{F}_k^X \cup \mathcal{F}_k^Y \cup \mathcal{F}_{k-1}^\gamma) = D(Y_k)\gamma_{k-1},$$

where $\mathcal{F}_{k-1}^\gamma$ is the completion of the σ-field on Ω generated by γ_0, γ_1, $\ldots$,γ_{k-1}. There-fore, the event-trigger γ_k is actually a two-state Markov chain with a dynamic trans-formation matrix $D(Y_k)$ parameterized as

$$D(Y_k) = \sum_{i=1}^M D^i \langle Y_k, f_i \rangle,$$

to reflect the effect of the measurement Y_k, where

$$D^i := [d_u^i]_{u \in \mathbb{N}_{1:2}}, d_u^i = [d_{u,v}^i]_{v \in \mathbb{N}_{1:2}}^\top,$$

with

$$d_{u,v}^i := \mathrm{P}(\gamma_k = \alpha_u | Y_k = f_i, \gamma_{k-1} = \alpha_v),$$

so that $\sum_{u=1}^2 d_{u,v}^i = 1$ and $d_{u,v}^i \geq 0$. Therefore, γ_k can be fully described by

$$\mathrm{E}(\gamma_k | \mathcal{F}_k^X \cup \mathcal{F}_k^Y \cup \mathcal{F}_{k-1}^\gamma) = \sum_{i=1}^M D^i \langle Y_k, f_i \rangle \gamma_{k-1}. \tag{2.28}$$

On the other hand, since packet dropout is normally inevitable for a practical commu-nication channel, it is also considered in this work. Following the standard procedure of modeling packet dropout, we model the packet dropout process ζ_k as an i.i.d. Bernoulli process under P satisfying

$$\mathrm{P}(\zeta_k = 1) = \mathrm{P}(\zeta_k = 1 | \mathcal{F}_k^X \cup \mathcal{F}_k^Y \cup \mathcal{F}_k^\gamma \cup \mathcal{F}_{k-1}^\zeta) = \lambda,$$
$$\mathrm{P}(\zeta_k = 0) = \mathrm{P}(\zeta_k = 0 | \mathcal{F}_k^X \cup \mathcal{F}_k^Y \cup \mathcal{F}_k^\gamma \cup \mathcal{F}_{k-1}^\zeta) = 1 - \lambda,$$

where $\mathcal{F}_{k-1}^\zeta$ is the completion of the σ-field on Ω generated by ζ_0, ζ_1, $\ldots$,ζ_{k-1}. If $\zeta_k = 1$, there is no packet loss at k; if $\zeta_k = 0$, there is a packet loss at k. The model above indicates that ζ is not related to X, Y and γ. Therefore,

$$\mathrm{E}(\zeta_k) = \mathrm{E}(\zeta_k | \mathcal{F}_k^X \cup \mathcal{F}_k^Y \cup \mathcal{F}_k^\gamma \cup \mathcal{F}_{k-1}^\zeta) = \lambda. \tag{2.29}$$

The main goal of this work is to estimate the state X_k based on the event-triggered measurement information up to time k available to the remote estimator. If the event is triggered and there is no packet loss on the communication channel, measurement Y_k is received at the remote estimator; however, if no measurement is received, the cause can be either that the no event is triggered or that the measurement is sent by the sensor but dropped by the communication channel. Due to the packet dropout effect, the event-triggering condition is no longer known to the remote estimator, which makes the estimation problem even more complicated.

2.2.3 Map from the reference measure to the "real-world" measure

In this work, the reference probability approach is used to solve the considered remote estimation problem. To simplify the structure of the event-triggered measurement information, a reference measure is proposed in this subsection. The idea is to derive the unnormalized conditional probability distribution of the state under the reference measure, then map it back to the "real-world" measure.

Consider a reference measure $\bar{\mathrm{P}}$, under which the following relationships hold for X_k, Y_k, γ_k and ζ_k under the reference measure $\bar{\mathrm{P}}$,

$$\bar{\mathrm{E}}(X_{k+1}|\mathcal{F}_k^X) = \bar{\mathrm{E}}(X_{k+1}|X_k) = AX_k, \tag{2.30}$$

$$\bar{\mathrm{P}}(Y_k = f_i|\mathcal{F}_k^X \cup \mathcal{F}_{k-1}^Y) = \bar{\mathrm{P}}(Y_k = f_i) = 1/M, \tag{2.31}$$

$$\bar{\mathrm{E}}(\gamma_k = \alpha_j|\mathcal{F}_k^X \cup \mathcal{F}_k^Y \cup \mathcal{F}_{k-1}^\gamma) = \bar{\mathrm{P}}(\gamma_k = \alpha_j) = 1/2, \tag{2.32}$$

$$\bar{\mathrm{E}}(\zeta_k|\mathcal{F}_k^X \cup \mathcal{F}_k^Y \cup \mathcal{F}_k^\gamma \cup \mathcal{F}_{k-1}^\zeta) = \bar{\mathrm{E}}(\zeta_k) = \lambda. \tag{2.33}$$

Notice the probability distributions of X_k and ζ_k are reserved, while Y_k and γ_k become independent uniform distributed random variables under $\bar{\mathrm{P}}$. The relationships above are important and will be exploited during the derivation of the conditional probability distribution of the state under the reference measure $\bar{\mathrm{P}}$. Then we can map the results derived under $\bar{\mathrm{P}}$ back to the "real-world" space P, which solves the original estimation problem. Based on this idea, we propose the following Radon-Nikodym derivative defined over $\bar{\mathcal{G}}_k := \mathcal{F}_k^X \cup \mathcal{F}_k^Y \cup \mathcal{F}_k^\gamma \cup \mathcal{F}_k^\zeta$:

$$\left.\frac{\mathrm{dP}}{\mathrm{d\bar{P}}}\right|_{\bar{\mathcal{G}}_k} = \bar{\Lambda}_k, \tag{2.34}$$

with

$$\bar{\Lambda}_k := \prod_{l=1}^k \bar{\lambda}_l, \quad \bar{\lambda}_k := \bar{\lambda}_{k,1} \cdot \bar{\lambda}_{k,2},$$

$$\bar{\lambda}_{k,1} := M \sum_{j_1=1}^M \langle CX_k, f_{j_1}\rangle\langle Y_k, f_{j_1}\rangle,$$

$$\bar{\lambda}_{k,2} := 2 \sum_{j_2=1}^2 \left\langle \sum_{i=1}^M D^i \langle Y_k, f_i\rangle\gamma_{k-1}, \alpha_{j_2}\right\rangle\langle\gamma_k, \alpha_{j_2}\rangle.$$

For the Radon-Nikodym derivative defined above, we have the following result.

Theorem 2.3. *If the model in (2.30)-(2.33) under the reference measure $\bar{\mathrm{P}}$ is mapped to P via the Radon-Nikodym derivative in (2.34), then the obtained model under P satisfies (2.26)-(2.29).*

Proof. The detailed proof is omitted due to length limitation. The proof sketch is given in the following. Firstly, we need the value of $\bar{\mathrm{E}}\big[\bar{\lambda}_{k+1}\big|\bar{\mathcal{G}}_k\big]$. According to the definition of $\bar{\lambda}_{k+1}$ and equation (2.31), based on some calculations, we have

$$\bar{\mathrm{E}}\big[\bar{\lambda}_{k+1}\big|\bar{\mathcal{G}}_k\big] = 1. \tag{2.35}$$

Then we move on to prove the main result of Theorem 2.3. Let's start with X_k. According to the definition of X_{k+1}, Theorem 3.2 in Chapter 2 of [17], and $\bar{\Lambda}_k$ is measurable on $\bar{\mathcal{G}}_k$, we obtain

$$\mathrm{E}\big[X_{k+1}|\bar{\mathcal{G}}_k\big] = \frac{\bar{\mathrm{E}}\big[X_{k+1}\bar{\Lambda}_{k+1}\big|\bar{\mathcal{G}}_k\big]}{\bar{\mathrm{E}}\big[\bar{\Lambda}_{k+1}\big|\bar{\mathcal{G}}_k\big]} = \bar{\mathrm{E}}\big[X_{k+1}\bar{\lambda}_{k+1}\big|\bar{\mathcal{G}}_k\big].$$

Using $\bar{\mathrm{P}}(Y_{k+1} = f_i) = 1/M$ and equation (2.30), based on some further calculations, we have

$$\mathrm{E}\big[X_{k+1}|\bar{\mathcal{G}}_k\big] = \bar{\mathrm{E}}\big[X_{k+1}\bar{\lambda}_{k+1}\big|\bar{\mathcal{G}}_k\big] = \bar{\mathrm{E}}\big[X_{k+1}\big|\bar{\mathcal{G}}_k\big] = AX_k.$$

By repeated conditioning and $\mathcal{F}_k^X \subset \bar{\mathcal{G}}_k$, we obtain

$$\mathrm{E}\big[X_{k+1}\big|\mathcal{F}_k^X\big] = \mathrm{E}\big[\mathrm{E}\big[X_{k+1}|\bar{\mathcal{G}}_k\big]\big|\mathcal{F}_k^X\big] = AX_k.$$

Following a similar procedure as in evaluating $\mathrm{E}\big[X_{k+1}\big|\mathcal{F}_k^X\big]$, we have the following relationships for Y_k, γ_k and ζ_k under the "real-world" measure P,

$$\mathrm{E}\big[Y_k|\mathcal{F}_k^X \cup \mathcal{F}_{k-1}^Y\big] = CX_{k+1},$$
$$\mathrm{E}\big[\gamma_k|\mathcal{F}_k^X \cup \mathcal{F}_k^Y \cup \mathcal{F}_{k-1}^\gamma\big] = \sum_{i=1}^M D^i\langle Y_k, f_i\rangle\gamma_{k-1},$$
$$\mathrm{E}\big[\zeta_k|\mathcal{F}_k^X \cup \mathcal{F}_k^Y \cup \mathcal{F}_k^\gamma \cup \mathcal{F}_{k-1}^\zeta\big] = \lambda.$$

The proof is then completed. $\square$

To describe the information received at the remote estimator, we introduce an instrumental measurement process $\bar{Y}_k$ taking values in $\{g_1, g_2, \ldots, g_N\}$, where g_i is the unit vector in $\mathbb{R}^{M+1}$ with the ith element equal to 1. Define

$$\bar{Y}_k = \begin{cases} g_i, & \text{if } \{Y_k = f_i\} \cap \{\gamma_k = \alpha_1\} \cap \{\zeta_k = 1\}, \\ g_{M+1}, & \text{if } \{\gamma_k = \alpha_2\} \cup \{\{\gamma_k = \alpha_1\} \cap \{\zeta_k = 0\}\}, \end{cases}$$

where $Y_k = g_i$ implies that the measurement is received by the remote estimator, and $Y_k = g_{M+1}$ implies that the measurement is not received. Let $\mathcal{F}_k^{\bar{Y}_k}$ be the completion of the σ-field on Ω generated by $\bar{Y}_0, \bar{Y}_1, \ldots, \bar{Y}_k$. Notice $\bar{Y}_k$ is measurable under $\bar{\mathcal{G}}_k$. Based on

the above theorem, the problem of estimating state X_k given $\mathcal{F}_k^{\bar{Y}}$ under P can be solved by considering a state estimation problem under $\bar{P}$ and mapping the results back to P. In the next subsection, we derive the unnormalized conditional probability distribution of the state under $\bar{P}$, based on which the conditional probability distribution under P can be obtained.

2.2.4 Recursive estimate of the state

Write $\bar{q}_k = [\bar{q}_k^1, \ \ldots, \ \bar{q}_k^N]^{\mathrm{T}}$ and $\bar{p}_k = [\bar{p}_k^1, \ \ldots, \ \bar{p}_k^N]^{\mathrm{T}}$, where

$$\bar{q}_k^r = \bar{\mathrm{E}}\big[\bar{\Lambda}_k \langle X_k, e_r \rangle \big| \mathcal{F}_k^{\bar{Y}}\big], \tag{2.36}$$

$$\bar{p}_k^r = \mathrm{E}\big[\langle X_k, e_r \rangle \big| \mathcal{F}_k^{\bar{Y}}\big], \tag{2.37}$$

for $r \in \mathbb{N}_{1:N}$. In this way, we have

$$\bar{p}_k^r = \bar{q}_k^r / \|\bar{q}_k\|_1. \tag{2.38}$$

Note that $\bar{p}_k$ denotes the conditional distribution of the state on the event-triggered measurement information under P. In the following, we show that the unnormalized conditional probability distribution $\bar{q}_k$ evolves in a linear recursive form, which is the main result of this work.

Theorem 2.4. *For the event-triggered system described by (2.26)-(2.29), the unnormalized probability distribution $\bar{q}_k$ of the state conditioned on the information available at the remote estimator under $\bar{P}$ satisfies the following linear recursive form:*

$$\bar{q}_{k+1} = \mathrm{diag}(\bar{o}_{k+1}) A \bar{q}_k, \tag{2.39}$$

where

$$\bar{o}_{k+1} := [\bar{o}_{k+1}^1 \ \ldots \ \bar{o}_{k+1}^N]^{\mathrm{T}} \in \mathbb{R}^N, \tag{2.40}$$

$$\bar{o}_{k+1}^r := 2M \sum_{j_1=1}^M c_{j_1, r}$$
$$\cdot \left\{ \left[\frac{\left(d_{1,1}^{j_1}(1-\lambda) + d_{2,1}^{j_1}\right)}{2M - M\lambda} \bar{Y}_{k+1}^{M+1} + d_{1,1}^{j_1} \bar{Y}_{k+1}^{j_1} \right] \left[\frac{1-\lambda}{2-\lambda} \bar{Y}_k^{M+1} + \sum_{i_2=1}^M \bar{Y}_k^{i_2} \right] \right.$$
$$\left. + \left[\frac{\left(d_{1,2}^{j_1}(1-\lambda) + d_{2,2}^{j_1}\right)}{2M - M\lambda} \bar{Y}_{k+1}^{M+1} + d_{1,2}^{j_1} \bar{Y}_{k+1}^{j_1} \right] \frac{1}{2-\lambda} \bar{Y}_k^{M+1} \right\}. \tag{2.41}$$

Proof. According to the definition of $\bar{q}_{k+1}^r$, we have

$$
\begin{aligned}
&\bar{q}_{k+1}^r \\
&= \bar{E}\left[\bar{\Lambda}_{k+1}\langle X_{k+1}, e_r\rangle\middle|\mathcal{F}_{k+1}^{\bar{Y}}\right] \\
&= \bar{E}\left[\bar{\Lambda}_k \bar{\lambda}_{k+1}\langle X_{k+1}, e_r\rangle\middle|\mathcal{F}_{k+1}^{\bar{Y}}\right] \\
&= \bar{E}\left[\bar{\Lambda}_k \langle X_{k+1}, e_r\rangle M \sum_{j_1=1}^M \langle CX_{k+1}, f_{j_1}\rangle\langle Y_{k+1}, f_{j_1}\rangle\right. \\
&\quad \left. \cdot 2\sum_{j_2=1}^2 \left\langle \sum_{i=1}^M D^i\langle Y_{k+1}, f_i\rangle\gamma_k, \alpha_{j_2}\right\rangle\langle\gamma_{k+1}, \alpha_{j_2}\rangle\middle|\mathcal{F}_{k+1}^{\bar{Y}}\right].
\end{aligned}
$$

Recalling that $\langle CX_{k+1}, f_{j_1}\rangle = \sum_{j_3=1}^N c_{j_1,j_3}\langle X_{k+1}, e_{j_3}\rangle$, we have

$$
\begin{aligned}
\bar{q}_{k+1}^r = {}&2M\bar{E}\left[\bar{\Lambda}_k \langle X_{k+1}, e_r\rangle \sum_{j_1=1}^M \sum_{j_3=1}^N c_{j_1,j_3}\langle X_{k+1}, e_{j_3}\rangle\right. \\
&\left. \cdot \langle Y_{k+1}, f_{j_1}\rangle \sum_{j_2=1}^2 \left\langle \sum_{i=1}^M D^i\langle Y_{k+1}, f_i\rangle\gamma_k, \alpha_{j_2}\right\rangle\langle\gamma_{k+1}, \alpha_{j_2}\rangle\middle|\mathcal{F}_{k+1}^{\bar{Y}}\right], \quad (2.42)
\end{aligned}
$$

where the term in (2.42) equals to 0 if $e_r \neq e_{j_3}$ or $f_{j_1} \neq f_i$, then we obtain

$$
\begin{aligned}
\bar{q}_{k+1}^r = {}&2M\bar{E}\left[\bar{\Lambda}_k \langle X_{k+1}, e_r\rangle \sum_{j_1=1}^M c_{j_1,r}\langle Y_{k+1}, f_{j_1}\rangle \sum_{j_2=1}^2 \left\langle D^{j_1}\gamma_k, \alpha_{j_2}\right\rangle\langle\gamma_{k+1}, \alpha_{j_2}\rangle\middle|\mathcal{F}_{k+1}^{\bar{Y}}\right] \\
= {}&2M\bar{E}\left[\bar{\Lambda}_k \langle X_{k+1}, e_r\rangle \sum_{j_1=1}^M c_{j_1,r}\langle Y_{k+1}, f_{j_1}\rangle\right. \\
&\left. \cdot \sum_{i_0=1}^2\langle\gamma_k, \alpha_{i_0}\rangle \sum_{j_2=1}^2 d_{j_2,i_0}^{j_1}\langle\gamma_{k+1}, \alpha_{j_2}\rangle\middle|\mathcal{F}_{k+1}^{\bar{Y}}\right].
\end{aligned}
$$

By repeated conditioning and $\mathcal{F}_{k+1}^{\bar{Y}} \subset \{\mathcal{F}_k^X \cup \mathcal{F}_{k+1}^Y \cup \mathcal{F}_{k+1}^\gamma \cup \mathcal{F}_{k+1}^\zeta\}$, combining

$$
\begin{aligned}
\bar{E}(X_{k+1}|\mathcal{F}_k^X) &= AX_k, \\
\langle AX_k, e_r\rangle &= \sum_{v=1}^N a_{rv}\langle X_k, e_v\rangle, \quad (2.43)
\end{aligned}
$$

$\bar{q}_{k+1}^r$ becomes

$$
\begin{aligned}
&2M\bar{E}\left[\sum_{v=1}^N a_{rv}\langle X_k, e_v\rangle\bar{\Lambda}_k \sum_{j_1=1}^M c_{j_1,r} \sum_{i_0=1}^2 \sum_{j_2=1}^2\right. \\
&\left. \cdot d_{j_2,i_0}^{j_1}\langle\gamma_k, \alpha_{i_0}\rangle\langle\gamma_{k+1}, \alpha_{j_2}\rangle\langle Y_{k+1}, f_{j_1}\rangle\middle|\mathcal{F}_{k+1}^{\bar{Y}}\right], \\
&= 2M\sum_{v=1}^N a_{rv}\bar{q}_k^v \sum_{j_1=1}^M c_{j_1,r} \sum_{i_0=1}^2 \sum_{j_2=1}^2 d_{j_2,i_0}^{j_1} \\
&\quad \cdot \bar{E}\left[\langle\gamma_{k+1}, \alpha_{j_2}\rangle\langle Y_{k+1}, f_{j_1}\rangle\middle|\bar{Y}_{k+1}\right]\bar{E}\left[\langle\gamma_k, \alpha_{i_0}\rangle\middle|\bar{Y}_k\right], \quad (2.44)
\end{aligned}
$$

where equation (2.44) is due to that γ_k is only related with $\bar{Y}_k$; γ_{k+1} and Y_{k+1} are only related with $\bar{Y}_{k+1}$. Now the key issue of calculating $\bar{q}_{k+1}^r$ is to obtain

$$
\sum_{i_0=1}^2 \sum_{j_2=1}^2 d_{j_2,i_0}^{j_1}\bar{E}\left[\langle\gamma_{k+1}, \alpha_{j_2}\rangle\langle Y_{k+1}, f_{j_1}\rangle\middle|\bar{Y}_{k+1}\right]\bar{E}\left[\langle\gamma_k, \alpha_{i_0}\rangle\middle|\bar{Y}_k\right]. \quad (2.45)
$$

For the first step, we focus on

$$
\begin{aligned}
&\sum_{j_2=1}^2 d_{j_2,i_0}^{j_1}\bar{E}\left[\langle\gamma_{k+1}, \alpha_{j_2}\rangle\langle Y_{k+1}, f_{j_1}\rangle\middle|\bar{Y}_{k+1}\right] \\
&= \sum_{j_2=1}^2 d_{j_2,i_0}^{j_1}\bar{P}\left[Y_{k+1} = f_{j_1} \cap \gamma_{k+1} = \alpha_{j_2}\middle|\bar{Y}_{k+1}\right]. \quad (2.46)
\end{aligned}
$$

36

1) For $\bar{Y}_{k+1} = g_{M+1}$, we have

$$\sum_{j_2=1}^2 d_{j_2,i_0}^{j_1} \bar{P}\big[Y_{k+1} = f_{j_1} \cap \gamma_{k+1} = \alpha_{j_2}\big|\bar{Y}_{k+1}\big].$$

$$=\sum_{j_2=1}^2 d_{j_2,i_0}^{j_1} \frac{\bar{P}\big[Y_{k+1} = f_{j_1} \cap \gamma_{k+1} = \alpha_{j_2} \cap \bar{Y}_{k+1} = g_{M+1}\big]}{\bar{P}\big[\bar{Y}_{k+1} = g_{M+1}\big]}. \qquad (2.47)$$

According to Bayes' rule, the numerator in (2.47) becomes

$$\sum_{j_2=1}^2 d_{j_2,i_0}^{j_1} \bar{P}\big[Y_{k+1} = f_{j_1} \cap \gamma_{k+1} = \alpha_{j_2} \cap \bar{Y}_{k+1} = g_{M+1}\big] = \frac{\big[d_{1,i_0}^{j_1}(1-\lambda) + d_{2,i_0}^{j_1}\big]}{2M}. \qquad (2.48)$$

According to (2.33), γ_{k+1} is independent of ζ_{k+1}, and thus the denominator in (2.47) satisfies

$$\bar{P}\big[\bar{Y}_{k+1} = g_{M+1}\big]$$
$$=\bar{P}\big[\{\gamma_k = \alpha_2\} \cup \{\{\gamma_k = \alpha_1\} \cap \{\zeta_k = 0\}\}\big]$$
$$=1 - \bar{P}\big[\gamma_k = \alpha_1 \cap \zeta_k = 0\big] = 1 - \frac{\lambda}{2}. \qquad (2.49)$$

So we have

$$\sum_{j_2=1}^2 d_{j_2,i_0}^{j_1} \bar{P}\big[Y_{k+1} = f_{j_1} \cap \gamma_{k+1} = \alpha_{j_2}\big|\bar{Y}_{k+1}\big] = \frac{\big[d_{1,i_0}^{j_1}(1-\lambda) + d_{2,i_0}^{j_1}\big]}{2M - M\lambda}.$$

2) For $\bar{Y}_{k+1} = g_{j_1}$, we have

$$\sum_{j_2=1}^2 d_{j_2,i_0}^{j_1} \bar{E}\big[\langle \gamma_{k+1}, \alpha_{j_2}\rangle\langle Y_{k+1}, f_{j_1}\rangle\big|\bar{Y}_{k+1} = g_{j_1}\big] = d_{1,i_0}^{j_1}.$$

3) For $\bar{Y}_{k+1} = g_v$, ($v \neq j_1$ and $v \neq M+1$),

$$\sum_{j_2=1}^2 d_{j_2,i_0}^{j_1} \bar{E}\big[\langle \gamma_{k+1}, \alpha_{j_2}\rangle\langle Y_{k+1}, f_{j_1}\rangle\big|\bar{Y}_{k+1} = g_v\big] = 0.$$

Combining the above results, we have

$$\sum_{j_2=1}^2 d_{j_2,i_0}^{j_1} \bar{P}\big[Y_{k+1} = f_{j_1} \cap \gamma_{k+1} = \alpha_{j_2}\big|\bar{Y}_{k+1}\big].$$

$$=\frac{\big[d_{1,i_0}^{j_1}(1-\lambda) + d_{2,i_0}^{j_1}\big]}{2M - M\lambda} \bar{Y}_{k+1}^{M+1} + d_{1,i_0}^{j_1} \bar{Y}_{k+1}^{j_1}, \qquad (2.50)$$

where $\bar{Y}_{k+1}^{M+1} = \langle \bar{Y}_{k+1}, g_{M+1}\rangle$ and $\bar{Y}_{k+1}^{j_1} = \langle \bar{Y}_{k+1}, g_{j_1}\rangle$.

Next, we move on to calculate $\bar{E}\big[\langle \gamma_k, \alpha_{i_0}\rangle\big|\bar{Y}_k\big]$. Since the procedure is the same as above, the results are directly provided:

$$\bar{E}\big[\langle \gamma_k, \alpha_1\rangle\big|\bar{Y}_k\big] = \frac{1-\lambda}{2-\lambda}\bar{Y}_k^{M+1} + \sum_{i_2=1}^M \bar{Y}_k^{i_2}, \qquad (2.51)$$

$$\bar{E}\big[\langle \gamma_k, \alpha_2\rangle\big|\bar{Y}_k\big] = \frac{1}{2-\lambda}\bar{Y}_k^{M+1}. \qquad (2.52)$$

Now, combining (2.50), (2.51) and (2.52), equation (2.45) satisfies

$$\sum_{i_0=1}^{2} \sum_{j_2=1}^{2} d_{j_2,i_0}^{j_1} \bar{E}\big[\langle \gamma_{k+1}, \alpha_{j_2}\rangle \langle Y_{k+1}, f_{j_1}\rangle \big| \bar{Y}_{k+1}\big] \bar{E}\big[\langle \gamma_k, \alpha_{i_0}\rangle \big| \bar{Y}_k\big].$$

$$= \left[\frac{(d_{1,1}^{j_1}(1-\lambda) + d_{2,1}^{j_1})}{2M - M\lambda} \bar{Y}_{k+1}^{M+1} + d_{1,1}^{j_1} \bar{Y}_{k+1}^{j_1} \right] \left[\frac{1-\lambda}{2-\lambda} \bar{Y}_k^{M+1} + \sum_{i_2=1}^{M} \bar{Y}_k^{i_2} \right]$$

$$+ \left[\frac{(d_{1,2}^{j_1}(1-\lambda) + d_{2,2}^{j_1})}{2M - M\lambda} \bar{Y}_{k+1}^{M+1} + d_{1,2}^{j_1} \bar{Y}_{k+1}^{j_1} \right] \frac{1}{2-\lambda} \bar{Y}_k^{M+1}.$$

Finally, substituting (2.53) into (2.44) and according to the definition of $\bar{o}_{k+1}^r$, we obtain

$$\bar{q}_{k+1}^r = \bar{o}_{k+1}^r \sum_{v=1}^{N} a_{rv} \bar{q}_k^v. \tag{2.53}$$

In this way, by the definition of matrix A and $\bar{o}_{k+1}$, we have

$$\bar{q}_{k+1} = \text{diag}(\bar{o}_{k+1}) A \bar{q}_k, \tag{2.54}$$

which completes the proof. □

Remark 2.1. The theorem above indicates that when the measurement is not received by the remote estimator, two scenarios need to be considered, including the scenario that the event is not triggered and the scenario that the event-triggering condition is satisfied but there is a packet loss in the communication channel. The conditional probability distribution of the state can be easily used to obtain the MMSE estimate of the state. On the other hand, an interesting special case to consider is that the parameters in dynamic event-triggering condition are configured as $d_{u,1}^i = d_{u,2}^i = d_{u,1}^j = d_{u,2}^j$ for all i, j $\in \mathbb{N}_{1:M}$, which means that if the measurement is not received, no "useful" information about Y_k can be inferred, as every f_i in the measurement space has the same probability of being transmitted. This corresponds to the case of ignoring the information contained in the event-triggering condition, which is utilized for comparison in the numerical example.

2.2.5 Numerical example

In this subsection, a hidden Markov model with a finite number of real-valued states is considered to illustrate the effectiveness of the proposed method. Consider a scalar real-valued process x_k with $N = 41$ and state space $x^1, \ldots, x^N$, where $x^1 = -5$, $x^N = 5$, and $x^{i+1} - x_i = 0.25$. The transition matrix A is constructed as $A = [a_1, \ldots, a_N]$, where $a_i = [\phi(\tau, 0, 0.15, N)/\sum_{l\in\mathbb{Z}_{1-i:N-i}} \phi(l, 0, 0.15, N)]_{l\in\mathbb{Z}_{1-i:N-i}}$, and

$$\phi(\tau, \mu, \sigma, n) := \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{[(\tau - \mu)/n]^2}{2\sigma^2} \right],$$
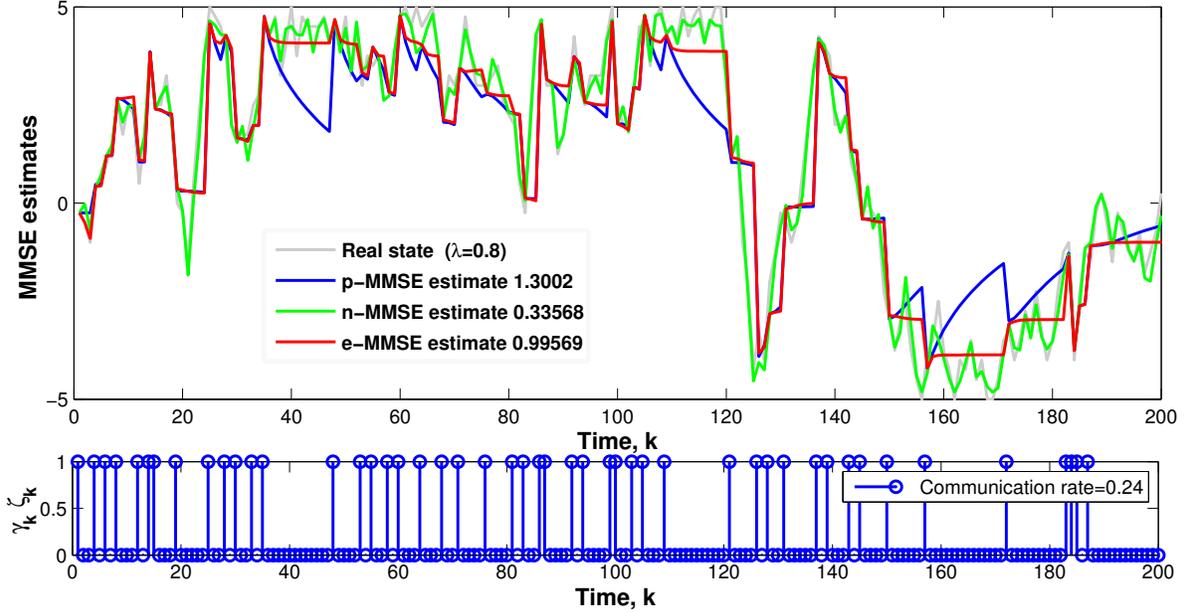
Figure 2.6: Estimation performance comparison for the system with i.i.d. packet dropout process ($\lambda = 0.8$).

so that $\|a_i\|_1 = 1$ holds. The measurement process y_k is a scalar real-valued process with $M = 41$ and state space $y^1, \ldots, y^M$, where $y^1 = -2$, $y^M = 2$, and $y^{i+1} - y_i = 0.1$. The measurement matrix C is constructed as $C = [c_1, \ldots, c_M]$, where $c_i := [\phi(\tau, 0, 0.05, M) / \sum_{l \in Z_{1-i:M-i}} \phi(l, 0, 0.05, M)]_{l \in Z_{1-i:M-i}}$. Recalling that

$$d^i_{u,v} := \mathrm{P}(\gamma_k = \alpha_u | Y_k = f_i, \gamma_{k-1} = \alpha_v),$$

the values of $d^i_{u,v}$'s at time instant k (which is written as $d^{i,k}_{u,v}$ below) depend on the previously received measurement and the state of previous event-trigger. The probability of sending $d^{i,k}_{1,v}$ is determined by the following stochastic event-triggering condition

$$d^{i,k}_{1,v} = 1 - \phi(i, \tilde{Y}, \sigma_v, M) / \phi(\tilde{Y}, \tilde{Y}, \sigma_v, M),$$

where $\tilde{Y}$ is the value of the previous received measurement at time instant k, and σ_v is the variance parameter that controls the transmission frequency. To implement the event-triggering process γ_k, we select a uniform random variable ϱ defined on $[0, 1]$, whose realization determines the value of γ as follows:

$$\gamma_k = \begin{cases} 1, & \text{if} \quad \varrho < \sum_{i=1}^M D^i \langle Y_k, f_i \rangle \gamma_{k-1}, \\ 0, & \text{if} \quad \varrho \geq \sum_{i=1}^M D^i \langle Y_k, f_i \rangle \gamma_{k-1}. \end{cases}$$

39
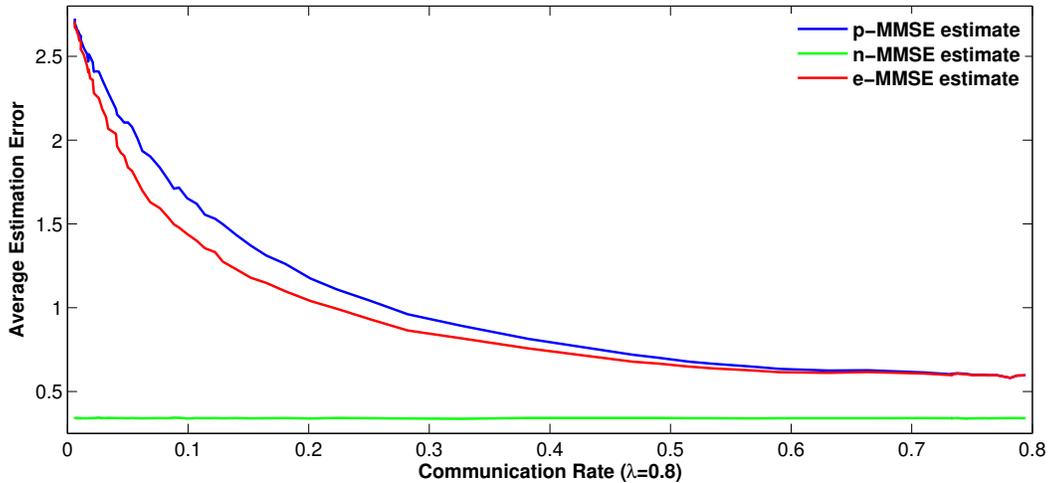
Figure 2.7: Tradeoff between the communication rate and estimation performance ($\lambda = 0.8$).

Intuitively, if $\gamma_{k-1} = \alpha_2$, we prefer to send the measurement at time instant k compared with the case when $\gamma_{k-1} = \alpha_1$, so the variance parameter σ_1 should be larger than σ_2. We consider the scenario with unreliable communication channel and utilize i.i.d. packet dropout model with packet dropout rate $1 - \lambda = 20\%$.

The MMSE estimate of X_k is obtained from the conditional probability distribution of the state obtained in Theorem 2.4. For notational brevity, the proposed event-triggered estimate is termed as "e-MMSE estimate", which is compared with the other two estimators. One is the MMSE estimate obtained by ignoring information contained in the dynamic event-triggering condition based on the parameterization of D^i in Remark 2.1, termed as "p-MMSE estimate"; the other one is the MMSE estimate obtained using all past measurements, termed as "n-MMSE estimate". For $\delta_1 = 2\delta_2 = 0.25$, the performance comparison of the n-MMSE estimate, the proposed e-MMSE estimate and the p-MMSE estimate obtained using the same sequence $\{\gamma_k \zeta_k\}$ is shown in Fig. 2.6. The average communication rate is 0.24 and the resulting estimation errors for p-MMSE, n-MMSE and e-MMSE estimates are 1.3002, 0.3357, 0.9957, respectively. The tradeoff between the average estimation error and the average communication rate obtained by keeping $\delta_1 = 2\delta_2$ and using different δ_1 is further shown in Fig. 2.7, indicating that the exploration of the set-valued information contained in the dynamic event-triggering condition helps improve the estimation performance in terms of the average communication rate. In Fig. 2.7, notice there is still a gap between the results of e-MMSE estimate and n-MMSE estimate when communication rate is almost 0.8, which is caused
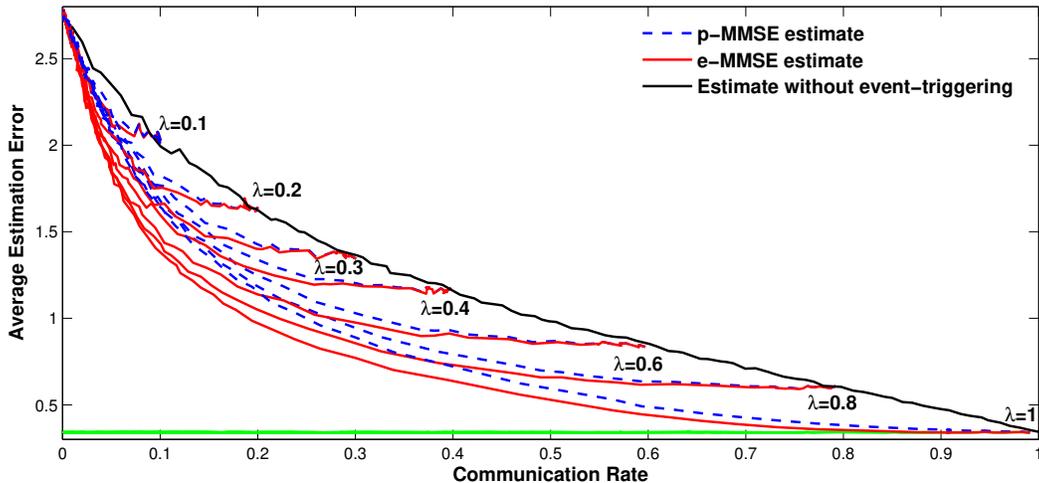
Figure 2.8: Tradeoff between the communication rate and estimation performance.

by packet loss. The tradeoff between average estimation error and average communication rate for the cases $\lambda = 0.1,\ 0.2,\ 0.3,\ 0.4,\ 0.6,\ 0.8$, and 1 are shown in Fig. 2.8, where the black line is the tradeoff between average estimation error and packet dropout rate without using event-triggering condition, namely, $\gamma_k = \alpha_1$ all the time. Fig. 2.8 shows that, when the average communication rate is close to $1 - \lambda$, p-MMSE and e-MMSE estimates become undistinguishable and converge to the result from without using event-triggering condition on the black curve.

2.2.6   Conclusion

In this work, a problem of state estimation for hidden Markov models subject to dynamically event-triggered measurements and packet dropouts has been considered and solved. Utilizing the reference probability measure approach, closed-form expressions for the unnormalized and normalized conditional distributions of the states on the event-triggered measurement information are obtained, based on which the optimal event-based state estimates can be calculated. For future work, the effect of the dynamic event-trigger and packet dropout process on the closed-loop risk-sensitive control will be investigated.

Chapter 3

Event-triggered state estimation with an energy harvesting sensor[*]

In this chapter, we investigate an event-triggered state estimation problem for a linear Gaussian system with an energy harvesting sensor, where the system state is not directly measurable but is observed through a linear measurement equation with some Gaussian measurement noise.

3.1 Introduction

For energy harvesting sensors, although the energy is inexhaustible from a long-term perspective as the sensor can absorb energy from the environment itself, energy shortage still occurs occasionally due to the randomness of the energy harvested from the environment or rapid energy consumption when the systems have significant deviations. To deal with this potential energy shortage, the sensor should adjust the frequency/probability of measurement transmission based on both the importance of the data and the amount of energy available, according to appropriately designed event-triggering conditions. This alternative transmission policy, however, in turn adds to the difficulty of solving the state estimation problem since the event-triggering condition is no longer fixed and not known to the remote estimator anymore.

The main contributions of the work in this chapter are summarized as follows. Firstly, we propose a stochastic energy-dependent event-triggered transmission protocol. The proposed event-triggering condition adds one degree of freedom by considering the energy level of the sensor and helps trading estimation performance for low com-

---

munication rates when the sensor's energy is not sufficient and vice versa, which is different from the conditions that normally only depend on the measurement information considered in the existing literature [62]. Secondly, we derive the joint conditional probability distribution of the state and energy on the set-valued and point-valued combined information at the estimator, based on which the recursive MMSE estimates of the state and energy are obtained. Thirdly, the feature of the communication rate of the proposed transmission protocol is presented by analyzing the relationship between the average communication rate and the average energy harvesting rate. The analysis indicates that these two average rates will match when the average energy harvesting rate of the sensor is large enough but smaller than 1, regardless of the choice of the other parameters, which is also verified by numerical examples.

For many event-triggered remote state estimation problems, the optimal estimates normally bear a Kalman-like structure. The results in this chapter, however, do not coincide with the existing results on state estimation problems using an energy harvesting sensor (e.g., [46]), since in this work the main results focus on how the Kalman gain evolves under the considered energy-dependent transmission protocol and provide the specific structure of the estimator by exploiting the event-triggered measurement information, which contains coupled information of the system state and the energy level. In addition, the considered sensor's decision rule in this work is different from the one obtained in [46]; due to the feature of the energy harvesting process considered in this work, analytical results of the MMSE estimate of the energy level are also provided.

Notation: $\mathbb{R}$ denotes the set of real numbers. $\mathbb{N}$ denotes the set of nonnegative 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 $n \in \mathbb{N}$, let $n! := \prod_{t=1}^{n} t$. $\mathbf{1}_{\{A\}}$ is the indicator function of a set A. Let $\delta(x)$ denote the probability distribution of a random variable x which only takes a single value $x = 0$.

3.2 Problem formulation

We consider a remote event-triggered estimation scheme in Fig. 3.1. Consider a discrete-time linear time-invariant process driven by white noise:

$$x_{k+1} = Ax_k + w_k, \tag{3.1}$$

where $x_k \in \mathbb{R}^n$ is the state, and $w_k \in \mathbb{R}^n$ is zero-mean Gaussian with covariance $Q > 0$. The initial value x_0 is Gaussian with mean $\hat{x}_0^- = \mu_0$, and covariance P_0^-. The state
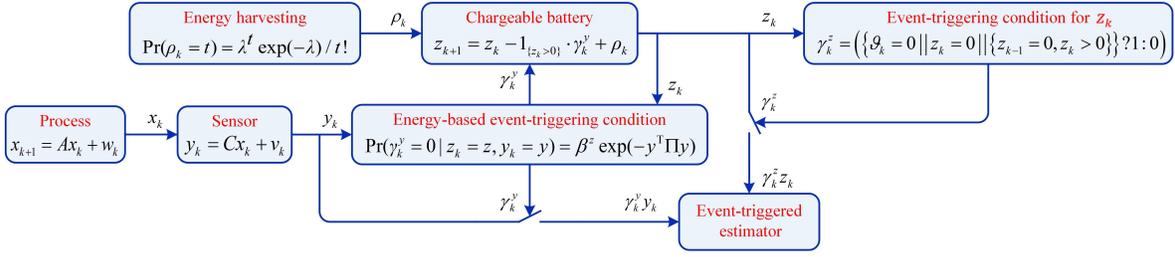
Figure 3.1: Block diagram of the event-based remote estimation system.

information is measured by an energy harvesting sensor, which communicates with the state estimator through a wireless channel, and the output equation is

$$y_k = Cx_k + v_k, \tag{3.2}$$

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. Let ρ_k denote the amount of energy harvested at time k, and we assume ρ_k is an i.i.d. Poisson process with parameter λ such that

$$\Pr(\rho_k = t) = \frac{\lambda^t \exp(-\lambda)}{t!}. \tag{3.3}$$

This assumption is based on the fact that energy harvesting modules usually contain small sub-modules harvesting energy independently, where the net energy harvested can be modeled as a binomial process, which approaches to the Poisson process in the limit when the number of sub-modules becomes large [15]; the Poisson energy harvesting model is widely used in engineering applications, e.g., cellular mobile communication networks [68], WSN [43] and traffic steering for green cellular networks [81]. Furthermore, we assume the energy harvesting sensor is equipped with a chargeable battery with initial power z_0. Denote the energy available to the sensor at time instant k as z_k. We assume z_k satisfies

$$z_{k+1} = z_k - \mathbf{1}_{\{z_k > 0\}} \cdot \gamma_k^y + \rho_k. \tag{3.4}$$

Here γ_k^y is a binary-valued event-trigger indicating whether a sensor transmission of y_k is performed at time instant k, with $\gamma_k^y = 0$ meaning no transmission and $\gamma_k^y = 1$ otherwise; We assume z_0, x_0, w_k, v_k and ρ_k are independent with each other. We assume when $z_k > 0$, the energy consumption is 1 for each measurement transmission from the sensor to the remote estimator. Although sensors equipped with energy harvesting modules allow more transmissions from the sensor to the estimator for better performance,

non-chargeable backup batteries are still necessary for many application scenarios to maintain low-rate transmission and guarantee basic estimation performance. In this regard, we assume a backup battery is equipped to the system to allow necessary transmissions when $z_k = 0$. The capacity of the backup battery can be determined from the worst-case scenario, namely when the energy harvested is always zero; in this way, the necessary capacity of the backup battery equals to the length of the battery maintenance period multiply the average communication rate under the worst case (which can be estimated based on the results for communication rate analysis obtained in [23]).

At each time instant k, the sensor produces a measurement y_k and generates a random variable ζ_k, which is uniformly distributed over $[0, 1]$. The scheduler of the sensor tests an energy-dependent stochastic event-triggering condition and the observation y_k is transmitted if and only if $\gamma_k^y = 1$, namely,

$$\gamma_k^y = \begin{cases} 0, & \text{if } \zeta_k \leq \beta^z \exp\left(-y^\top \Pi y / 2\right), \\ 1, & \text{if } \zeta_k > \beta^z \exp\left(-y^\top \Pi y / 2\right), \end{cases}$$

where $0 < \beta < 1$ and $\Pi > 0$. In other words, we have

$$\Pr(\gamma_k^y = 0 | z_k = z, y_k = y) = \beta^z \exp\left(-y^\top \Pi y / 2\right). \tag{3.5}$$

In equation (3.5), β is a control parameter which decides how the energy level influences the event-triggering condition, e.g., if β is very close to 1, β^z does not vary too much when z takes different levels, indicating the influence of energy level on the event-triggering condition is relatively weaker compared with the case when β is smaller. As $\beta < 1$, when the energy storage of the sensor z_k is large, the scheduler tends to send the measurement more often and sacrifice communication rate for estimation performance as the energy issue is not serious in this scenario. To provide the remote estimator some partial information of the battery status, we assume the value of z_k is transmitted to the estimator only if $\gamma_k^z = 1$ and not transmitted if $\gamma_k^z = 0$, where γ_k^z is a transmission decision variable. Define a sequence $\{\vartheta_k\}$ for $k \geq 0$ such that

$$\vartheta_k = \gamma_k^z z_k + (1 - \gamma_k^z)(\vartheta_{k-1} - \gamma_{k-1}^y), \tag{3.6}$$

where ϑ_k denotes the lower bound of the possible energy level at k. We consider the following transmission policy for z_k:

$$\gamma_k^z = \begin{cases} 1, & \text{if } \{\vartheta_{k-1} - \gamma_{k-1}^y = 0\} \cup \{z_k = 0\} \cup \{z_{k-1} = 0, z_k > 0\}; \\ 0, & \text{otherwise.} \end{cases} \tag{3.7}$$

At each time instant k, the sensor first updates γ_k^z according to equation (3.7), then obtains ϑ_k according to equation (3.6). This policy acts like an alarm protocol, indicating that z_k is transmitted to the estimator when its value or its lower bound estimate $\vartheta_{k-1} - \gamma_{k-1}^y$ becomes 0, or when some energy is harvested to reactivate the sensor from an out-of-power status; when z_k is sent to the estimator, the value of ϑ_k is updated as well. Since z_k is scalar- and integer-valued (as opposed to the vector- and real-valued sensor measurements y_k) and the event $\gamma_k^z = 1$ is less likely to happen than $\gamma_k^y = 1$ in many scenarios (e.g. Fig. 3.2), the total energy consumption caused by energy-level transmission is much smaller than that used for measurement transmission, and thus we assume the energy used in transmitting z_k is negligible to simplify the analysis in this work. In addition, the case that the energy consumption for z_k is non-negligible can be considered along a similar line of arguments for estimator design, as the consideration of energy consumption for transmitting z_k does not further complicate the information pattern provided by the event-triggering conditions. To simplify the derivations and focus on the main event-based estimation problem, we consider the simpler case that the energy consumption of transmitting z_k can be ignored in this work. We assume $\gamma_0^z = 1$. Let $\mathcal{I}_0 = \{z_0, \gamma_0^y, \gamma_0^y y_0\}$ and for $k \in \mathbb{N}$, $\mathcal{I}_{k+1} = \mathcal{I}_k \cup \{\gamma_{k+1}^y, \gamma_{k+1}^z, \gamma_{k+1}^y y_{k+1}, \gamma_{k+1}^z z_{k+1}\}$, which denotes the information available at the remote estimator. The goal is to estimate x_k and z_k based on $\mathcal{I}_k$. Specifically, the following two problems will be investigated in this work:

1. What form will the MMSE event-based estimate for x_k and z_k have, given the event-triggered measurement information under the energy-dependent event-triggered transmission policy?

2. How will the energy harvesting process affect the average measurement transmission rate between the sensor and the estimator?

3.3 Main results

In this section, the event-based estimation problem is solved for the proposed energy-dependent event-triggering condition. To do this, the joint probability distribution of the state and energy level is derived, which is summarized in the following theorem. Before continuing, we introduce three γ-dependent sequences $\{L_k\}$, $\{\alpha_{i,k}, i \in \mathbb{N}_{1:L_k}\}$

46

and $\{\eta_{i,k}, i \in \mathbb{N}_{1:L_k}\}$ for $k \geq 0$ such that for $\gamma_k^z = 0$,

$$L_{k+1} = L_k, \tag{3.8}$$

$$\alpha_{i,k+1} = \beta \cdot (\alpha_{i,k} + \lambda), \tag{3.9}$$

$$\eta_{i,k+1} = \frac{\eta_{i,k} \exp\left[-(1-\beta)(\alpha_{i,k}+\lambda)\right]}{\sum_{j=1}^{L_k} \eta_{j,k} \exp\left[-(1-\beta)(\alpha_{j,k}+\lambda)\right]} \tag{3.10}$$

for $\gamma_{k+1}^y = 0$, and

$$L_{k+1} = 2L_k, \tag{3.11}$$

$$\alpha_{i,k+1} = \begin{cases} \alpha_{i,k} + \lambda, & \text{if } 1 \leq i \leq L_k; \\ \beta(\alpha_{i-L_k,k} + \lambda), & \text{if } L_k < i \leq L_{k+1}, \end{cases} \tag{3.12}$$

$$\eta_{i,k+1} = \begin{cases} \dfrac{\eta_{i,k}}{1-\sum_{j=1}^{L_k} \eta_{j,k}\beta^\vartheta \exp\left[-\frac{1}{2}y^\top \Pi y - (1-\beta)(\alpha_{j,k}+\lambda)\right]}, & \text{if } 1 \leq i \leq L_k; \\ \dfrac{-\eta_{i-L_k,k}\beta^{\vartheta k+1} \exp\left[-\frac{1}{2}y^\top \Pi y - (1-\beta)(\alpha_{i-L_k,k}+\lambda)\right]}{1-\sum_{j=1}^{L_k} \eta_{j,k}\beta^{\vartheta k+1} \exp\left[-\frac{1}{2}y^\top \Pi y - (1-\beta)(\alpha_{j,k}+\lambda)\right]}, & \text{if } L_k < i \leq L_{k+1} \end{cases} \tag{3.13}$$

for $\gamma_{k+1}^y = 1$, where y denotes the measurement received at $k+1$ for simplicity; finally, for $\gamma_k^z = 1$,

$$L_{k+1} = 1, \ \alpha_{1,k+1} = \beta\lambda, \ \eta_{1,k+1} = 1, \tag{3.14}$$

for $\gamma_{k+1}^y = 0$, and $L_{k+1} = 2$,

$$\alpha_{1,k+1} = \lambda, \ \alpha_{2,k+1} = \beta\lambda, \tag{3.15}$$

$$\eta_{1,k+1} = \frac{1}{1-\beta^{\vartheta_1} \exp\left[-\frac{1}{2}y^\top \Pi y - (1-\beta)\lambda\right]}, \tag{3.16}$$

$$\eta_{2,k+1} = \frac{-\beta^{\vartheta_1} \exp\left[-\frac{1}{2}y^\top \Pi y - (1-\beta)\lambda\right]}{1-\beta^{\vartheta_1} \exp\left[-\frac{1}{2}y^\top \Pi y - (1-\beta)\lambda\right]}, \tag{3.17}$$

for $\gamma_{k+1}^y = 1$. In addition, for $k \geq 0$, denote $\{\hat{x}_k\}$ and $\{P_k\}$ as sequences that evolve according to

$$P_k = P_k^- - P_k^- C^\top [R + (1-\gamma_k^y)\Pi^{-1} + CP_k^- C^\top]^{-1} CP_k^-, \tag{3.18}$$

$$\hat{x}_k = P_k(P_k^-)^{-1}\hat{x}_k^- + \gamma_k^y P_k C^\top R^{-1} y_k, \tag{3.19}$$

$$P_{k+1}^- = AP_k A^\top + Q, \tag{3.20}$$

$$\hat{x}_{k+1}^- = A\hat{x}_k, \tag{3.21}$$

where $\hat{x}_0^-$ and P_0^- are the mean and covariance of the initial value x_0. Based on the above definitions, we are ready to present our first result.

Theorem 3.1. *For the event-based state estimation problem in Fig. 3.1 with the energy-dependent event-triggering scheme in (3.4)-(3.5), the joint probability distribution of z_k and x_k $(k \geq 1)$ conditioned on the event-triggered measurement information $\mathcal{I}_k$ has the form:*

$$f_{z_k,x_k}(z, x | \mathcal{I}_k) = \left[\gamma_k^z \cdot \delta(z - z_k) + (1 - \gamma_k^z) \sum_{i=1}^{L_k} \eta_{i,k} \frac{\alpha_{i,k}^{z-\vartheta_k} \exp(-\alpha_{i,k})}{(z - \vartheta_k)!} \right]$$
$$\cdot \frac{1}{(2\pi)^{\frac{n}{2}} \det^{\frac{1}{2}} P_k} \exp\left[-(x - \hat{x}_k)^\top P_k^{-1}(x - \hat{x}_k)/2 \right], \qquad (3.22)$$

where the paramters L_k, $\alpha_{i,k}$, $\eta_{i,k}$, $\hat{x}_k$ and P_k evolve in recursive forms according to (3.8)-(3.21).

Proof of Theorem 3.1. To prove the result, we consider the general case first and the basis when $k = 1$ can be proved by following a similar procedure as the general case, which will be shown at the end of the proof. We assume at time instant k, the probability distribution function of $z_k = z, x_k = x | \mathcal{I}_k$ has the form in equation (3.22) and then we prove this form holds for the joint probability distribution of z_{k+1} and x_{k+1} conditioned on $\mathcal{I}_{k+1}$.

(a) Firstly, we consider the case $\gamma_k^z = 0$. According to equation (3.22), we have

$$f_{z_k,x_k}(z, x | \mathcal{I}_k) = \sum_{i=1}^{L_k} \eta_{i,k} \frac{\alpha_{i,k}^{z-\vartheta_k} \exp(-\alpha_{i,k})}{(z - \vartheta_k)!(2\pi)^{\frac{n}{2}} \det^{\frac{1}{2}} P_k} \cdot \exp\left[-(x - \hat{x}_k)^\top P_k^{-1}(x - \hat{x}_k)/2 \right].$$

For time instant $k + 1$, if $\gamma_{k+1}^z = 0$, we have

$$f_{z_{k+1},x_{k+1}}(z, x | \mathcal{I}_k)$$
$$= \sum_{t=\vartheta_k}^{z+\gamma_k^y} \int_{\mathbb{R}^n} f_{z_{k+1},x_{k+1}}(z, x | \mathcal{I}_k, z_k = t, x_k = \xi) f_{z_k,x_k}(t, \xi | \mathcal{I}_k) d\xi.$$

Since $\mathcal{I}_k$ is determined by x_0, z_0, $w_{0:k-1}$, $v_{0:k}$ and $\rho_{0:k-1}$, we have

$$f_{z_{k+1},x_{k+1}}(z, x | \mathcal{I}_k, z_k = t, x_k = \xi)$$
$$= f_{w_k}(x - Ax_k | x_k = \xi) \Pr(z_{k+1} = z | z_k = t, \gamma_k^y)$$
$$= \frac{1}{(2\pi)^{\frac{n}{2}} \det^{\frac{1}{2}} Q} \exp\left[-(x - A\xi)^\top Q^{-1}(x - A\xi)/2 \right] \frac{\lambda^{z-t+\gamma_k^y} \exp(-\lambda)}{(z - t + \gamma_k^y)!}.$$

48

Therefore, for $\gamma_{k+1}^z = 0$, we have

$$
f_{z_{k+1}, x_{k+1}}(z, x | \mathcal{I}_k)
$$

$$
= \sum_{t=\vartheta_k}^{z+\gamma_k^y} \int_{\mathbb{R}^n} \frac{1}{(2\pi)^{\frac{n}{2}} \det^{\frac{1}{2}} Q} \exp\left[-\tfrac{1}{2}(x - A\xi)^\top Q^{-1}(x - A\xi)\right]
$$

$$
\cdot \frac{\lambda^{z-t+\gamma_k^y} \exp(-\lambda)}{(z - t + \gamma_k^y)!} \sum_{i=1}^{L_k} \eta_{i,k} \frac{\alpha_{i,k}^{t-\vartheta_k} \exp(-\alpha_{i,k})}{(t - \vartheta_k)!(2\pi)^{\frac{n}{2}} \det^{\frac{1}{2}} P_k}
$$

$$
\cdot \exp\left[-\tfrac{1}{2}(\xi - \hat{x}_k)^\top P_k^{-1}(\xi - \hat{x}_k)\right] \mathrm{d}\xi
$$

$$
\propto \sum_{i=1}^{L_k} \sum_{t=\vartheta_k}^{z+\gamma_k^y} \eta_{i,k} \frac{\alpha_{i,k}^{t-\vartheta_k} \lambda^{z-t+\gamma_k^y} \exp(-\alpha_{i,k} - \lambda)}{(t - \vartheta_k)!(z - t + \gamma_k^y)!}
$$

$$
\cdot \int_{\mathbb{R}^n} \exp\left[-\tfrac{1}{2}(x - A\xi)^\top Q^{-1}(x - A\xi)\right] \exp\left[-\tfrac{1}{2}(\xi - \hat{x}_k)^\top P_k^{-1}(\xi - \hat{x}_k)\right] \mathrm{d}\xi \qquad (3.23)
$$

$$
\propto \sum_{i=1}^{L_k} \eta_{i,k} \frac{(\alpha_{i,k} + \lambda)^{z-\vartheta_{k+1}} \exp(-\alpha_{i,k} - \lambda)}{(2\pi)^{\frac{n}{2}} (z - \vartheta_{k+1})! \det^{\frac{1}{2}}(AP_k A^\top + Q)}
$$

$$
\cdot \exp\left[-\tfrac{1}{2}(x - A\hat{x}_k)^\top (AP_k A^\top + Q)^{-1}(x - A\hat{x}_k)\right], \qquad (3.24)
$$

where Lemma 15 in [57] is used to derive equation (3.24) from (3.23). For $\Pr(\gamma_{k+1} = 0 | x_{k+1} = x, z_{k+1} = z, \mathcal{I}_k)$, we have

$$
\Pr(\gamma_{k+1} = 0 | x_{k+1} = x, z_{k+1} = z, \mathcal{I}_k)
$$

$$
= \int_{\mathbb{R}^m} \Pr(\gamma_{k+1} = 0 | y_{k+1} = y, x_{k+1} = x, z_{k+1} = z, \mathcal{I}_k) f_{y_{k+1}}(y | x_{k+1} = x, z_{k+1} = z, \mathcal{I}_k) \mathrm{d}y
$$

$$
= \int_{\mathbb{R}^m} \beta^z \exp\left(-\tfrac{1}{2} y^\top \Pi y\right) \frac{1}{(2\pi)^{\frac{m}{2}} \det^{\frac{1}{2}} R} \exp\left[-\tfrac{1}{2}(y - Cx)^\top R^{-1}(y - Cx)\right] \mathrm{d}y
$$

$$
= \frac{\beta^z}{\det^{\frac{1}{2}}(\Pi R + I)} \exp\left[-\tfrac{1}{2} x^\top C^\top (R + \Pi^{-1})^{-1} Cx\right].
$$

49

Based on the above calculations, for $\gamma^z_{k+1} = 0$ and $\gamma^y_{k+1} = 0$, we have

$$
\begin{aligned}
&f_{z_{k+1},x_{k+1}}(z,x|\mathcal{I}_{k+1})\\
&\propto \Pr(\gamma_{k+1}=0|x_{k+1}=x,z_{k+1}=z,\mathcal{I}_k)f_{z_{k+1},x_{k+1}}(z,x|\mathcal{I}_k)\\
&\propto \frac{\beta^z}{\det^{\frac{1}{2}}(\Pi R + I)} \exp\left[-\tfrac{1}{2}x^\top C^\top (R+\Pi^{-1})^{-1}Cx\right]\\
&\quad \cdot \sum_{i=1}^{L_k} \eta_{i,k} \frac{(\alpha_{i,k}+\lambda)^{z-\vartheta_{k+1}}\exp\left(-\alpha_{i,k}-\lambda\right)}{(2\pi)^{\frac{n}{2}}(z-\vartheta_{k+1})!\det^{\frac{1}{2}}(AP_kA^\top + Q)}\\
&\quad \cdot \exp\left[-\tfrac{1}{2}(x-A\hat{x}_k)^\top (AP_kA^\top + Q)^{-1}(x-A\hat{x}_k)\right] \qquad\qquad (3.25)\\
&\propto \sum_{i=1}^{L_k} \eta_{i,k} \frac{\beta^z(\alpha_{i,k}+\lambda)^{z-\vartheta_{k+1}}\exp\left(-\alpha_{i,k}-\lambda\right)}{(2\pi)^{\frac{n}{2}}(z-\vartheta_{k+1})!\det^{\frac{1}{2}}P_{k+1}}\\
&\quad \cdot \exp\left[-\tfrac{1}{2}(x-\hat{x}_{k+1})^\top P_{k+1}^{-1}(x-\hat{x}_{k+1})\right], \qquad\qquad (3.26)
\end{aligned}
$$

where Lemma 15 in [57] is used to derive equation (3.26) from (3.25) and

$$
\begin{aligned}
P_{k+1} &= \left[(AP_kA^\top + Q)^{-1} + C^\top(R+\Pi^{-1})^{-1}C\right]^{-1}\\
&= P_{k+1}^- - P_{k+1}^- C^\top[(R+\Pi^{-1}) + CP_{k+1}^- C^\top]^{-1}CP_{k+1}^-,\\
\hat{x}_{k+1} &= \left[(AP_kA^\top + Q)^{-1} + C^\top(R+\Pi^{-1})^{-1}C)\right]^{-1}(AP_kA^\top + Q)^{-1}A\hat{x}_k\\
&= P_{k+1}(P_{k+1}^-)^{-1}A\hat{x}_k.
\end{aligned}
$$

Based on some further calculations, it is easy to verify that for $\gamma^z_{k+1} = 0$ and $\gamma^y_{k+1} = 0$,

$$
\begin{aligned}
&f_{z_{k+1},x_{k+1}}(z,x|\mathcal{I}_{k+1})\\
&= \sum_{i=1}^{L_{k+1}} \eta_{i,k+1} \frac{\alpha_{i,k+1}^{z-\vartheta_{k+1}}\exp\left(-\alpha_{i,k+1}\right)}{(z-\vartheta_{k+1})!(2\pi)^{\frac{n}{2}}\det^{\frac{1}{2}}P_{k+1}} \exp\left[-\tfrac{1}{2}(x-\hat{x}_{k+1})^\top P_{k+1}^{-1}(x-\hat{x}_{k+1})\right],
\end{aligned}
$$

where $L_{k+1} = L_k$, $\alpha_{i,k+1} = \beta\cdot(\alpha_{i,k}+\lambda)$ and

$$
\eta_{i,k+1} = \frac{\eta_{i,k}\exp\left[-(1-\beta)(\alpha_{i,k}+\lambda)\right]}{\sum_{j=1}^{L_k}\eta_{j,k}\exp\left[-(1-\beta)(\alpha_{j,k}+\lambda)\right]}.
$$

Now we move on to consider the case $\gamma^z_{k+1} = 0$ and $\gamma^y_{k+1} = 1$. To do this, we first consider the term $f_{z_{k+1},x_{k+1}}(z,x|\mathcal{I}_k,y_{k+1}=y)$. Based on the Bayes' rule and equation

(3.24), we have

$$f_{z_{k+1}, x_{k+1}}(z, x | \mathcal{I}_k, y_{k+1} = y)$$

$$\propto f_{y_{k+1}}(y | z_{z+1} = z, x_{k+1} = x, \mathcal{I}_k) f_{z_{k+1}, x_{k+1}}(z, x | \mathcal{I}_k)$$

$$= \frac{1}{(2\pi)^{\frac{m}{2}} \det^{\frac{1}{2}} R} \exp\left[-\frac{1}{2}(y - Cx)^\top R^{-1}(y - Cx)\right]$$

$$\cdot \sum_{i=1}^{L_k} \eta_{i,k} \frac{(\alpha_{i,k} + \lambda)^{z - \vartheta_{k+1}} \exp\left(-\alpha_{i,k} - \lambda\right)}{(2\pi)^{\frac{n}{2}} (z - \vartheta_{k+1})! \det^{\frac{1}{2}}(AP_k A^\top + Q)}$$

$$\cdot \exp\left[-\frac{1}{2}(x - A\hat{x}_k)^\top (AP_k A^\top + Q)^{-1}(x - A\hat{x}_k)\right] \tag{3.27}$$

$$\propto \sum_{i=1}^{L_k} \eta_{i,k} \frac{(\alpha_{i,k} + \lambda)^{z - \vartheta_{k+1}} \exp\left(-\alpha_{i,k} - \lambda\right)}{(2\pi)^{\frac{n}{2}} (z - \vartheta_{k+1})! \det^{\frac{1}{2}} P_{k+1}}$$

$$\cdot \exp\left[-\frac{1}{2}(x - \hat{x}_{k+1})^\top P_{k+1}^{-1}(x - \hat{x}_{k+1})\right], \tag{3.28}$$

where Lemma 15 in [57] is used from (3.27) to (3.28), and

$$P_{k+1}^- = AP_k A^\top + Q,$$

$$P_{k+1} = P_{k+1}^- - P_{k+1}^- C^\top (R + C P_{k+1}^- C^\top)^{-1} C P_{k+1}^-,$$

$$\hat{x}_{k+1}^- = A\hat{x}_k,$$

$$\hat{x}_{k+1} = \left[(P_{k+1}^-)^{-1} + C^\top R^{-1} C\right]^{-1} \left[(P_{k+1}^-)^{-1} \hat{x}_{k+1}^- + C^\top R^{-1} y\right]$$

$$= P_{k+1}(P_{k+1}^-)^{-1} \hat{x}_{k+1}^- + P_{k+1} C^\top R^{-1} y.$$

Based on the above results, for $\gamma_{k+1}^z = 0$ and $\gamma_{k+1}^y = 1$, we have

$$f_{z_{k+1}, x_{k+1}}(z, x | \mathcal{I}_{k+1})$$

$$\propto \Pr(\gamma_{k+1} = 1 | x_{k+1} = x, z_{k+1} = z, y_{k+1} = y, \mathcal{I}_k)$$

$$\cdot f_{z_{k+1}, x_{k+1}}(z, x | \mathcal{I}_k, y_{k+1} = y)$$

$$\propto \left[1 - \beta^z \exp\left(-\frac{1}{2} y^\top \Pi y\right)\right] \sum_{i=1}^{L_k} \eta_{i,k} \frac{(\alpha_{i,k} + \lambda)^{z - \vartheta_{k+1}} \exp\left(-\alpha_{i,k} - \lambda\right)}{(2\pi)^{\frac{n}{2}} (z - \vartheta_{k+1})! \det^{\frac{1}{2}} P_{k+1}}$$

$$\cdot \exp\left[-\frac{1}{2}(x - \hat{x}_{k+1})^\top P_{k+1}^{-1}(x - \hat{x}_{k+1})\right].$$

It is easy to verify that for $\gamma_{k+1}^z = 0$ and $\gamma_{k+1}^y = 1$,

$$f_{z_{k+1}, x_{k+1}}(z, x | \mathcal{I}_{k+1})$$

$$= \sum_{i=1}^{L_{k+1}} \eta_{i,k+1} \frac{\alpha_{i,k+1}^{z - \vartheta_{k+1}} \exp\left(-\alpha_{i,k+1}\right)}{(2\pi)^{\frac{n}{2}} (z - \vartheta_{k+1})! \det^{\frac{1}{2}} P_{k+1}} \exp\left[-\frac{1}{2}(x - \hat{x}_{k+1})^\top P_{k+1}^{-1}(x - \hat{x}_{k+1})\right],$$

where $L_{k+1} = 2L_k$,

$$\alpha_{i,k+1} = \begin{cases} \alpha_{i,k} + \lambda, & \text{if } 1 \le i \le L_k; \\ \beta(\alpha_{i-L_k,k} + \lambda), & \text{if } L_k < i \le L_{k+1}; \end{cases}$$

$$\eta_{i,k+1} = \begin{cases} \dfrac{\eta_{i,k}}{1-\sum_{j=1}^{L_k} \eta_{j,k}\beta^{\vartheta_{k+1}}\exp\left[-\frac{1}{2}y^\top\Pi y-(1-\beta)(\alpha_{j,k}+\lambda)\right]}, & \text{if } 1 \le i \le L_k; \\ \dfrac{-\eta_{i-L_k,k}\beta^{\vartheta_{k+1}}\exp\left[-\frac{1}{2}y^\top\Pi y-(1-\beta)(\alpha_{i-L_k,k}+\lambda)\right]}{1-\sum_{j=1}^{L_k} \eta_{j,k}\beta^{\vartheta_{k+1}}\exp\left[-\frac{1}{2}y^\top\Pi y-(1-\beta)(\alpha_{j,k}+\lambda)\right]}, & \text{if } L_k < i \le L_{k+1}. \end{cases}$$

For time instant $k + 1$, if $\gamma_{k+1}^z = 1$, the exact value of the energy level at time $k + 1$ is known by the estimator, thus, it is straightforward to obtain

$$\begin{aligned} &f_{z_{k+1},x_{k+1}}(z, x|\mathcal{I}_{k+1}) \\ &= \frac{\delta(z-z_{k+1})}{(2\pi)^{\frac{n}{2}} \det^{\frac{1}{2}} P_{k+1}} \exp\left[-\tfrac{1}{2}(x - \hat{x}_{k+1})^\top P_{k+1}^{-1}(x - \hat{x}_{k+1})\right], \end{aligned} \tag{3.29}$$

where $\hat{x}_{k+1}$ and P_{k+1} follow equations (3.18)-(3.21).

(b) Secondly, we consider the case $\gamma_k^z = 1$. According to equation (3.22), we have

$$\begin{aligned} &f_{z_k,x_k}(z, x|\mathcal{I}_k) \\ &= \frac{\delta(z-z_k)}{(2\pi)^{\frac{n}{2}} \det^{\frac{1}{2}} P_k} \exp\left[-\tfrac{1}{2}(x - \hat{x}_k)^\top P_k^{-1}(x - \hat{x}_k)\right]. \end{aligned}$$

For time instant $k + 1$, if $\gamma_{k+1}^z = 0$, following a similar argument as in part (a), we have

$$\begin{aligned} &f_{z_{k+1},x_{k+1}}(z, x|\mathcal{I}_k) \\ &= f_{z_{k+1}}(z|z_k, \gamma_k^y) f_{x_{k+1}}(x|\mathcal{I}_k) \\ &= \frac{\lambda^{z-\vartheta_{k+1}} \exp(-\lambda)}{(z-\vartheta_{k+1})!(2\pi)^{\frac{n}{2}} \det^{\frac{1}{2}} P_{k+1}^-} \exp\left[-\tfrac{1}{2}(x - \hat{x}_{k+1}^-)^\top (P_{k+1}^-)^{-1}(x - \hat{x}_{k+1}^-)\right], \end{aligned}$$

where $\vartheta_{k+1} = \vartheta_k - \gamma_k^y$, $\vartheta_k = z_k$, $P_{k+1}^- = AP_kA^\top + Q$, and $\hat{x}_{k+1}^- = A\hat{x}_k$. Based on the above results, we have for $\gamma_{k+1}^y = 0$,

$$\begin{aligned} &f_{z_{k+1},x_{k+1}}(z, x|\mathcal{I}_{k+1}) \\ &= \frac{(\beta\lambda)^{z-\vartheta_{k+1}} \exp(-\beta\lambda)}{(z-\vartheta_{k+1})!(2\pi)^{\frac{n}{2}} \det^{\frac{1}{2}} P_{k+1}} \exp\left[-\tfrac{1}{2}(x - \hat{x}_{k+1})^\top P_{k+1}^{-1}(x - \hat{x}_{k+1})\right]; \end{aligned}$$

for $\gamma_{k+1}^y = 1$,

$$\begin{aligned} &f_{z_{k+1},x_{k+1}}(z, x|\mathcal{I}_{k+1}) \\ &= \left[1 - \beta^z \exp\left(-\tfrac{1}{2}y^\top\Pi y\right)\right] \frac{\lambda^{z-\vartheta_{k+1}} \exp(-\lambda)}{(z-\vartheta_{k+1})!(2\pi)^{\frac{n}{2}} \det^{\frac{1}{2}} P_{k+1}} \\ &\quad \cdot \exp\left[-\tfrac{1}{2}(x - \hat{x}_{k+1})^\top P_{k+1}^{-1}(x - \hat{x}_{k+1})\right] \bigg/ \left\{1 - \beta^{\vartheta_{k+1}} \exp\left[-\tfrac{1}{2}y^\top\Pi y - (1-\beta)\lambda\right]\right\}. \end{aligned}$$

Therefore we have for $\gamma_{k+1}^z = 0$,

$$f_{z_{k+1}, x_{k+1}}(z, x | \mathcal{I}_{k+1})$$
$$= \sum_{i=1}^{L_{k+1}} \eta_{i,k+1} \frac{\alpha_{i,k+1}^{z-\vartheta_{k+1}} \exp(-\alpha_{i,k+1})}{(z-\vartheta_{k+1})! (2\pi)^{\frac{n}{2}} \det^{\frac{1}{2}} P_{k+1}} \exp\left[-\tfrac{1}{2}(x - \hat{x}_{k+1})^\top P_{k+1}^{-1}(x - \hat{x}_{k+1})\right],$$

where for $\gamma_{k+1}^y = 0$, we have

$$L_{k+1} = 1, \ \eta_{1,k+1} = 1, \ \alpha_{1,k+1} = \beta\lambda;$$

for $\gamma_{k+1}^y = 1$, we have $L_{k+1} = 2$,

$$\alpha_{1,k+1} = \lambda, \ \alpha_{2,k+1} = \beta\lambda,$$
$$\eta_{1,k+1} = \frac{1}{1 - \beta^{\vartheta_{k+1}} \exp\left[-\tfrac{1}{2}y^\top \Pi y - (1-\beta)\lambda\right]},$$
$$\eta_{2,k+1} = \frac{-\beta^{\vartheta_{k+1}} \exp\left[-\tfrac{1}{2}y^\top \Pi y - (1-\beta)\lambda\right]}{1 - \beta^{\vartheta_{k+1}} \exp\left[-\tfrac{1}{2}y^\top \Pi y - (1-\beta)\lambda\right]}.$$

For time instant $k + 1$, if $\gamma_{k+1}^z = 1$, it is straightforward to obtain the joint probability distribution of z_{k+1} and x_{k+1} as described in (3.29), where $\hat{x}_{k+1}$ and P_{k+1} follow equations (3.18)-(3.21). Recall that we assume $\gamma_0^z = 1$. For $k = 1$, following a similar argument as that for the general case in (b), we obtain $f_{z_1, x_1}(z, x | \mathcal{I}_1)$, which satisfies equation (3.22). It is not difficult to obtain that $\hat{x}_1^- = A\hat{x}_0$, $P_1^- = AP_0A^\top + Q$, and the other parameters follow (3.14)-(3.19). This completes the proof. $\qquad\square$

Remark 3.1. Since a Poisson energy harvesting process is considered, it is natural to expect that the distribution of z_k may follow an exponential form. The issue, however, is that due to the energy-based transmission protocol, the event-triggered information $\{\gamma_k^y, \gamma_k^y y_k\}$ provides not only information about x_k, but also information about z_k; we need to investigate how the conditional probability distribution of z_k evolves when $\gamma_k^y = 0$ and how it is related to y_k if $\gamma_k^y = 1$, which are shown in equations (3.8)-(3.17). According to equations (3.8)-(3.13), for $\gamma_k^z = 0$, the theorem indicates that $L_{k+1} = L_k$ if $\gamma_k^y = 0$, where $\alpha_{i,k+1}$ and $\eta_{i,k+1}$ can be calculated following equations (3.9)-(3.10), and that $L_k = 2L_k$ if $\gamma_k^y = 1$, where $\alpha_{i,k+1}$ and $\eta_{i,k+1}$ evolve according to piecewise functions, depending on the index i as shown in equations (3.12)-(3.13). The parameters are updated at time instant $k + 1$ according to equations (3.14)-(3.17) for $\gamma_k^z = 1$, which can be seen as the re-initialization of the scaled exponential distribution.

In Theorem 3.1, the joint probability distribution of the state and energy level derived is in a recursive form, based on which, the marginal probability distributions

of the state and energy level can be obtained by integration, then it is straightforward to derive the MMSE estimates of the state and energy level, which are summarized in Theorem 3.2.

Theorem 3.2. *For the event-based state estimation problem in Fig. 3.1 with the energy-dependent event-triggering scheme in (3.4)-(3.5) and $k \geq 1$, the probability distribution $f_{x_k}(x|\mathcal{I}_k)$ is $\mathcal{N}(\hat{x}_k, P_k)$, where $\hat{x}_k$ and P_k evolve according to equations (3.18)-(3.21). The MMSE estimate of the state $\bar{x}_k$ satisfies*

$$\bar{x}_k = \hat{x}_k. \tag{3.30}$$

The MMSE estimate $\bar{z}_k$ of the sensor energy level satisfies

$$\bar{z}_k = \gamma_k^z z_k + (1 - \gamma_k^z)(\vartheta_k + \sum_{i=1}^{L_k} \eta_{i,k} \alpha_{i,k}). \tag{3.31}$$

Remark 3.2. Notice that in equation (3.31), the estimate of z_k is recursive and depends on $\sum_{i=1}^{L_k} \eta_{i,k} \alpha_{i,k}$ if $\gamma_k^z = 0$, where $L_{k+1} = 2L_k$ if $\gamma_k^y = 1$, which indicates that the computational complexity of the estimate of z_k will increase exponentially with respect to the number of instants when $\gamma_k^y = 1$. One solution to this issue is to transmit the energy level z_k to the remote estimator when the computational complexity reaches a specific limit, i.e., we have $\gamma_k^z = 1$ if $\gamma_{k_r}^z = 1$, $\gamma_t^z = 0$ for $t = \mathbb{N}_{k_r:k-1}$ and $\sum_{t=k_r}^{k-1} \gamma_t^y \geq T$, where T is a predefined nonnegative constant. We note that this sensor transmission of z_k only affects the implementation of the estimator, but does not affect the event-triggering process $\{\gamma_k^y\}$.

3.4 Relationship between the communication rate and energy harvesting rate

In this section, we analyze the relationship between the communication rate and energy harvesting rate for systems with our proposed transmission protocol. We focus on the analysis of the average measurement transmission rate, since the transmission of y_k is the main cause of sensor energy consumption and more likely to happen compared with the transmission of z_k, which will also be shown in the numerical example. Define the average communication rate as $\gamma := \limsup_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \gamma_k^y$. To allow asymptotic analysis, we assume the system in (3.1) is stable.

Specifically, if $\lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} (\gamma_k^y - \rho_k) < -\varepsilon$ holds with a sufficiently large probability for some small and positive ε, we have from equation (3.4) that $\lim_{k \to \infty} z_k \to \infty$

w.p. $1 - \zeta$ for some sufficiently small $\zeta > 0$. Recall that $\lim_{k \to \infty} \frac{1}{N} \sum_{k=1}^{N} \rho_k = \lambda$ due to the nature of Poisson process. From the definition of γ_k^y, this further implies

$$\lim_{k \to \infty} \Pr(\gamma_k^y = 0) = \lim_{k \to \infty} \int_{\mathbb{R}^m} \sum_{z=0}^{\infty} \Pr(\gamma_k^y = 0 | z_k = z, y_k = y) f(y_k = y | z_k = z) \Pr(z_k = z) \mathrm{d}y$$

$$= \lim_{k \to \infty} \int_{\mathbb{R}^m} \sum_{z=0}^{\infty} \beta^z \exp\left(-\tfrac{1}{2} y^\top \Pi y\right) f(y_k = y | z_k = z) \Pr(z_k = z) \mathrm{d}y$$

$$\approx \lim_{k \to \infty} \lim_{z \to \infty} \int_{\mathbb{R}^m} \beta^z \exp\left(-\tfrac{1}{2} y^\top \Pi y\right) f(y_k = y | z_k = z) \mathrm{d}y = 0, \qquad (3.32)$$

where (3.32) is due to that ζ is close to 0, the probability of $\lim_{k \to \infty} z_k \to \infty$ is $1 - \zeta$ and the summation of probabilities for z_k taking all the other values equals ζ. In other words, we have $\limsup_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \gamma_k^y \approx 1$. Notice that we assume $\lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} (\gamma_k^y - \rho_k) < -\varepsilon$ holds with a sufficiently large probability for some small ε, which requires $1 - \lambda < -\varepsilon$, or $\lambda > 1 + \varepsilon$.

On the other hand, if $\lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} (\gamma_k^y - \rho_k) > \varepsilon$ with a sufficiently large probability for some small $\varepsilon > 0$, $\lim_{k \to \infty} z_k \to 0$ holds w.p. $1 - \zeta$ for some sufficiently small $\zeta > 0$. In this way, we have

$$\lim_{k \to \infty} \Pr(\gamma_k^y = 0) = \lim_{k \to \infty} \int_{\mathbb{R}^m} \sum_{z=0}^{\infty} \beta^z \exp\left(-\tfrac{1}{2} y^\top \Pi y\right) f(y_k = y | z_k = z) \Pr(z_k = z) \mathrm{d}y$$

$$\approx \int_{\mathbb{R}^m} \exp\left(-\tfrac{1}{2} y^\top \Pi y\right) f(y_k = y | z_k = z) \mathrm{d}y. \qquad (3.33)$$

Noticing that when $\lim_{k \to \infty} z_k \to 0$ holds w.p. $1 - \zeta$ for some sufficiently small $\zeta > 0$, $\lim_{k \to \infty} z_k$ behaves approximately like a deterministic variable such that $f(y_k = y | z_k = z) \approx f(y_k = y)$. Since the system in (3.1) is stable, y_k becomes a stationary Gaussian process as $k \to \infty$, thus when $\lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} (\gamma_k^y - \rho_k) < \varepsilon$ with a sufficiently large probability, $\lim_{k \to \infty} \Pr(\gamma_k^y = 0)$ will approximately stay at a constant level, which can be obtained according to Theorem 3 in [23]. Combining the above discussions, another interesting observation is that when $\lim_{k \to \infty} \Pr(\gamma_k^y = 0)$ is neither 0 nor as described in (3.33), we would infer that the event $\lim_{N \to \infty} \frac{1}{N} | \sum_{k=1}^{N} (\gamma_k^y - \rho_k) | > \varepsilon$ would hold for a small probability, which means that $\lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \gamma_k^y \approx \lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \rho_k = \lambda$. Therefore, the average communication rate will match the energy harvesting rate when λ is large enough but smaller than 1. The discussions here will be further illustrated and verified in the next section.

3.5   Simulation example

The effectiveness of the proposed method is shown by a numerical example in this section. We consider a stable second-order system, which can be described by equations (3.1) and (3.2) with parameters $A = [0.7\ 0.1; 0.2\ 0.8]$ and $C = [0.5\ 0.1; 0.1\ 0.4]$. The process noise w_k and measurement noise v_k are zero-mean Gaussians with covariances $Q = [0.8\ 0; 0\ 0.8]$ and $R = [0.3\ 0; 0\ 0.3]$.

The initial value x_0 of the state is Gaussian distributed with mean $\hat{x}_0^- = \mu_0 = [0\ 0]^\top$ and covariance $P_0 = [0.2\ 0; 0\ 0.2]$. The state is measured by an energy harvesting sensor satisfying equation (3.4), with initial energy $z_0 = 15$, and ρ_k follows the Poisson process described in equation (3.3) with $\lambda = 0.55$, which corresponds to the average harvested energy. The measurement will be transmitted to the remote estimator whenever the energy-dependent event-triggering condition in (3.5) is satisfied, where we choose $\beta = 0.96$ and

$$\Pi_0 = \begin{bmatrix} 0.5 & 0.1 \\ 0.2 & 0.6 \end{bmatrix}.$$

The sensor transmits the exact energy level to the estimator whenever the condition in (3.7) or the specific criterion discussed in Remark 4 is satisfied. We assume $T = 15$ indicating that the maximum value of L_k is $2^{15} = 32768$. The performance comparison of the proposed event-triggered estimator obtained for $N = 200$ runs, the estimates of the Kalman filter and the Kalman filter with intermittent observations (KF with IO) [67] obtained using the same communication sequence is shown in Fig. 3.2. The resultant average estimation errors for the Kalman filter, the proposed estimator and the KF with IO are 1.3779, 1.6892 and 1.969, respectively, where we observe that the proposed method outperforms the KF with IO, although not as good as the Kalman filter. When $\gamma_k^z = 1$, the altitude of $\gamma_k^z = 1$ is shown as 0.5 for the purpose of easy distinction between γ_k^z and γ_k^y in Fig. 3.2. The average transmission rate of measurements $\sum_{k=1}^N \gamma_k^y = 0.575$ is more than 9 times larger than the average transmission rate of the energy level $\sum_{k=1}^N \gamma_k^z = 0.06$.

Next, we numerically evaluate the tradeoff between the communication rates and the estimation performances for different energy levels. We run the system for $800,000$ time instants and keep all parameters the same. The corresponding communication rate for one specific energy level i is defined as $\frac{1}{N_i} \sum_{z_k=i} \gamma_k^y$, where N_i is the total number of time instants with energy level $z_k = i$. The average estimation errors and communication rates for different energy levels are evaluated and compared in Fig. 3.3, where the blue
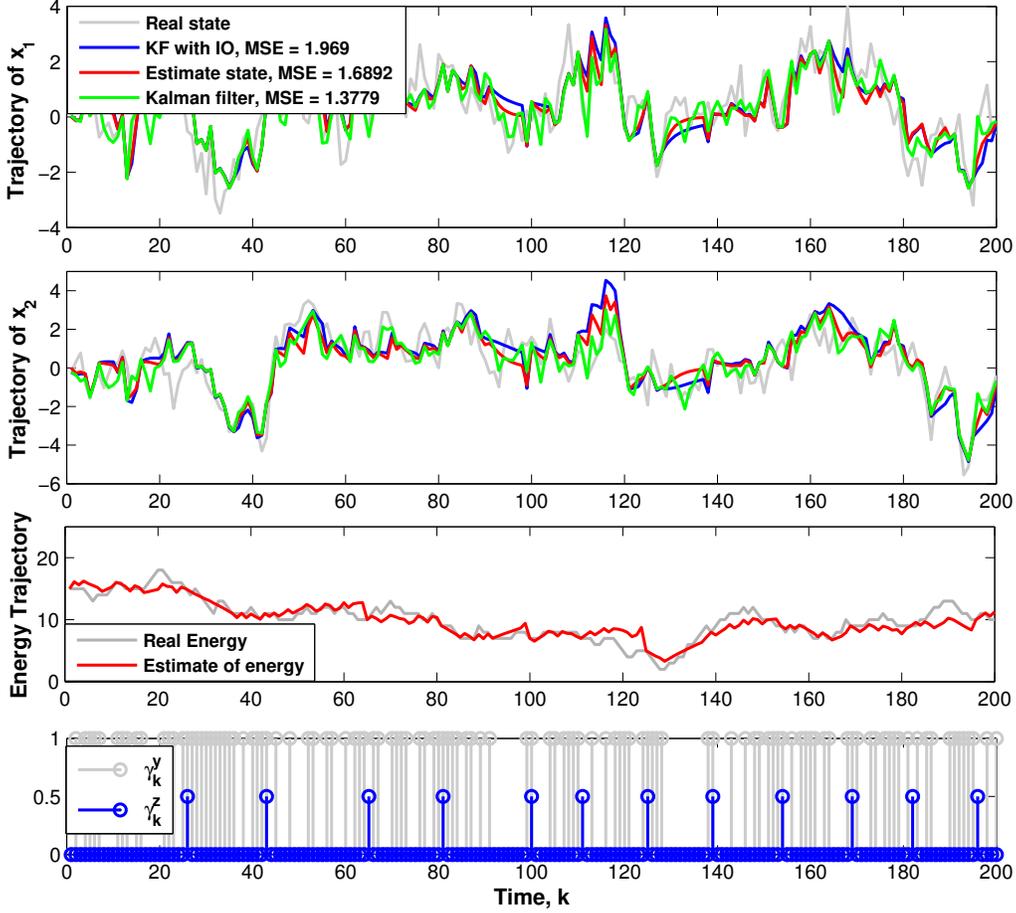
Figure 3.2: Simulation results of the second-order system.

line shows the variation of the communication rate by the left vertical axis. The other lines in Fig. 3.3 are the estimation errors by the right vertical axis. Fig. 3.3 indicates that the estimation performance is sacrificed to maintain low average communication rate when limited energy is available.

Finally, we numerically evaluate the relationship between the average communication rate γ and average energy harvesting rate λ. Recall that the average communication rate is defined as $\gamma = \lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \gamma_k^y$, where we choose runtime $N = 80,000$. The average communication rate is shown in Fig. 3.4 when λ varies from 0 to 1.2 for $\Pi = 0.2\Pi_0, 0.4\Pi_0, 0.6\Pi_0, 0.8\Pi_0, 1.0\Pi_0$, where the other parameters remain the same. We observe that when $\lambda \to 0$, the average communication rates in Fig. 3.4 are approximately equal to the results calculated according to Equation (28) of Theorem 3 in [23], which are 0.1090, 0.1929, 0.2601, 0.3154 and 0.3620, respectively. When $\lambda > 1$, the average communication rates γ reach 1. Finally, an interesting phenomenon observed
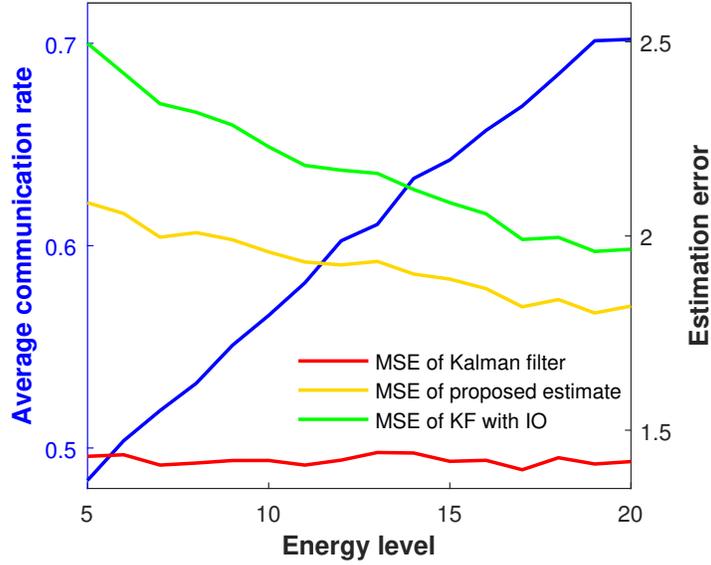
Figure 3.3: Average communication rates and estimation performance with respect to different energy levels.

is that when λ is relatively large (approximately 0.55 in Fig. 3.4) but less than 1, the average communication rates coincide for all choices of Π and becomes approximately equal to λ, which is consistent with our analysis in Section 3.4.

3.6   Conclusion

In this chapter, a remote state estimation problem for a linear Gaussian system with an energy harvesting sensor and an energy-dependent event-triggered scheduling strategy is investigated. The joint conditional probability distribution of the state and energy based on the information received at the remote estimator is derived. It is proved that the estimate of the state evolves in a recursive form. The analysis in Section 3.4 shows that the average communication rate will match the energy harvesting rate if λ is large enough but less than 1 due to the proposed event-triggering condition. Finally, the performance improvement of the proposed estimator compared with the KF with IO is illustrated and the features of the average communication rate analyzed in Section 3.4 are verified by numerical simulations.
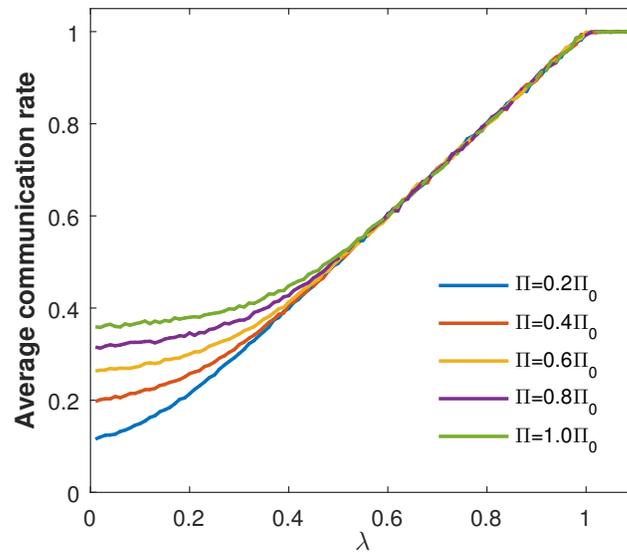
Figure 3.4: Relationship between the average communication rate and the energy harvesting rate.

Chapter 4

Optimal risk-sensitive state estimation

In this chapter, we consider the risk-sensitive filtering problem, which is an optimal state estimation problem in a certain sense. The risk-sensitive approach counteracts against disturbances and model-plant mismatches by considering an exponential-form cost function, which penalizes all the higher order moments of the estimation error energy. Thus, the risk-sensitive filter is more robust compared with the MMSE filter due to the efforts made at the design stage.

4.1 Robust event-triggered state estimation: A risk-sensitive approach[*]

In this section, we investigate the robust event-triggered state estimation problem utilizing the risk-sensitive filtering approach.

4.1.1 Introduction

It is well known in the control community that the optimal MMSE filter [66] and the robust risk-sensitive filter [70] for time-triggered linear Gaussian systems evolve in similar recursive Kalman-like forms parameterized by Riccati equations. Recently, it was shown that the optimal MMSE filter of linear Gaussian systems with event-triggered scheduling followed a Kalman-like structure [23, 59, 77], which motivates us to investigate if this structural similarity of the filters for time-triggered and event-triggered scenarios holds for risk-sensitive state estimation. We consider linear Gaussian systems

---

[*]Parts of the results in this section were submitted to *Automatica*, 2018.

with stochastic event-triggering conditions. The robust event-triggered estimation problem is formulated by minimizing the risk-sensitive error criterion, which refers to the expectation of the exponential of the sum of the squared estimation error, thus penalizing all the higher order moments of the estimation error energy. The index of the exponential is weighted by a risk-sensitive parameter, which introduces one degree of freedom in system design to allow the desired tradeoff between optimal filtering for the nominal model case and robustness to worst-case noise and model uncertainty.

The main contributions of this work are summarized in two folds. First, we obtain the risk-sensitive event-triggered estimator using the reference measure approach. We propose a reference measure, under which, the measurements are i.i.d., also independent of the state process, then the reference measure is linked to the "real-world" measure via the Radon-Nikodym derivative. To obtain the state estimates, we reformulate the risk-sensitive cost criterion under the reference measure, then introduce the so-called information states and derive their recursive forms under the reference measure, which is parameterized by an event-trigger dependent time-varying Riccati equation. Based on these results, the closed-form expressions of the prior and posterior risk-sensitive event-triggered estimates are derived, which are shown to evolve in recursive Kalman-like forms. However, the obtained risk-sensitive prior and posterior estimates do not have a simple relationship, which is different from the Kalman filter and the MMSE event-triggered estimator. In addition, it is shown that the open-loop MMSE event-triggered estimators proposed in [23] can be recovered from the risk-sensitive event-triggered estimators as a special case when the risk-sensitive parameter tends to zero.

Second, we present the sufficient stability conditions for the proposed risk-sensitive event-triggered estimators. We show that, for each time instant, if one can verify that the solution to the time-varying Riccati equation remains bounded and satisfies predefined inequalities, the proposed prior and posterior estimators are exponentially stable. Moreover, based on the contraction analysis of the time-triggered risk-sensitive filters in [39], we obtain the second sufficient stability condition, according to which, the range of the risk-sensitive parameter and covariance of the initial state for which the proposed risk-sensitive event-triggered estimators are stable can be estimated in advance. Numerical examples are included to verify the effectiveness of the stability conditions. Comparative simulation results show that the proposed estimators indeed achieve robustness to system uncertainty compared with the optimal MMSE event-triggered estimator.

Notation: $\mathbb{R}$ denotes the set of real numbers. $\mathbb{N}$ denotes the set of nonnegative 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 $z, \xi \in \mathbb{R}^m$, z^i is the ith element of z and the event $\{\xi \leq z\} = \{\xi^i \leq z^i, i = 1, \ldots, m\}$. For a probability measure P ($\bar{\text{P}}$), we use E ($\bar{\text{E}}$) to represent the expectation operator. $\mathbf{1}\{\mathbf{A}\}$ is the indicator function of set $\mathbf{A}$. I_n is the $n \times n$ identity matrix. $\lambda_1(P)$ and $\lambda_2(P)$ are the largest eigenvalue and second largest eigenvalue of matrix P, respectively.

4.1.2   Problem formulation

First, we introduce a remote event-triggered estimation scheme. We consider a discrete-time linear time-invariant process driven by white noise on the probability space $(\Omega, \mathcal{F}, \text{P})$:

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

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 $\sigma_w > 0$. The probability distribution of w_k is described by $\Phi_w(\cdot)$. The initial value x_0 of the state is Gaussian with mean μ_0, and covariance P_0. The probability distribution function of the initial state x_0 is denoted as $\pi_0(\cdot)$. Let $\mathcal{F}_k^0 := \sigma\{x_0, \cdots, x_k\}$, and let $\mathcal{F}_k^x$ be the complete filtration generated by $\mathcal{F}_k^0$. The state information is measured by a sensor, which communicates with the state estimator through a network, and the output equation is

$$y_k = Cx_k + v_k, \tag{4.2}$$

where $v_k \in \mathbb{R}^m$ is zero-mean Gaussian with covariance $\sigma_v > 0$. The probability distribution of v_k is described by $\Phi_v(\cdot)$. In addition, x_0, w_k and v_k are uncorrelated with each other. Assume the pair (C, A) is observable. Let $\mathcal{F}_k^y$ be the completion of the σ-field on Ω generated by y_0, y_1, ..., y_k. At each time instant k, the sensor produces a measurement y_k and the scheduler of the sensor tests a stochastic event-triggering condition

$$\Pr(\gamma_k = 0 | y_k = y) = \delta(y) = \exp\left(-\frac{1}{2}y^\top \Pi y\right), \tag{4.3}$$

where $\Pi > 0$ and γ_k is a binary-valued event-trigger indicating whether a sensor transmission of y_k is performed or not at time instant k, with $\gamma_k = 0$ representing no transmission performed and $\gamma_k = 1$ otherwise. Let $\mathcal{F}_k^\gamma$ be the completion of the σ-field generated by $\gamma_1, \ldots, \gamma_k$. Let $\mathcal{I}_0 = \{\gamma_0, \gamma_0 y_0\}$ and for $k \in \mathbb{N}$, $\mathcal{I}_{k+1} = \mathcal{I}_k \cup \{\gamma_{k+1}, \gamma_{k+1} y_{k+1}\}$, which denotes the information available at the remote estimator.

Existing results on optimal event-triggered state estimation commonly minimize the MSE cost function, or equivalently,

$$\hat{x}_{k|k}^{MSE} = \arg\min_{\xi \in \mathbb{R}^n} \mathrm{E}\left[\tfrac{1}{2}\sum_{t=0}^{k-1}(x_t - \hat{x}_{t|t}^{MSE})^T Q_t(x_t - \hat{x}_{t|t}^{MSE}) + \tfrac{1}{2}(x_k - \xi)^T Q_k(x_k - \xi)\big|\mathcal{I}_k\right].$$

One risk-sensitive generalization of this problem is to find the risk-sensitive event-triggered posterior estimate $\hat{x}_{k|k}$ based on $\mathcal{I}_k$ such that

$$\hat{x}_{k|k} := \arg\min_{\xi \in \mathbb{R}^n} J_k(\xi), \tag{4.4}$$

where the risk-sensitive cost function is defined as

$$J_k(\xi) = \mathrm{E}\left[\exp[\theta \Psi_{0,k}(\xi)]|\mathcal{I}_k\right],$$

where $\theta > 0$ is the risk-sensitive parameter and $\Psi_{0,k}(\xi)$ is defined as

$$\Psi_{0,k}(\xi) = \hat{\Psi}_{0,k-1} + \frac{1}{2}(x_k - \xi)^T Q_k(x_k - \xi).$$

where Q_k is a symmetric, positive-definite weighting matrix for $k \geq 1$ and

$$\hat{\Psi}_{0,k-1} = \frac{1}{2}\sum_{t=0}^{k-1}(x_t - \hat{x}_{t|t})^T Q_t(x_t - \hat{x}_{t|t}).$$

The main difference between the risk-sensitive approach and the MMSE approach lies in the cost function to be minimized. The risk-sensitive approach counteracts against disturbances and model-plant mismatches by considering an exponential form in the cost function, which penalizes all the higher order moments of the estimation error energy. Thus, the risk-sensitive filter is more robust due to the efforts made at the design stage; the robust property of a risk-sensitive filter was illustrated in [7]. Motivated by the existing results on the robustness of the risk-sensitive filter in the literature, we solve the event-triggered robust estimation problem by utilizing the risk-sensitive filtering approach. For the MMSE estimator, it is well known that the following two optimization tasks are equivalent,

$$\hat{x}_{k|k}^{MSE} = \arg\min_{\xi \in \mathbb{R}^n} \mathrm{E}\left[\tfrac{1}{2}\sum_{t=0}^{k-1}(x_t - \hat{x}_{t|t}^{MSE})^T Q_t(x_t - \hat{x}_{t|t}^{MSE}) + \tfrac{1}{2}(x_k - \xi)^T Q_k(x_k - \xi)\big|\mathcal{I}_k\right]$$

$$= \arg\min_{\xi \in \mathbb{R}^n} \mathrm{E}\left[\tfrac{1}{2}(x_k - \xi)^T Q_k(x_k - \xi)|\mathcal{I}_k\right],$$

which results from the linearity property of the expectation operator. However, due to the exponential form of the risk-sensitive cost criteria, this equivalence does not hold

for the risk-sensitive estimator; it has been proved that the estimator that minimizes the exponential of a quadratic sum of the estimation errors is not equivalent to the estimator that minimizes the exponential of the current squared estimation error in the pioneering work [70] on risk-sensitive state estimation. Thus, one needs to solve the optimization problem in (4.4) in order to obtain the risk-sensitive event-triggered posterior state estimate.

Similar to the posterior estimator defined in (4.4), the risk-sensitive event-triggered prior estimate $\hat{x}_{k|k-1}$ based on $\mathcal{I}_{k-1}$ is defined as

$$\hat{x}_{k|k-1} := \arg\min_{\xi \in \mathbb{R}^n} J_k^-(\xi), \tag{4.5}$$

where the prior risk-sensitive cost function is defined as

$$J_k^-(\xi) = \mathrm{E}\big[\exp[\theta\Psi_{0,k}^-(\xi)]|\mathcal{I}_{k-1}\big],$$

with

$$\Psi_{0,k}^-(\xi) = \hat{\Psi}_{0,k-1}^- + \frac{1}{2}(x_k - \xi)^T Q_k(x_k - \xi),$$

$$\hat{\Psi}_{0,k-1}^- = \frac{1}{2}\sum_{t=0}^{k-1}(x_t - \hat{x}_{t|t-1})^T Q_t(x_t - \hat{x}_{t|t-1}).$$

One main goal of this work is to obtain the risk-sensitive event-triggered posterior and prior estimates of the system state based on partially observed measurements at the remote estimator. However, due to the exponential form of its cost criteria, the risk-sensitive event-triggered estimators cannot be derived by using a similar method as that for the MMSE estimator, where one usually first derives the conditional probability distribution of the state on the information available at the estimator, based on which, the MMSE estimator can be obtained. In the following subsection, we propose a reference probability measure under which, the state and observations are independent with each other and introduce a map to link the "real-world" measure to the reference measure. Then, we derive the recursive form of the information state of an augmented plant where the state includes the actual state of the system and part of the risk-sensitive cost under the reference measure conditioned on the information available at the remote estimator. Based on these results, the recursive form of the posterior and prior estimates are obtained respectively.

4.1.3 The reference measure

To solve this event-triggered risk-sensitive estimation problem, we use the reference measure approach. We now propose a new probability measure $\bar{\mathrm{P}}$ (also called the ref-

erence measure) on $(\Omega, \mathcal{F})$. We recall that the "real-world" measure P is defined on $(\Omega, \mathcal{F})$. The "real-world" measure P and the reference measure $\bar{P}$ are defined under the same sample space Ω and set of events $\mathcal{F}$. Under the reference measure $\bar{P}$, the variables satisfy the following relationships:

$$x_{k+1} = Ax_k + w_k, \tag{4.6}$$

where $w_k \in \mathbb{R}^n$ is i.i.d. and $w_k \sim N(0, \sigma_w)$. The measurement y_k is an i.i.d. variable under the reference measure $\bar{P}$ and $y_k \sim N(0, \sigma_v)$, which simplifies the estimation problem under the reference measure. We assume the following relationship under $\bar{P}$

$$v_k = y_k - Cx_k. \tag{4.7}$$

Under $\bar{P}$, the measurement is sent to the remote estimator according to the following stochastic event- triggering condition

$$\bar{P}(\gamma_k = 0|y_k = y) = \delta(y) = \exp\left(-\tfrac{1}{2}y^\top \Pi y\right). \tag{4.8}$$

For a random variable f defined on the probability space $(\Omega, \mathcal{F}, \bar{P})$, we further define $\bar{E}(f) = \int f d\bar{P}$, which is the expected value of f under the reference measure. In addition, we define

$$\bar{P}[A|B] = \frac{\bar{P}[A \cap B]}{\bar{P}[B]},$$

which is the conditional probability under the reference measure $\bar{P}$. To link the reference measure with the original "real-world" measure, we define a map from $\bar{P}$ to P:

$$\left.\frac{dP}{d\bar{P}}\right|_{\mathcal{G}_k} = \bar{\Lambda}_k, \tag{4.9}$$

where $\mathcal{G}_k$ is $\mathcal{F}_k^x \cup \mathcal{F}_k^y \cup \mathcal{F}_k^\gamma$, and the Radon-Nikodym derivative is defined as

$$\bar{\Lambda}_k = \prod_{l=0}^{k} \bar{\lambda}_l, \quad \bar{\lambda}_k = \frac{\Phi_v(y_k - Cx_k)}{\Phi_v(y_k)}. \tag{4.10}$$

The conditional Bayes theorem (Theorem 3.3 in Chapter 2 of [17]) provides a way of mapping the probability distribution under reference measure $\bar{P}$ back to the "real-world" measure P. Based on the conditional Bayes theorem, we investigate the properties of the map defined in (4.9)-(4.10) from the reference measure to the original probability measure, which is summarized in the following result.

Lemma 4.1. *If the model in (4.6)-(4.8) is mapped from the reference measure $\bar{\mathrm{P}}$ back to the "real-world" measure P via the Radon-Nikodym derivative in (4.9)-(4.10), then the obtained model satisfies the properties under measure P as stated in (4.1)-(4.3).*

Proof. According to the definition of $\bar{\lambda}_k$ and equation (4.6), we have

$$\bar{\mathrm{E}}\left[\bar{\lambda}_{k+1}|\mathcal{G}_k\right] = \bar{\mathrm{E}}\left[\frac{\Phi_v(y_{k+1}-C(Ax_k+w_k))}{\Phi_v(y_{k+1})}\bigg|\mathcal{G}_k\right].$$

Since under the reference measure $\bar{\mathrm{P}}$, y_{k+1} and x_k are independent and y_{k+1} is i.i.d. with probability distribution $N(0,\sigma_v)$, we have

$$\begin{aligned}
&\bar{\mathrm{E}}\left[\bar{\lambda}_{k+1}|\mathcal{G}_k\right]\\
&= \int_{\mathbb{R}^n}\int_{\mathbb{R}^m}\frac{\Phi_v(y_{k+1}-C(Ax_k+w_k))}{\Phi_v(y_{k+1})}\Phi_v(y_{k+1})\Phi_w(w_k)\mathrm{d}y_{k+1}\mathrm{d}w_k\\
&= \int_{\mathbb{R}^n}\int_{\mathbb{R}^m}\Phi_v(y_{k+1}-C(Ax_k+w_k))\\
&\quad\mathrm{d}(y_{k+1}-C(Ax_k+w_k))\cdot\Phi_w(w_k)\mathrm{d}w_k\\
&=1.
\end{aligned}$$

Following a similar procedure, we have

$$\bar{\mathrm{E}}\left[\bar{\lambda}_{k+1}|\mathcal{F}^x_{k+1}\cup\mathcal{F}^y_k\cup\mathcal{F}^\gamma_k\right] = 1.$$

According to the conditional Bayes theorem, we have

$$\begin{aligned}
&\mathrm{P}[x_{k+1}-Ax_k\leq z|\mathcal{G}_k]\\
&=\frac{\bar{\mathrm{E}}\left[\bar{\Lambda}_{k+1}\mathbf{1}\{x_{k+1}-Ax_k\leq z\}|\mathcal{G}_k\right]}{\bar{\mathrm{E}}\left[\bar{\Lambda}_{k+1}|\mathcal{G}_k\right]}\\
&=\frac{\bar{\Lambda}_k\bar{\mathrm{E}}\left[\bar{\lambda}_{k+1}\mathbf{1}\{x_{k+1}-Ax_k\leq z\}|\mathcal{G}_k\right]}{\bar{\Lambda}_k\bar{\mathrm{E}}\left[\bar{\lambda}_{k+1}|\mathcal{G}_k\right]}\\
&=\bar{\mathrm{E}}\left[\bar{\lambda}_{k+1}\mathbf{1}\{x_{k+1}-Ax_k\leq z\}|\mathcal{G}_k\right],
\end{aligned}$$

where we use the fact that $\bar{\mathrm{E}}\left[\bar{\lambda}_{k+1}|\mathcal{G}_k\right] = 1$. According to (4.6) and the fact that $w_k\in\mathbb{R}^n$ is i.i.d. and $w_k\sim N(0,\sigma_w)$ under the reference measure $\bar{\mathrm{P}}$, we further have

$$\begin{aligned}
&\mathrm{P}[x_{k+1}-Ax_k\leq z|\mathcal{G}_k]\\
&=\bar{\mathrm{E}}\left[\frac{\Phi_v[y_{k+1}-C(Ax_k+w_k)]}{\Phi_v(y_{k+1})}\mathbf{1}\{w_k\leq z\}\big|\mathcal{G}_k\right]\\
&=\int_{\mathbb{R}^n}\int_{\mathbb{R}^m}\frac{\Phi_v[y_{k+1}-C(Ax_k+w_k)]}{\Phi_v(y_{k+1})}\mathbf{1}\{w_k\leq z\}\Phi_v(y_{k+1})\Phi_w(w_k)\ \mathrm{d}y_{k+1}\ \mathrm{d}w_k\\
&=\int_{-\infty}^{z^1}\cdots\int_{-\infty}^{z^n}\Phi_w(w_k)\ \mathrm{d}w_k,
\end{aligned}$$

which recovers the relationship described in equation (4.1) under the "real-world" measure P. Similarly, for the measurement process, according to the conditional Bayes theorem and (4.50), we have

$$
\begin{aligned}
&P[y_k - Cx_k \leq z | \mathcal{F}_k^x \cup \mathcal{F}_{k-1}^y \cup \mathcal{F}_{k-1}^\gamma] \\
=& \bar{E}\big[\tfrac{\Phi_v(y_k - Cx_k)}{\Phi_v(y_k)} \mathbf{1}\{y_k - Cx_k \leq z\} \big| \mathcal{F}_k^x \cup \mathcal{F}_{k-1}^y \cup \mathcal{F}_{k-1}^\gamma\big] \\
=& \int_{\mathbb{R}^m} \Phi_v(y_k - Cx_k) \mathbf{1}\{y_k - Cx_k \leq z\} \, \mathrm{d}(y_k - Cx_k) \\
=& \int_{-\infty}^{z^1} \cdots \int_{-\infty}^{z^n} \Phi_v(v_k) \, \mathrm{d}v_k,
\end{aligned}
$$

which recovers the relationship described in equation (4.2) under the "real-world" measure P. For the event-trigger γ_k, we have

$$
\begin{aligned}
&P(\gamma_k = 1 | \mathcal{F}_k^x \cup \mathcal{F}_{k-1}^y \cup \mathcal{F}_{k-1}^\gamma, y_k = y) \\
=& \frac{\bar{E}(\bar{\Lambda}_k \gamma_k | \mathcal{F}_k^x \cup \mathcal{F}_{k-1}^y \cup \mathcal{F}_{k-1}^\gamma, y_k = y)}{\bar{E}(\bar{\Lambda}_k | \mathcal{F}_k^x \cup \mathcal{F}_{k-1}^y \cup \mathcal{F}_{k-1}^\gamma, y_k = y)} \\
=& \bar{E}(\gamma_k | \mathcal{F}_k^x \cup \mathcal{F}_{k-1}^y \cup \mathcal{F}_{k-1}^\gamma, y_k = y) \\
=& 1 - \delta(y).
\end{aligned}
$$

By repeated conditioning and $\{y_k = y\} \subset \{\mathcal{F}_k^x \cup \mathcal{F}_{k-1}^y \cup \mathcal{F}_{k-1}^\gamma, y_k = y\}$,

$$
\begin{aligned}
&P(\gamma_k = 1 | y_k = y) \\
=& E[\gamma_k | y_k = y] \\
=& E[E[\gamma_k | \mathcal{F}_k^x \cup \mathcal{F}_{k-1}^y \cup \mathcal{F}_{k-1}^\gamma, y_k = y] | y_k = y] \\
=& 1 - \delta(y),
\end{aligned}
$$

which recovers the event-triggered condition described in equation (4.3) and completes the proof. $\qquad \square$

4.1.4 Risk-sensitive event-triggered state estimates

In this subsection, we reformulate the estimation problem under the reference measure $\bar{P}$ and introduce the so-called information states. Then we derive the risk-sensitive state estimates in terms of this information states, based on which, we obtain the recursive form of the event-triggered risk-sensitive state estimates.

4.1.4.1 The information states

As stated in Section 4.1.2, our goal is to find the posterior estimate $\hat{x}_{k|k}$ and prior estimate $\hat{x}_{k|k-1}$ such that the corresponding risk-sensitive cost functions are minimized.

According to the definition of $J_k(\xi)$, the cost criterion can be reformulated under the reference measure $\bar{\mathrm{P}}$ as

$$J_k(\xi) = \mathrm{E}\left[\exp[\theta\Psi_{0,k}(\xi)]|\mathcal{I}_k\right] = \frac{\bar{\mathrm{E}}\left[\bar{\Lambda}_k \exp[\theta\Psi_{0,k}(\xi)]|\mathcal{I}_k\right]}{\bar{\mathrm{E}}\left[\bar{\Lambda}_k|\mathcal{I}_k\right]},$$

where we use Lemma 4.1 and the fact $\mathcal{I}_k$ is a sub-σ-field of $\mathcal{G}_k$ in the last equality. Since $\bar{\mathrm{E}}\left[\bar{\Lambda}_k|\mathcal{I}_k\right]$ is independent of ξ, the problem becomes to determine

$$\hat{x}_{k|k} \in \arg\min_{\xi\in\mathbb{R}^n}\bar{\mathrm{E}}\left[\bar{\Lambda}_k \exp[\theta\Psi_{0,k}(\xi)]|\mathcal{I}_k\right]. \tag{4.11}$$

Following a similar argument as above, we have

$$\hat{x}_{k|k-1} \in \arg\min_{\xi\in\mathbb{R}^n}\bar{\mathrm{E}}\left[\bar{\Lambda}_k \exp[\theta\Psi_{0,k}^-(\xi)]|\mathcal{I}_{k-1}\right]. \tag{4.12}$$

Definition 4.1 (Information States). *The information states $\alpha_k(x)$ and $\alpha_k^-(x)$ are defined as the unnormalized density functions such that*

$$\alpha_k(x)\mathrm{d}x = \bar{\mathrm{E}}[\bar{\Lambda}_{k-1}\exp\{\theta\hat{\Psi}_{0,k-1}\}\mathbf{1}(x_k \in \mathrm{d}x)|\mathcal{I}_{k-1}],$$
$$\alpha_k^-(x)\mathrm{d}x = \bar{\mathrm{E}}[\bar{\Lambda}_{k-1}\exp\{\theta\hat{\Psi}_{0,k-1}^-\}\mathbf{1}(x_k \in \mathrm{d}x)|\mathcal{I}_{k-1}].$$

Note that $\alpha_k(x)\mathrm{d}x$ can be considered as the information state of an augmented plant where the state includes the actual state of the system and part of the risk-sensitive cost [34]; so does $\alpha_k^-(x)\mathrm{d}x$. The term $\alpha_k(x)\mathrm{d}x$ is an infinitesmall displacement scaled by the unnormalized density $\alpha_k(x)$. Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is any Borel test function defined on the space $(\Omega, \mathcal{F}, \bar{\mathrm{P}})$. Based on the definition of $\alpha_k(x)$ and $\alpha_k^-(x)$, we have

$$\bar{\mathrm{E}}[f(x_k)\bar{\Lambda}_{k-1}\exp\{\theta\hat{\Psi}_{0,k-1}\}|\mathcal{I}_{k-1}] = \int_{\mathbb{R}^n} f(x)\alpha_k(x)\mathrm{d}x,$$
$$\bar{\mathrm{E}}[f(x_k)\bar{\Lambda}_{k-1}\exp\{\theta\hat{\Psi}_{0,k-1}^-\}|\mathcal{I}_{k-1}] = \int_{\mathbb{R}^n} f(x)\alpha_k^-(x)\mathrm{d}x.$$

To solve the risk-sensitive state estimation problem, it is beneficial in the literature to first derive the unnormalized conditional density functions (the information states in Definition 4.1); with the obtained iterative forms of the information states, the closed-from expressions of the risk-sensitive state estimates can be further derived, e.g., [28], [17], [14] and [7]. Before continuing, we introduce the technique of completing the square for vectors as follows:

Lemma 4.2. *[55] Suppose G is a symmetric invertible matrix. The following equation exists for the quadratic form of vector V:*

$$d + 2E^\top V + V^\top GV = (V - r)^\top M(V - r) + s,$$

where $M = G$, $r = -G^{-1}E$ and $s = d - E^\top G^{-1}E$.

Based on the above results, the recursive closed-form expressions of the information states are presented as follows.

Lemma 4.3. *For $k \geq 0$, the information states $\alpha_k(x)$ and $\alpha_k^-(x)$ are unnormalized Gaussian densities given by*

$$\alpha_k(x) = Z_k \exp\left[-\tfrac{1}{2}(x - \eta_k)^\top \Sigma_k^{-1}(x - \eta_k)\right], \tag{4.13}$$

$$\alpha_k^-(x) = Z_k^- \exp\left[-\tfrac{1}{2}(x - \eta_k^-)^\top \Sigma_k^{-1}(x - \eta_k^-)\right], \tag{4.14}$$

where Σ_k, η_k^- and η_k evolve according to the following recursions:

$$R_k = [\Sigma_k^{-1} - \theta Q_k + C^\top (\sigma_v + (1 - \gamma_k)\Pi^{-1})^{-1} C + A^\top \sigma_w^{-1} A]^{-1}, \tag{4.15}$$

$$\Sigma_{k+1} = A[\Sigma_k^{-1} - \theta Q_k + C^\top (\sigma_v + (1 - \gamma_k)\Pi^{-1})^{-1} C]^{-1} A^\top + \sigma_w, \tag{4.16}$$

$$\eta_{k+1} = \Sigma_{k+1}\sigma_w^{-1} A R_k (\Sigma_k^{-1}\eta_k + \gamma_k C^\top \sigma_v^{-1} y_k - \theta Q_k^\top \hat{x}_{k|k}), \tag{4.17}$$

$$\eta_{k+1}^- = \Sigma_{k+1}\sigma_w^{-1} A R_k (\Sigma_k^{-1}\eta_k^- + \gamma_k C^\top \sigma_v^{-1} y_k - \theta Q_k^\top \hat{x}_{k|k-1}), \tag{4.18}$$

where $\Sigma_0 = P_0$, $\eta_0^- = \mu_0$, $\eta_0 = \mu_0$ and Z_k^- and Z_k are constants and are independent of x.

Proof. We give detailed proofs for $\alpha_k(x)$, then $\alpha_k^-(x)$ can be obtained by following similar arguments. To prove Lemma 4.3, we utilize the mathematical induction. We consider the general case first and the basis when $k = 1$ can be proved by following a similar procedure as the general case, which will be shown at the end of the proof. We assume at time instant k, $\alpha_k(x)$ has the form in equation (4.13) and then we prove this form holds for $\alpha_{k+1}(x)$. Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is any Borel test function. According to the definition of $\bar{\Lambda}_k$ and $\hat{\Psi}_{0,k}$, we have

$$\begin{aligned}
&\bar{E}[f(x_{k+1})\bar{\Lambda}_k \exp(\theta\hat{\Psi}_{0,k})|\mathcal{I}_k] \\
=&\bar{E}\left[f(Ax_k + w_k)\tfrac{\Phi_v(y_k - Cx_k)}{\Phi_v(y_k)}\bar{\Lambda}_{k-1} \exp(\theta\hat{\Psi}_{0,k-1}) \exp\left[\tfrac{\theta}{2}(x_k - \hat{x}_{k|k})^\top Q_k(x_k - \hat{x}_{k|k})\right] \Big| \mathcal{I}_k\right].
\end{aligned} \tag{4.19}$$

(a) We consider the case $\gamma_k = 0$ that the event is not triggered and the measurement is not received by the remote estimator at time instant k. Since y_k is an i.i.d. variable under the reference measure $\bar{P}$, we have

$$\bar{E}[\mathbf{1}(y_k \in \mathrm{d}y)|\mathcal{I}_k] = \bar{E}[\mathbf{1}(y_k \in \mathrm{d}y)|\gamma_k, \gamma_k y_k].$$

For $\gamma_k = 0$, we have

$$\bar{E}[\mathbf{1}(y_k \in \mathrm{d}y)|\gamma_k = 0] = \tfrac{\Phi_v(y)\delta(y)\mathrm{d}y}{\int_{\mathbb{R}^m} \Phi_v(y)\delta(y)\mathrm{d}y} = \tfrac{1}{c_1}\Phi_v(y)\delta(y)\mathrm{d}y, \tag{4.20}$$

where we notice that $\int_{\mathbb{R}^m} \Phi_v(y)\delta(y)\mathrm{d}y$ is a constant, independent of y, thus denoted as c_1. For $\gamma_k = 0$, since x_k, $\hat{x}_{k|k}$ and $\bar{\Lambda}_{k-1}$ are independent of y_k and γ_k, we substitute the above equation into (4.19) and obtain

$$
\begin{aligned}
&\bar{\mathrm{E}}[f(x_{k+1})\bar{\Lambda}_k \exp(\theta\hat{\Psi}_{0,k})|\mathcal{I}_k] \\
&= \tfrac{1}{c_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} f(Az + w)\Phi_v(q - Cz)\delta(q)\mathrm{d}q \ \Phi_w(w) \\
&\quad \cdot \mathrm{d}w \exp\left[\tfrac{\theta}{2}(z - \hat{x}_{k|k})^\top Q_k(z - \hat{x}_{k|k})\right] \alpha_k(z)\mathrm{d}z \\
&= \tfrac{1}{c_1} \int_{\mathbb{R}^n} f(x) \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} \alpha_k(z)\Phi_w(x - Az)\Phi_v(q - Cz)\delta(q) \\
&\quad \cdot \exp\left[\tfrac{\theta}{2}(z - \hat{x}_{k|k})^\top Q_k(z - \hat{x}_{k|k})\right] \mathrm{d}q\mathrm{d}z\mathrm{d}x, \quad\quad\quad (4.21)
\end{aligned}
$$

where we let $x = Az + w$ to obtain (4.21). According to the definition of α_{k+1}, we have

$$
\bar{\mathrm{E}}[f(x_{k+1})\bar{\Lambda}_k \exp(\theta\hat{\Psi}_{0,k})|\mathcal{I}_k] = \int_{\mathbb{R}^n} f(x)\alpha_{k+1}(x)\mathrm{d}x.
$$

Combining (4.21) and the last equality, we obtain the recursive form of $\alpha_{k+1}(x)$ for $\gamma_k = 0$,

$$
\begin{aligned}
\alpha_{k+1}(x) =& \tfrac{1}{c_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} \Phi_w(x - Az)\alpha_k(z)\Phi_v(q - Cz)\delta(q) \\
&\cdot \exp\left[\tfrac{\theta}{2}(z - \hat{x}_{k|k})^\top Q_k(z - \hat{x}_{k|k})\right] \mathrm{d}q\mathrm{d}z.
\end{aligned}
$$

We recall that Φ_w and Φ_v follow Gaussian distributions. Substituting (4.3) into the above equation, we obtain

$$
\begin{aligned}
&\alpha_{k+1}(x) \\
&\propto \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} \exp\left[-\tfrac{1}{2}(q - Cz)^\top\sigma_v^{-1}(q - Cz) - \tfrac{1}{2}q^\top\Pi q\right] \\
&\quad \cdot \mathrm{d}q \exp\left[-\tfrac{1}{2}(x - Az)^\top\sigma_w^{-1}(x - Az) - \tfrac{1}{2}(z - \eta_k)^\top\Sigma_k^{-1}(z - \eta_k)\right. \\
&\quad \left. + \tfrac{\theta}{2}(z - \hat{x}_{k|k})^\top Q_k(z - \hat{x}_{k|k})\right]\mathrm{d}z. \quad\quad\quad (4.22)
\end{aligned}
$$

According to Lemma 4.2, we have

$$
\begin{aligned}
&\int_{\mathbb{R}^m} \exp\left[-\tfrac{1}{2}(q - Cz)^\top\sigma_v^{-1}(q - Cz) - \tfrac{1}{2}q^\top\Pi q\right] \mathrm{d}q \\
&= \int_{\mathbb{R}^m} \exp\left[-\tfrac{1}{2}(q - r_1)^\top(\sigma_v^{-1} + \Pi)(q - r_1) - \tfrac{1}{2}s_1\right] \mathrm{d}q \\
&\propto \int_{\mathbb{R}^m} \exp\left[-\tfrac{1}{2}s_1\right],
\end{aligned}
$$

with $r_1 = (\sigma_v^{-1} + \Pi)^{-1}\sigma_v^{-1}Cz$ and $s_1 = z^\top C^\top(\sigma_v + \Pi^{-1})^{-1}Cz$, where the matrix inversion lemma ([31], A.1(vii)) is used to obtain s_1. Substituting the above result into (4.22),

70

we have

$$\alpha_{k+1}(x)$$
$$\propto \int_{\mathbb{R}^n} \exp\big[-\tfrac{1}{2}z^\top C^\top(\sigma_v + \Pi^{-1})^{-1}Cz$$
$$- \tfrac{1}{2}(x - Az)^\top \sigma_w^{-1}(x - Az) - \tfrac{1}{2}(z - \eta_k)^\top \Sigma_k^{-1}(z - \eta_k)$$
$$+ \tfrac{\theta}{2}(z - \hat{x}_{k|k})^\top Q_k(z - \hat{x}_{k|k})\big]\mathrm{d}z$$
$$\propto \int_{\mathbb{R}^n} \exp\big\{ -\tfrac{1}{2}\big[(z - r_2)^\top M_2(z - r_2) + s_2\big]\big\}\mathrm{d}z.$$

The exact values of r_2 and M_2 are not necessary to be known since the integral of a Gaussian density equals to 1 no matter what values the mean and variance terms are. According to Lemma 4.2, s_2 is a function of x as follows:

$$s_2 = x^\top \sigma_w^{-1}x + \eta_k^\top \Sigma_k^{-1}\eta_k - \theta\hat{x}_k^\top Q_k\hat{x}_k - 2[\eta^\top \Sigma_k^{-1} - \theta\hat{x}_k^\top Q_k]R_kA^\top\sigma_w^{-1}x$$
$$- x^\top\sigma_w^{-1}AR_kA^\top\sigma_w^{-1}x$$
$$= (x - \eta_{k+1})^\top\Sigma_{k+1}(x - \eta_{k+1}) + s_3, \tag{4.23}$$

with $R_k = \big[A^\top\sigma_w^{-1}A + \Sigma_k^{-1} - \theta Q_k + C^\top(\sigma_v + \Pi^{-1})^{-1}C\big]^{-1}$, where Lemma 4.2 is used to obtain (4.23), s_3 is a constant independent of x and

$$\Sigma_{k+1} = A[\Sigma_k^{-1} - \theta Q_k + C^\top(\sigma_v + \Pi^{-1})^{-1}C]^{-1}A^\top + \sigma_w,$$
$$\eta_{k+1} = \Sigma_{k+1}\sigma_w^{-1}AR_k(\Sigma_k^{-1}\eta_k - \theta Q_k^\top\hat{x}_{k|k})$$

for $\gamma_k = 0$. Based on the above results, we have

$$\alpha_{k+1}(x) \propto \exp\big[-\tfrac{1}{2}(x - \eta_{k+1})^\top\Sigma_{k+1}(x - \eta_{k+1}) - \tfrac{1}{2}s_3\big].$$

Thus, for $\gamma_k = 0$, we have

$$\alpha_{k+1}(x) = Z_{k+1}\exp\big[-\tfrac{1}{2}(x - \eta_{k+1})^\top\Sigma_{k+1}(x - \eta_{k+1})\big],$$

where Z_{k+1} is a constant independent of x.

(b) Next, we consider the scenario $\gamma_k = 1$, which indicates that the measurement y_k is received by the remote estimator. Following a similar argument as that in part (a) of this proof, for $\gamma_k = 1$, we obtain

$$\alpha_{k+1}(x) = Z_{k+1}\exp\big[-\tfrac{1}{2}(x - \eta_{k+1})^\top\Sigma_{k+1}(x - \eta_{k+1})\big],$$

where Z_k is a constant independent of x and

$$R_k = [\Sigma_k^{-1} - \theta Q_k + C^\top\sigma_v^{-1}C + A^\top\sigma_w^{-1}A]^{-1},$$
$$\Sigma_{k+1} = A[\Sigma_k^{-1} - \theta Q_k + C^\top\sigma_v^{-1}C]^{-1}A^\top + \sigma_w,$$
$$\eta_{k+1} = \Sigma_{k+1}\sigma_w^{-1}AR_k(\Sigma_k^{-1}\eta_k + C^\top\sigma_v^{-1}y_k - \theta Q_k^\top\hat{x}_{k|k}).$$

For the basis $k = 1$, according to (4.13), we have

$$\bar{E}[f(x_1)\bar{\Lambda}_0 \exp(\theta\hat{\Psi}_{0,0})|\mathcal{I}_0]$$
$$=\bar{E}\left[f(Ax_0 + w_0)\frac{\Phi_v(y_0 - Cx_0)}{\Phi_v(y_0)} \cdot \exp\left\{\frac{\theta}{2}(x_0 - \hat{x}_{0|0})^T Q_0(x_0 - \hat{x}_{0|0})\right\} \Big|\mathcal{I}_0\right]$$
$$= \int_{\mathbb{R}^n} f(x)\alpha_1(x)\mathrm{d}x.$$

Recall that the initial value x_0 of the state is Gaussian with mean $\hat{x}_{0|-1} = \mu_0$, and covariance P_0. By following a similar argument as that for the general case, we have

$$\alpha_1(x) = Z_1 \exp\left[-\tfrac{1}{2}(x - \eta_1)^\top \Sigma_1 (x - \eta_1)\right],$$

where Σ_1 and η_1 follow (4.15)-(4.18) with $\Sigma_0 = P_0$ and $\eta_0 = \mu_0$; Z_1 is a constant independent of x.

For $\alpha_k^-(x)$, the results can be obtained similarly and the detailed proof is omitted due to length limitation. This completes the proof.

$\square$

The above lemma proves that the information states $\alpha_k(x)$ and $\alpha_k^-(x)$ are unnormalized Gaussian densities, where Σ_k represents the variance of the Gaussian densities. For the classical Kalman filter, the cost function to be minimized is the mean square error cost function, leading to a nice property that the Kalman filter not only obtains the posterior state estimate $\hat{x}_{k|k}^{KF}$, but also calculates the posterior error covariance matrix $P_{k|k}$ at each time instant, which can measure the estimated accuracy of the state estimate $\hat{x}_{k|k}^{KF}$. However, due to the exponential-form cost function, this nice property does not hold for the risk-sensitive state estimator. To the authors' best knowledge, Σ_k cannot be expressed as any covariance matrix related to the estimation error.

4.1.4.2 The risk-sensitive event-triggered state estimate

Based on the obtained results of the information states, we derive the risk-sensitive event-triggered posterior and prior estimtators.

Lemma 4.4. *For the event-triggered risk-sensitive problem with the event-triggering scheme in (4.3), if $\Sigma_k^{-1} - \theta Q_k \geq 0$ for $k \geq 1$, we have the following results:*
(a) the posterior estimate of the state $\hat{x}_{k|k}$ satisfies

$$\hat{x}_{k|k} = [\Sigma_k^{-1} + C^\top(\Pi^{-1} + \sigma_v)^{-1}C]^{-1}\Sigma_k^{-1}\eta_k \tag{4.24}$$

for $\gamma_k = 0$ and

$$\hat{x}_{k|k} = \eta_k + (\Sigma_k^{-1} + C^\top\sigma_v^{-1}C)^{-1}C^\top\sigma_v^{-1}(y_k - C\eta_k) \tag{4.25}$$

for $\gamma_k = 1$, where the recursion of η_k is described by (4.15)-(4.17);
(b) the prior estimate of the state $\hat{x}_{k|k-1}$ satisfies

$$\hat{x}_{k|k-1} = \eta_k^-, \tag{4.26}$$

where the recursion of η_k^- is described by (4.18).

Proof. (a) In this part, we focus on the posterior estimate $\hat{x}_{k|k}$ and present the detailed proof. First, we recall that

$$\hat{x}_{k|k} \in \arg\min_{\xi \in \mathbb{R}^n} \bar{\mathrm{E}}\left[\bar{\Lambda}_k \exp[\theta\Psi_{0,k}(\xi)]|\mathcal{I}_k\right].$$

According to the definition of $\bar{\Lambda}_k$, we have

$$\bar{\mathrm{E}}\left[\bar{\Lambda}_k \exp[\theta\Psi_{0,k}(\xi)]|\mathcal{I}_k\right]$$
$$= \bar{\mathrm{E}}\left[\frac{\Phi_v(y_k - Cx_k)}{\Phi_v(y_k)}\bar{\Lambda}_{k-1} \exp\left\{\frac{\theta}{2}(x_k - \xi)^\top Q_k(x_k - \xi)\right\} \cdot \exp(\theta\hat{\Psi}_{0,k-1})\big|\mathcal{I}_k\right].$$

1) First, we consider the scenario $\gamma_k = 0$. According to (4.20) and the definition of $\alpha_k(x)$, we have

$$\bar{\mathrm{E}}\left[\bar{\Lambda}_k \exp[\theta\Psi_{0,k}(\xi)]|\mathcal{I}_k\right]$$
$$= \frac{1}{c_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} \delta(q)\Phi_v(q - Cx) \cdot \exp\left\{\frac{\theta}{2}(x - \xi)^\top Q_k(x - \xi)\right\} \alpha_k(x)\mathrm{d}q\mathrm{d}x$$
$$\propto \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} \exp\left[-\frac{1}{2}q^\top \Pi q - \frac{1}{2}(q - cz)^\top \sigma_v^{-1}(q - cz)\right] \mathrm{d}q$$
$$\exp\left[\frac{1}{2}\theta(x - \xi)^\top Q_k(x - \xi) - \frac{1}{2}(x - \eta_k)^\top \Sigma_k^{-1}(x - \eta_k)\right]\mathrm{d}x. \tag{4.27}$$

Using Lemma 4.2, we have

$$\int_{\mathbb{R}^m} \exp\left[-\frac{1}{2}q^\top \Pi q - \frac{1}{2}(q - cz)^\top \sigma_v^{-1}(q - cz)\right] \mathrm{d}q$$
$$\propto \exp\left[-\frac{1}{2}x^\top C^\top \sigma_v^{-1}Cx + \frac{1}{2}(xC\sigma_v^{-1})^\top (\Pi + \sigma_v^{-1})^{-1}\sigma_v^{-1}Cx\right].$$

Substituting the above results into (4.27), we have

$$\bar{\mathrm{E}}\left[\bar{\Lambda}_k \exp[\theta\Psi_{0,k}(\xi)]|\mathcal{I}_k\right]$$
$$\propto \int_{\mathbb{R}^n} \exp\left[-\frac{1}{2}x^\top C^\top \sigma_v^{-1}Cx + \frac{1}{2}(xC\sigma_v^{-1})^\top (\Pi + \sigma_v^{-1})^{-1}\sigma_v^{-1}Cx\right.$$
$$\left. + \frac{\theta}{2}(x - \xi)^\top Q_k(x - \xi) - \frac{1}{2}(x - \eta_k)^\top \Sigma_k^{-1}(x - \eta_k)\right]\mathrm{d}x$$
$$\propto \int_{\mathbb{R}^n} \exp\left[-\frac{1}{2}\left\{(x - r_4)^\top M_4(x - r_4) + s_4\right\}\right]\mathrm{d}x,$$

where we use Lemma 4.2, s_4 is independent of x and

$$M_4 = (\Sigma_k^{-1} - \theta Q_k + C^\top(\Pi^{-1} + \sigma_v)^{-1}C),$$
$$s_4 = \eta_k^\top \Sigma_k^{-1}\eta_k - \theta\xi^\top Q_k\xi - (\theta\xi^\top Q_k - \eta_k^\top \Sigma_k^{-1})M_4^{-1}(\theta\xi^\top Q_k - \eta_k^\top \Sigma_k^{-1})^\top.$$

We assume the risk-sensitive parameter θ is chosen to satisfy $\Sigma_k^{-1} - \theta Q_k + C^\top(\Pi^{-1} + \sigma_v)^{-1}C > 0$. Based on some further calculations, it is easy to verify that for $\gamma_k = 0$,

$$\bar{E}\left[\bar{\Lambda}_k \exp[\theta\Psi_{0,k}(\xi)]|\mathcal{I}_k\right] \propto \exp\left[-\tfrac{1}{2}s_4\right]$$
$$\propto \exp\left[\tfrac{1}{2}\{(\xi - r_5)^\top M_5(\xi - r_5) + s_5\}\right],$$

where s_5 is independent of ξ and

$$M_5 = [(\theta Q_k)^{-1} - (\Sigma_k^{-1} + C^\top(\Pi^{-1} + \sigma_v)^{-1}C)^{-1}]^{-1},$$
$$r_5 = [\Sigma_k^{-1} + C^\top(\Pi^{-1} + \sigma_v)^{-1}C]^{-1}\Sigma_k^{-1}\eta_k.$$

Thus, if $\Sigma_k^{-1} - \theta Q_k > 0$, which definitely can guarantee $M_5 > 0$, for $\gamma_k = 0$, the risk-sensitive event-triggered posterior estimate is

$$\hat{x}_{k|k} = [\Sigma_k^{-1} + C^\top(\Pi^{-1} + \sigma_v)^{-1}C]^{-1}\Sigma_k^{-1}\eta_k.$$

2) For the scenario $\gamma_k = 1$, the measurement y_k is received by the remote estimator and we have

$$\bar{E}\left[\bar{\Lambda}_k \exp[\theta\Psi_{0,k}(\xi)]|\mathcal{I}_k\right]$$
$$= \int_{\mathbb{R}^n} \frac{\Phi_v(y_k - Cx)}{\Phi_v(y_k)} \exp\left\{\tfrac{\theta}{2}(x - \xi)^\top Q_k(x - \xi)\right\}\alpha_k(x)\mathrm{d}x$$
$$\propto \int_{\mathbb{R}^n} \exp\left[-\tfrac{1}{2}(x - \eta_k)^\top\Sigma_1(x - \eta_k) - \tfrac{1}{2}(y_k - Cx)^\top\sigma_v^{-1}(y_k - Cx)\right.$$
$$\left. + \tfrac{\theta}{2}(x - \xi)^\top Q_k(x - \xi)\right]\mathrm{d}x$$
$$\propto \int_{\mathbb{R}^n} \exp\left[-\tfrac{1}{2}\{(x - r_6)^\top M_6(x - r_6) + s_6\}\right]\mathrm{d}x,$$

with $M_6 = \Sigma_k^{-1} - \theta Q_k + C^\top\sigma_v^{-1}C$ and

$$s_6 = \eta_k^\top\Sigma_k^{-1}\eta_k + y_k\sigma_v^{-1}y_k - \theta\xi^\top Q_k\xi$$
$$- (\xi^\top\theta Q_k - \eta_k^\top\Sigma_k^{-1} - y_k^\top\sigma_v^{-1}C)M_6^{-1}(\xi^\top\theta Q_k - \eta_k^\top\Sigma_k^{-1} - y_k^\top\sigma_v^{-1}C)^\top.$$

Using Lemma 4.2 and following similar arguments as in the first part of the proof, we have that for $\gamma_k = 1$

$$\bar{E}\left[\bar{\Lambda}_k \exp[\theta\Psi_{0,k}(\xi)]|\mathcal{I}_k\right] \propto \exp\left[\tfrac{1}{2}(\xi - r_7)^\top M_7(\xi - r_7) + s_7\right],$$

where $M_7 = [(\theta Q_k)^{-1} - (\Sigma_k^{-1} + C^\top\sigma_v^{-1}C)^{-1}]^{-1}$ and

$$r_7 = \eta_k + (\Sigma_k^{-1} + C^\top\sigma_v^{-1}C)^{-1}C^\top\sigma_v^{-1}(y_k - C\eta_k).$$

74

Therefore, if $\Sigma_k^{-1} - \theta Q_k > 0$, which definitely can guarantee $M_7 > 0$, for $\gamma_k = 1$, the risk-sensitive event-triggered posterior estimate is

$$\hat{x}_{k|k} = \eta_k + (\Sigma_k^{-1} + C^\top \sigma_v^{-1} C)^{-1} C^\top \sigma_v^{-1}(y_k - C\eta_k).$$

(b) Next, we consider the prior estimate, where the problem is to determine

$$\hat{x}_{k|k-1} \in \arg\min_{\xi \in \mathbb{R}^n} \bar{\mathrm{E}}\big[\bar{\Lambda}_k \exp[\theta \Psi_{0,k}^-(\xi)]|\mathcal{I}_{k-1}\big].$$

Due to the definition of $\bar{\Lambda}_k$ and that fact that y_k is an i.i.d. variable under the reference measure $\bar{\mathrm{P}}$, we have

$$\bar{\mathrm{E}}\big[\bar{\Lambda}_k \exp[\theta \Psi_{0,k}^-(\xi)]|\mathcal{I}_{k-1}\big]$$
$$=\bar{\mathrm{E}}\big[\tfrac{\Phi_v(y_k - Cx_k)}{\Phi_v(y_k)}\bar{\Lambda}_{k-1} \exp\big\{\tfrac{\theta}{2}(x_k - \xi)^\top Q_k(x_k - \xi)\big\} \cdot \exp(\theta \hat{\Psi}_{0,k-1}^-)|\mathcal{I}_{k-1}\big]$$
$$= \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} \Phi_v(q)\tfrac{\Phi_v(q - Cx)}{\Phi_v(q)} \exp\big\{\tfrac{\theta}{2}(x - \xi)^\top Q_k(x - \xi)\big\} \cdot \alpha_k^-(x)\mathrm{d}q\mathrm{d}x. \qquad (4.28)$$

By substituting the expression of α_k^- as described in (4.14) into (4.28), we have

$$\bar{\mathrm{E}}\big[\bar{\Lambda}_k \exp[\theta \Psi_{0,k}^-(\xi)]|\mathcal{I}_{k-1}\big]$$
$$\propto \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} \exp\big[-\tfrac{1}{2}(q - Cx)^\top \sigma_v^{-1}(q - Cx)\big]\mathrm{d}q$$
$$\cdot \exp\big[\tfrac{\theta}{2}(x - \xi)^\top Q_k(x - \xi) - \tfrac{1}{2}(x - \eta_k^-)^\top \Sigma_k^{-1}(x - \eta_k^-)\big]\mathrm{d}x$$
$$\propto \int_{\mathbb{R}^n} \exp\big[-\tfrac{1}{2}\big\{(x - r_8)^\top M_8(x - r_8) + s_8\big\}\big]\mathrm{d}x,$$

where $M_8 = (\Sigma_k^{-1} - \theta Q_k)$ and

$$s_8 = \eta_k^{-\top}\Sigma_k^{-1}\eta_k^- - \theta \xi^\top Q_k \xi - (\theta \xi^\top Q_k - \eta_k^{-\top}\Sigma_k^{-1})(\Sigma_k^{-1} - \theta Q_k)^{-1}(\theta \xi^\top Q_k - \eta_k^{-\top}\Sigma_k^{-1})^\top,$$

where the risk-sensitive parameter θ needs to be chosen to satisfy $\Sigma_k^{-1} - \theta Q_k \geq 0$. Using the matrix inversion lemma, we have

$$[\theta Q_k + \theta Q_k(\Sigma_k^{-1} - \theta Q_k)^{-1}\theta Q_k^\top] = [(\theta Q_k)^{-1} - \Sigma_k]^{-1}.$$

Since $\bar{\mathrm{E}}\big[\bar{\Lambda}_k \exp[\theta \Psi_{0,k}^-(\xi)]|\mathcal{I}_{k-1}\big]$ is a function of ξ, we have

$$\bar{\mathrm{E}}\big[\bar{\Lambda}_k \exp[\theta \Psi_{0,k}^-(\xi)]|\mathcal{I}_{k-1}\big] \propto \exp\big[-\tfrac{1}{2}s_8\big]$$
$$\propto \exp\big[\tfrac{1}{2}\big\{(\xi - r_9)^\top[(\theta Q_k)^{-1} - \Sigma_k]^{-1}(\xi - r_9) + s_9\big\}\big],$$

where we use Lemma 4.2 and s_9 is independent of ξ. By some math manipulations, we have

$$r_9 = [(\theta Q_k)^{-1} - \Sigma_k][\eta_k^{-\top}\Sigma_k^{-1}(\Sigma_k^{-1} - \theta Q_k)^{-1}\theta Q_k^\top]^\top$$
$$= [(\theta Q_k)^{-1} - \Sigma_k]\theta Q_k[\Sigma_k + \Sigma_k((\theta Q_k)^{-1} - \Sigma_k)^{-1}\Sigma_k]\Sigma_k^{-1}\eta_k^- = \eta_k^-.$$

75

Thus, if $(\theta Q_k)^{-1} - \Sigma_k > 0$ for $k \geq 1$, the risk-sensitive prior estimate of the state $\hat{x}_{k|k-1}$ satisfies

$$\hat{x}_{k|k-1} = \eta_k^-,$$

which completes the proof. □

The above lemma gives the risk-sensitive event-triggered estimates, which depend on the mean and variance of the density of the information states. Combining (4.17), (4.18) and the above results, we obtain the recursive forms of risk-sensitive state estimates, which are summarized as follows.

Theorem 4.1. *For the event-triggered risk-sensitive problem with the event-triggering scheme in (4.3), if $\Sigma_k^{-1} - \theta Q_k \geq 0$ for $k \geq 1$, the risk-sensitive state estimates obey the following recursive form:*
(a) The posterior estimate $\hat{x}_{k|k}$ defined in (4.4) satisfies

$$\hat{x}_{k|k} = (I_n - K_k C)\eta_k + \gamma_k K_k y_k, \tag{4.29}$$

where

$$\eta_{k+1} = A\hat{x}_{k|k}, \tag{4.30}$$
$$K_k = \Sigma_k C^\top (C\Sigma_k C^\top + \sigma_v + (1 - \gamma_k)\Pi^{-1})^{-1} \tag{4.31}$$

and Σ_k evolves according to the following time-varying Riccati equation

$$\Sigma_{k+1} = A[\Sigma_k^{-1} - \theta Q_k + C^\top(\sigma_v + (1 - \gamma_k)\Pi^{-1})^{-1}C]^{-1}A^\top + \sigma_w. \tag{4.32}$$

(b) The prior estimate $\hat{x}_{k|k-1}$ defined in (4.5) satisfies

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k-1} + K_k^-(\gamma_k y_k - C\hat{x}_{k|k-1}), \tag{4.33}$$

where

$$K_k^- = A[\Sigma_k^{-1} - \theta Q_k + C^\top(\sigma_v + (1 - \gamma_k)\Pi^{-1})^{-1}C]^{-1}C^\top(\sigma_v + (1 - \gamma_k)\Pi^{-1})^{-1}. \tag{4.34}$$

Proof. (a) We first focus on the proof for the posterior estimate. For $\gamma_k = 0$, based on (4.17) and (4.24), it is trivial to obtain that

$$\eta_{k+1} = \Sigma_{k+1}\sigma_w^{-1}AR_{k+1}(\Sigma_k^{-1} + C^\top(\sigma_v + \Pi^{-1})^{-1}C - \theta Q_k^\top)\hat{x}_{k|k}.$$

According to the definition of R_{k+1} and using the matrix inversion lemma, we can verify that

$$\eta_{k+1} = A\hat{x}_{k|k}.$$

Based on some further calculations, for $\gamma_k = 0$, we have

$$\hat{x}_{k|k} = (I_n - K_k C)\eta_k,$$

with $K_k = \Sigma_k C^\top (C\Sigma_k C^\top + \sigma_v + \Pi^{-1})^{-1}$.

For $\gamma_k = 1$, combining (4.17) and (4.25), we have

$$\eta_k = (I_n + \Sigma_k C^\top \sigma_v^{-1} C)\hat{x}_{k|k} - \Sigma_k C^\top \sigma_v^{-1} y_k.$$

Substituting the above result and (4.15) into (4.17), we can easily obtain that

$$\eta_{k+1} = \Sigma_{k+1}\sigma_w^{-1} AR_{k+1}(\Sigma_k^{-1} + C^\top \sigma_v^{-1} C - \theta Q_k^\top)\hat{x}_{k|k}$$
$$= A\hat{x}_{k|k}.$$

Accordingly, for $\gamma_k = 1$, the posterior estimate follows

$$\hat{x}_{k|k} = (I_n - K_k C)\eta_k + K_k y_k,$$

where $K_k = \Sigma_k C^\top (C\Sigma_k C^\top + \sigma_v)^{-1}$. Thus, the posterior estimate follows (4.29).

(b) Next, we consider the prior estimate. According to (4.18) and (4.26), we substitute in Σ_{k+1} and R_k, then have

$$\eta_{k+1}^- = \Sigma_{k+1}\sigma_w^{-1} AR_k(\Sigma_k^{-1}\eta_k^- + \gamma_k C^\top \sigma_v^{-1} y_k - \theta Q_k^\top \eta_k^-)$$
$$= A[\Sigma_k^{-1} - \theta Q_k + C^\top (\sigma_v + (1-\gamma_k)\Pi^{-1})^{-1} C]^{-1}[(\Sigma_k^{-1} - \theta Q_k)\eta_k^- + \gamma_k C^\top \sigma_v^{-1} y_k].$$

For notational simplicity, we denote $O_k = [\Sigma_k^{-1} - \theta Q_k + C^\top (\sigma_v + (1-\gamma_k)\Pi^{-1})^{-1} C]^{-1}$. Using the matrix inversion lemma, we have

$$O_k = [\Sigma_k^{-1} - \theta Q_k + C^\top (\sigma_v + (1-\gamma_k)\Pi^{-1})^{-1} C]^{-1}$$
$$= (\Sigma_k^{-1} - \theta Q_k)^{-1} - (\Sigma_k^{-1} - \theta Q_k)^{-1} C^\top$$
$$\cdot [\sigma_v + (1-\gamma_k)\Pi^{-1} + C(\Sigma_k^{-1} - \theta Q_k)^{-1} C^\top]^{-1} C(\Sigma_k^{-1} - \theta Q_k)^{-1}.$$

Based on some further calculations, we have

$$AO_k(\Sigma_k^{-1} - \theta Q_k)\eta_k^-$$
$$= A\eta_k^- - AO_k C^\top (\sigma_v + (1-\gamma_k)\Pi^{-1})^{-1} C\eta_k^-.$$

Then we have

$$
\begin{aligned}
\eta_{k+1}^- =& A\eta_k^- - AO_kC^\top(\sigma_v + (1-\gamma_k)\Pi^{-1})^{-1}C\eta_k^- \\
&+ AO_kC^\top(\sigma_v + (1-\gamma_k)\Pi^{-1})^{-1}\gamma_k y_k \\
=& A\eta_k^- + K_k(\gamma_k y_k - C\eta_k^-).
\end{aligned}
$$

Recall that $\hat{x}_{k|k-1} = \eta_k^-$ as stated in Theorem 4.4. This completes the proof. $\qquad\square$

Remark 4.1. The above theorem presents the close-form expressions of the event-triggered estimators, which are parameterized by the time-varying Riccati equation in (4.32). Based on (4.31), (4.32) and (4.34), we can easily show that the gains K_k and K_k^- of the risk-sensitive event-triggered filter, which is supposed to be more robust to system uncertainties, are larger compared with those of the MMSE estimator with stochastic event-triggering conditions in [23] since the term θQ_k increases the observer gains. This is consistent with the results obtained from high-gain observers [1], which are shown to have the ability to reject modeling disturbance.

Next, we compare the risk-sensitive filter proposed in Theorem 4.1 with the well-known Kalman filter, for which the following relationship holds between the prior state estimate $\hat{x}_{k|k-1}^{KF}$ and the posterior estimate $\hat{x}_{k|k}^{KF}$:

$$
\begin{aligned}
\hat{x}_{k|k-1}^{KF} =& A\hat{x}_{k-1|k-1}^{KF}, \\
\hat{x}_{k|k}^{KF} =& \hat{x}_{k|k-1}^{KF} + K_k^{KF}(y_k - C\hat{x}_{k|k-1}^{KF}),
\end{aligned}
$$

where K_k^{KF} is the Kalman gain. The prior estimate $\hat{x}_{k|k-1}^{KF}$ can be easily obtained based on the posterior estimate $\hat{x}_{k-1|k-1}^{KF}$ at time instant $k-1$, then the posterior estimate $\hat{x}_{k|k}^{KF}$ can be calculated based on the prior estimate $\hat{x}_{k|k-1}^{KF}$, which indicates a simple relationship between the prior and posterior estimates of the Kalman filter. Similar relationship holds for the prior and posterior estimates of the event-triggered MMSE estimator in [23] as well. According to Theorem 6, for the risk-sensitive state estimator, the posterior estimate $\hat{x}_{k|k}$ evolves according to equations (4.29)-(4.32) and the prior estimate $\hat{x}_{k|k-1}$ evolves according to equations (4.33)-(4.34), thus the simple relationship between the Kalman estimates $\hat{x}_{k|k-1}^{KF}$ and $\hat{x}_{k|k}^{KF}$ does not hold for the risk-sensitive estimates $\hat{x}_{k|k-1}$ and $\hat{x}_{k|k}$ in this work.

As a special case, when the risk-sensitive parameter tends to zero, our proposed risk-sensitive estimators recover the open-loop MMSE event-triggered estimators proposed in [23], which sacrifice the robustness to uncertainties for optimal filtering under the nominal case. For the proposed risk-sensitive estimators, users can achieve desired

tradeoff between the optimality and robustness by tuning the risk-sensitive parameter. The risk-sensitive estimators are more robust as the risk-sensitive parameter becomes relatively larger.

4.1.5 Sufficient stability conditions

In this subsection, we study the stability issue of the event-triggered risk-sensitive posterior and prior estimators proposed in Theorem 4.1 and give the range of values of the risk-sensitive parameter and covariance of the initial state for which Σ_k is bounded and the proposed estimators are exponentially stable. We assume the weighting matrix $Q_k = Q > 0$, then there exists $Q = F^\top F$ with $F \in \mathbb{R}^{n \times n}$ and $F > 0$. For the variance of the process noise, we have $\sigma_w = BB^\top$ with $B \in \mathbb{R}^{n \times n}$ and $B > 0$.

Before we show the main stability results for the risk-sensitive event-triggered estimators, we present a lemma for one special case that no transmission is performed ($\gamma_k = 0$) for all $k \geq 0$. In this ideal case, the time-varying Riccati equation proposed in (4.32) becomes

$$\tilde{\Sigma}_{k+1} = A[\tilde{\Sigma}_k^{-1} - \theta Q + C^\top(\sigma_v + \Pi^{-1})^{-1}C]^{-1}A^\top + \sigma_w, \tag{4.35}$$

which is an algebraic Riccati equation at steady state. The steady-state characterization of $\tilde{\Sigma}_k$, in which $k \to \infty$, can be found by iterating the dynamic equation repeatedly until it converges; in other words, the steady state $\tilde{\Sigma}$ is characterized by removing the time subscripts from the dynamic equation (4.35), thus resulting in an algebraic Riccati equation. Since $(\sigma_v + \Pi^{-1})^{-1} > 0$, there exists $(\sigma_v + \Pi^{-1})^{-1} = D^\top D$ with $D \in \mathbb{R}^{m \times m}$ and $D > 0$. Next, we introduce parameters that will be used in the following lemma. For $N \geq n$, $\mathcal{H}_N$ is the $Nm \times Nn$ block Toeplitz matrix defined by

$$\mathcal{H}_N = \begin{bmatrix} 0 & H_1 & H_2 & \cdots & H_{N-1} \\ 0 & 0 & H_1 & \cdots & H_{N-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & H_1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix},$$

where

$$H_t = \begin{cases} CA^{t-1}B, & t \geq 1, \\ 0, & \text{otherwise.} \end{cases}$$

The N-step observability matrices of the pair (C, A) and (F, A) are denoted as

$$O_N = \begin{bmatrix} (CA^{N-1})^\top & \cdots & (CA)^\top & C^\top \end{bmatrix}^\top,$$
$$O_N^R = \begin{bmatrix} (FA^{N-1})^\top & \cdots & (FA)^\top & F^\top \end{bmatrix}^\top.$$

For $N \geq n$, $\mathcal{L}_N$ is the N block Toeplitz matrix taking the form

$$\mathcal{L}_N = \begin{bmatrix} 0 & L_1 & L_2 & \cdots & L_{N-1} \\ 0 & 0 & L_1 & \cdots & L_{N-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & L_1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix},$$

where

$$L_t = \begin{cases} FA^{t-1}B, & t \geq 1, \\ 0, & \text{otherwise.} \end{cases}$$

We define the $Nm \times Nm$ diagonal matrix

$$\Xi = \mathrm{diag}[\sigma_v + \Pi^{-1}, \ldots, \sigma_v + \Pi^{-1}]$$

The risk-sensitive Gramian Ω_N^θ is defined as

$$\Omega_N^\theta = O_N^\top [\Xi + \mathcal{H}_N \mathcal{H}_N^\top]^{-1} O_N + J_N^\top (S_N^\theta)^{-1} J_N, \tag{4.36}$$

where

$$J_N \triangleq O_N^R - \mathcal{L}_N \mathcal{H}_N^\top (\Xi + \mathcal{H}_N \mathcal{H}_N^\top)^{-1} O_N,$$
$$S_N^\theta \triangleq -\theta^{-1} I_{Nn} + \mathcal{L}_N (I_{Nn} + \mathcal{H}_N^\top \mathcal{H}_N)^{-1} \mathcal{L}_N^\top.$$

Let $\tau_N < \theta_N$ be the first value of θ for which Ω_N^θ becomes singular, where

$$\theta_N = \frac{1}{\lambda_1 (\mathcal{L}_N (I_{Nn} + \mathcal{H}_N^\top \mathcal{H}_N)^{-1}) \mathcal{L}_N^\top)}. \tag{4.37}$$

Next, we choose a free matrix G to ensure that the matrix $(A - GDC)$ is stable and

$$r \triangleq \max_{0 \leq i \leq n} |\lambda_i (A - GDC)|.$$

For $1 < \rho < 1/r$, we define

$$\beta_\rho \triangleq \frac{\rho^2 - 1}{\rho^2 \lambda_1 (F \Sigma_\rho F^\top)}, \tag{4.38}$$

where

$$\Sigma_\rho = \rho^2 (A - GDC)^k \Sigma_\rho ((A - GDC))^\top + Q + GG^\top. \tag{4.39}$$

Then, we present the following convergence lemma for the Riccati equation in (4.35).

Lemma 4.5. *For the system described in* (4.1) *and* (4.2), *we consider the scenario that* $\gamma_k = 0$ *for all* $k > 0$ *(none of the sensor's measurements is transmitted to the remote estimator) and obtain the Riccati equation in* (4.35). *With* θ_N *defined in* (4.37), $\tau_N < \theta_N$ *is the first value of* θ *for which* Ω_N^θ *in* (4.36) *becomes singular. If* $0 \leq \theta < \tau_N$ *and* $\theta \leq \beta_\rho$ *with* $N \geq n$ *and* β_ρ *defined in* (4.38), *the algebraic Riccati map as defined in* (4.35) *has a unique positive definite fixed point* $\tilde{\Sigma}$ *such that* $\tilde{\Sigma}^{-1} - \theta Q \geq 0$. *Furthermore, if the initial condition* $\tilde{\Sigma}_0$ *of the Riccati equation* (4.35) *satisfies* $0 < \tilde{\Sigma}_0 \leq \Sigma_\rho$ *with* Σ_ρ *given in* (4.39), *the entire trajectory of the Riccati map satisfies* $0 < \tilde{\Sigma}_k \leq \Sigma_\rho$ *and* $\tilde{\Sigma}_k^{-1} - \theta Q > 0$ *for all* $k > 0$, *and tends to* $\tilde{\Sigma}$.

Proof. This lemma is obtained by applying Theorem 5.3 in [39], which relies on a block implementation of risk-sensitive filters. Some modifications are made to the results in [39] since the variance of the measurement noise in our formulation is not the identity matrix considered in [39]. □

The above lemma provides the ranges of the risk-sensitive parameter θ and the variance $\tilde{\Sigma}_0$ of the initial state, for which the Riccati map in (4.35) is convergent. The results presented in Lemma 4.5 relies on the contraction analysis of the risk-sensitive filters [39], which requires only the observability of the system. Thus Lemma 4.5 still holds even if the system matrix A is unstable.

Lemma 4.6. *For the Riccati equation in* (4.35), *if the covariance of the initial state* $P_0 = \tilde{\Sigma}_0$ *and the risk-sensitive parameter are chosen within the range proposed in Lemma 4.5, which implies that* $\tilde{\Sigma}_k$ *in* (4.35) *satisfies* $0 < \tilde{\Sigma}_k \leq \Sigma_\rho$ *and* $\tilde{\Sigma}_k^{-1} - \theta Q > 0$ *for all* $k \geq 0$, *the following inequalities hold for the solution of the risk-sensitive Riccati equation* (4.32)

$$\Sigma_k \leq \Sigma_\rho, \ \Sigma_k^{-1} - \theta Q > 0.$$

Proof. We prove this lemma by induction. We first assume that $\Sigma_{k-1} \leq \tilde{\Sigma}_{k-1}$, then prove that $\Sigma_k \leq \tilde{\Sigma}_k$. Since $0 < \tilde{\Sigma}_{k-1}^{-1} - \theta Q$ and $\Sigma_{k-1} \leq \tilde{\Sigma}_{k-1}$, we have $0 \leq \tilde{\Sigma}_{k-1}^{-1} - \theta Q \leq \Sigma_{k-1}^{-1} - \theta Q$. Recalling that the transmission decision variable γ_k can be either 0 or 1, we have

$$0 < \tilde{\Sigma}_{k-1}^{-1} - \theta Q + C^\top (\sigma_v + \Pi^{-1})^{-1} C \leq$$
$$\Sigma_{k-1}^{-1} - \theta Q + C^\top (\sigma_v + (1 - \gamma_k)\Pi^{-1})^{-1} C.$$

Furthermore, we have

$$\sigma_w + A[\Sigma_{k-1}^{-1} - \theta Q + C^\top(\sigma_v + (1 - \gamma_k)\Pi^{-1})^{-1}C]^{-1}A^\top$$
$$\leq \sigma_w + A[\tilde{\Sigma}_{k-1}^{-1} - \theta Q + C^\top(\sigma_v + \Pi^{-1})^{-1}C]^{-1}A^\top,$$

which is $\Sigma_k \leq \tilde{\Sigma}_k$. Since $\tilde{\Sigma}_k \leq \Sigma_\rho$, we have $\Sigma_k \leq \Sigma_\rho$. For the basis, since $\Sigma_0 = P_0 = \tilde{\Sigma}_0$ satisfies $\Sigma_0 \leq \tilde{\Sigma}_0$. This completes the proof. $\qquad\square$

Lemma 4.7 ([3], Theorem 4.3). *Suppose that $[\hat{F}_k, \hat{G}_k]$ is uniformly stabilizable and that $\hat{F}_k$ and $\hat{G}_k$ are bounded. If there is a bounded nonnegative definite matrix sequence $\hat{P}_k$ satisfying*

$$\hat{P}_{k+1} = \hat{F}_k \hat{P}_k \hat{F}_k^\top + \hat{G}_k \hat{G}_k^\top,$$

for $k \geq 0$. Then $z_{k+1} = \hat{F}_k z_k$ is exponentially stable.

Lemma 4.8 ([76], Theorem 3.1). *An exponentially stable linear dynamic system $p_{k+1} = \hat{D}_k p_k + \hat{H}_k u_k$ is bounded-input, bounded-output (BIBO) stable if $\hat{H}_k$ is uniformly bounded.*

Based on above results, we are now ready to present the stability results on the posterior and prior estimators with the time-varying Riccati equation in (4.32).

Theorem 4.2. *For the system described in (4.1) and (4.2) with the event-triggering condition in (4.3), the event-triggered risk-sensitive posterior estimator in (4.29) and prior estimator in (4.33) are exponentially stable if either one of the following conditions holds:*

(a) there exists a sequence of solution Σ_k to the Riccati map (4.32) satisfying $\Sigma_k > 0$ and $\Sigma_k^{-1} - \theta Q > 0$ for all $k > 0$;

(b) the first condition to be satisfied is that $0 \leq \theta < \tau_N$ and $\theta \leq \beta_\rho$, where τ_N and β_ρ can be calculated according to Lemma 4.5; the second condition is that the covariance of the initial state P_0 satisfies $0 < P_0 \leq \Sigma_\rho$.

Proof. We prove the stability of the risk-sensitive estimators using the Lyapunov approach for linear time-varying systems introduced in Lemma 4.7 .

First, we prove (a) of the above theorem. For the posterior estimator in (4.29), we have

$$\eta_{k+1} = A\hat{x}_{k|k} = A(I_n - K_k C)\eta_k + \gamma_k A K_k y_k. \tag{4.40}$$

For notational simplicity, we define $\Gamma_k = (\Sigma_k^{-1} + C^\top(\sigma_v + (1 - \gamma_k)\Pi^{-1})^{-1}C)^{-1}$. By the matrix inversion lemma, the Riccati equation in (4.32) can be equivalently written as

$$\Sigma_{k+1} = A[\Gamma_k + \Gamma_k((\theta Q)^{-1} - \Gamma_k)^{-1}\Gamma_k]A^\top + \sigma_w.$$

Based on some further calculations, we have

$$A\Gamma_k A^\top = A(I_n - K_k C)\Sigma_k(I_n - K_k C)^\top A^\top + AK_k(\sigma_v + (1 - \gamma_k)\Pi^{-1})K_k^\top A^\top.$$

Then we obtain

$$\Sigma_{k+1} = A(I_n - K_k C)\Sigma_k(I_n - K_k C)^\top A^\top + N_k, \tag{4.41}$$

where

$$N_k = AK_k(\sigma_v + (1 - \gamma_k)\Pi^{-1})K_k^\top A^\top + A\Gamma_k((\theta Q)^{-1} - \Gamma_k)^{-1}\Gamma_k A^\top + \sigma_w.$$

Since $\Sigma_k^{-1} - \theta Q > 0$, it is easy to verify $(\theta Q)^{-1} - \Gamma_k > 0$, which indicates $N_k > 0$ and $N_k = G_k G_k^\top$. Applying Lemma 4.7 to (4.41), we obtain that $\eta_{k+1} = A(I_n - K_k C)\eta_k + \gamma_k AK_k y_k$ is exponentially stable. Since $(I_n - K_k C)$ is uniformly bounded from above, we apply Lemma 4.8 on the prior estimator in (4.29) and conclude that the risk-sensitive posterior estimator is exponentially stable.

From (4.33), the prior estimator follows

$$\hat{x}_{k+1|k} = (A - K_k^- C)\hat{x}_{k|k-1} + \gamma_k K_k^- y_k.$$

According to the definition of K_k^- and by some matrix manipulations, we have

$$A[\Sigma_k^{-1} - \theta Q + C^\top(\sigma_v + (1 - \gamma_k)\Pi^{-1})^{-1}C]^{-1}A^\top$$
$$= (A - K_k^- C)[\Sigma_k^{-1} - \theta Q]^{-1}(A - K_k^- C)^\top + K_k^-(\sigma_v + (1 - \gamma_k)\Pi^{-1})^{-1}K_k^{-\top}$$

Substituting the above results into the risk-sensitive Riccati equation in (4.32), we have

$$\Sigma_{k+1}$$
$$= A[\Sigma_k^{-1} - \theta Q + C^\top(\sigma_v + (1 - \gamma_k)\Pi^{-1})^{-1}C]^{-1}A^\top + \sigma_w$$
$$= (A - K_k^- C)(\Sigma_k^{-1} - \theta Q)^{-1}(A - K_k^- C)^\top + M_k,$$

where $M_k = K_k^-(\sigma_v + (1 - \gamma_k)\Pi^{-1})^{-1}K_k^{-\top} + \sigma_w$. By adding $(\Sigma_{k+1}^{-1} - \theta Q)^{-1}$ to both sides of the above equation and moving Σ_{k+1} to the right side, we have

$$(\Sigma_{k+1}^{-1} - \theta Q)^{-1}$$
$$= (A - K_k^- C)(\Sigma_k^{-1} - \theta Q)^{-1}(A - K_k^- C)^\top + M_k + (\Sigma_{k+1}^{-1} - \theta Q)^{-1} - \Sigma_{k+1}. \tag{4.42}$$

Since $\Sigma_{k+1}^{-1} > \Sigma_{k+1}^{-1} - \theta Q > 0$, we have $(\Sigma_{k+1}^{-1} - \theta Q)^{-1} - \Sigma_{k+1} > 0$. Also, due to that $K_k^-(\sigma_v + (1 - \gamma_k)\Pi^{-1})^{-1}K_k^{-\top} \geq 0$ and $\sigma_w > 0$, we have

$$M_k + (\Sigma_{k+1}^{-1} - \theta Q)^{-1} - \Sigma_{k+1} > 0,$$

indicating that $M_k + (\Sigma_{k+1}^{-1} - \theta Q)^{-1} - \Sigma_{k+1} = G_k^- G_k^{-\top}$, where $G_k^- \in \mathbb{R}^{n \times n}$ and G_k^- has full column rank. Furthermore, $[A - K_k^- C, G_k^-]$ is stabilizable. Since $\Sigma_k \leq \Sigma_\rho$, we have $(\Sigma_k^{-1} - \theta Q)^{-1} \leq (\Sigma_\rho^{-1} - \theta Q)^{-1}$, indicating that $(\Sigma_k^{-1} - \theta Q)^{-1}$ is bounded. Similarly, we can prove that $(A - K_k^- C)$ and $M_k + (\Sigma_{k+1}^{-1} - \theta Q)^{-1} - \Sigma_{k+1}$ are bounded. Thus, combining (4.42) and using Lemma 4.7, the risk-senstive event-triggered prior estimator in (4.33) is exponentially stable.

For part (b) of the above theorem, based on Lemmas 4.5 and 4.6, we can easily prove the boundedness of Σ_k and $\Sigma_k^{-1} - \theta Q > 0$ for all $k > 0$ within the range of values of θ and P_0 given in the second condition of the above theorem. Then, based on the obtained results in part (a) of the above theorem, we conclude that the risk-senstive event-triggered posterior estimator in (4.29) and prior estimator in (4.33) are exponentially stable, which completes the proof. □

Remark 4.2. The sufficient stability condition proposed in part (a) of Theorem 4.2 can be extended to the case when the weighting matrix Q_k is time-variant. The ranges of the risk-sensitive parameter and covariance of the initial state can be estimated in a prior based on part (b) of Theorem 4.2, which is easy to be examined and provides a reasonable choice of the risk-sensitive parameter for users, though is relatively conservative compared with part (a) of Theorem 4.2.

4.1.6 Simulation examples

In this subsection, we show the effectiveness and robustness of the proposed risk-sensitive event-triggered (RSET) prior estimator by simulation examples. The simulation results for the RSET posterior estimator are similar to that of the RSET prior estimator and are omitted. We consider a second-order system which is described by

$$x_{k+1} = (A + \Delta A)x_k + w_k,$$

where the nominal model $A = [0.99 \; 0.01; 0 \; 0.99]$ with $\Delta A = [0 \; \varsigma; 0 \; 0]$ representing the model uncertainty and $\varsigma = -0.05$. The measurement process is described by (4.2) with $C = [1 \; -1]$. The process noise w_k and measurement noise v_k are zero-mean Gaussians with covariances $\sigma_w = [0.5 \; 0; 0 \; 0.5]$ and $\sigma_v = 0.3$. The initial value x_0 of the state is
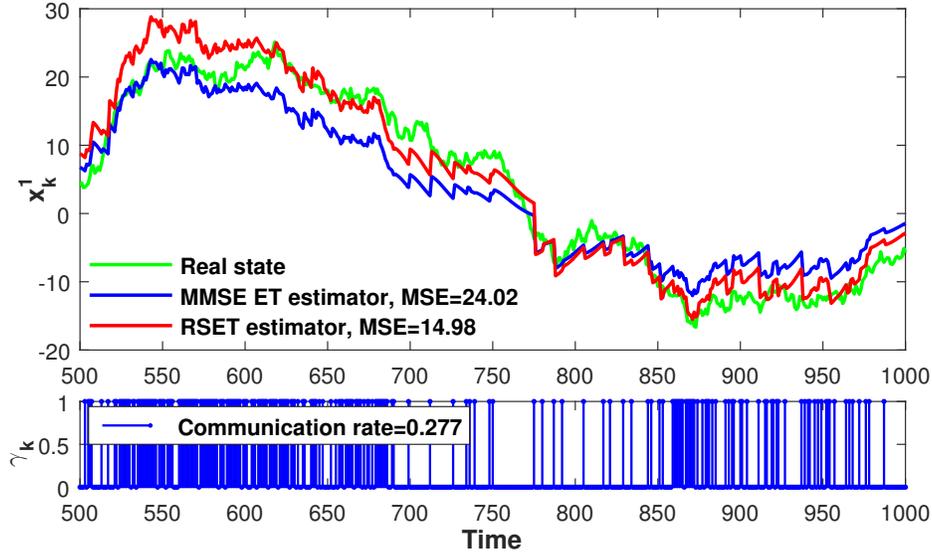
Figure 4.1: Performance comparison between the proposed RSET prior estimator and the MMSE ET prior estimator for $\varsigma = -0.05$.

Gaussian distributed with mean $\hat{x}_0^- = \mu_0 = [0\ 0]^\top$ and covariance $P_0 = [0.5\ 0; 0\ 0.5]$. Sensor measurements will be transmitted to the remote estimator whenever the event-triggering condition in (4.3) is satisfied, where we choose $\Pi = 0.005$. Also, there exists $(\sigma_v + \Pi^{-1})^{-1} = D^\top D$ where $D = 14.1527$. The risk-sensitive estimator is obtained according to (4.5), and we choose the weighting parameter to be constant, namely, $Q_k = Q = [1\ 0; 0\ 1]$.

First, we show the robustness of the proposed estimator to system uncertainty. We choose the same risk-sensitive parameter $\theta = 0.0001$ and verify that part (a) of Theorem 4.2 is satisfied for finite horizon $T = 1000$. The performance comparison of the RSET estimator obtained from time instant 500 to 1000 and the original MMSE event-triggered (ET) prior estimator in [23] obtained using the same communication sequence is shown in Fig. 4.1, where both the estimators are obtained based on the nominal model $x_{k+1} = Ax_k + w_k$. The results in Fig. 4.1 start from time instant 500 since it takes some time for the model uncertainty to be fully reflected in the system dynamics. The vertical axis of Fig. 4.1 represents the value of the first element of the system state, denoted as x_k^1. The resultant MSE for the proposed RSET estimator and the MMSE ET estimator are 14.98 and 24.02, respectively, indicating that the proposed RSET estimator outperforms the MMSE ET estimator when uncertainty exists and $\Delta A = [0\ -0.05; 0\ 0]$.

To evaluate the estimation performance of the proposed RSET estimator properly,

85
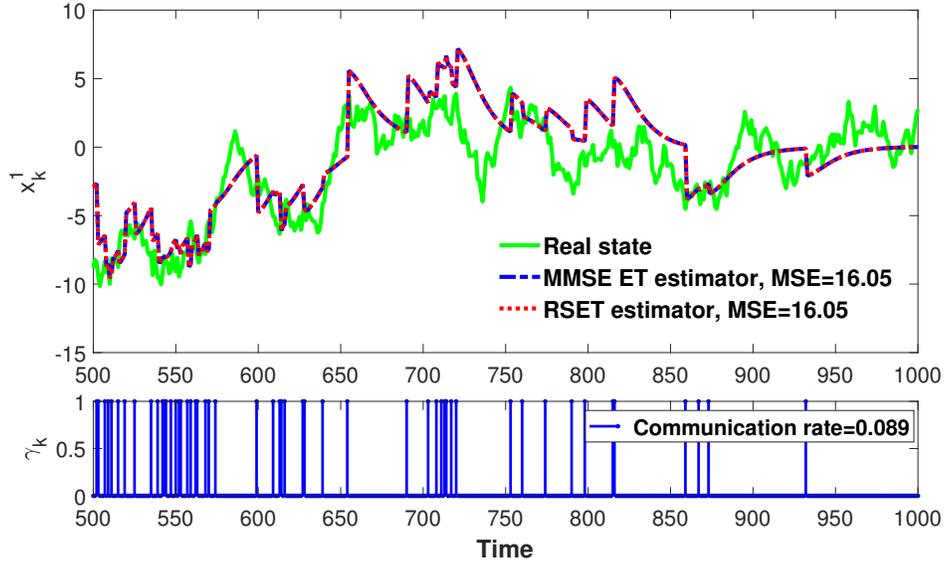
Figure 4.2: Performance comparison between the proposed REST prior estimator and the MMSE ET prior estimator for the scenario with no parameter uncertainty.

we consider the scenario without model uncertainty, namely, $\Delta A = [0\ 0; 0\ 0]$, where the other parameters remain exactly the same as in Fig. 4.1. We choose the same risk-sensitive parameter $\theta = 0.0001$ and verify that part (a) of Theorem 4.2 is satisfied for finite horizon $T = 1000$. The performance comparison of the RSET estimator obtained from time instant 500 to 1000 and the original MMSE event-triggered (ET) prior estimator in [23] obtained using the same communication sequence is shown in Fig. 4.2. The estimates of the proposed RSET estimator and the MMSE ET estimator almost coincide in Fig. 4.2 and the resultant MSEs for the two methods are very close – one cannot tell a difference in accuracy to the second decimal place. This phenomenon results from the specific system considered in this example and the relatively small risk-sensitive parameter. We can conclude that for the scenario without system uncertainty, the performance of the proposed RSET estimator is acceptable compared with the MMSE ET estimator.

To further show the merits of the proposed method, the variation of the MSEs for the RSET estimator and the MMSE ET estimator as the uncertainty parameter ς in ΔA ranges in value from -0.2 to 0.15, where $\theta = 0.00008$ and the other parameters remain the same, is shown in Fig. 4.3. Each point in Fig. 4.3 is obtained for finite horizon $T = 150000$, within which period, the condition in part (a) of Theorem 4.2 is verified to be satisfied. We notice that for the case that the uncertainty parameter ς

86
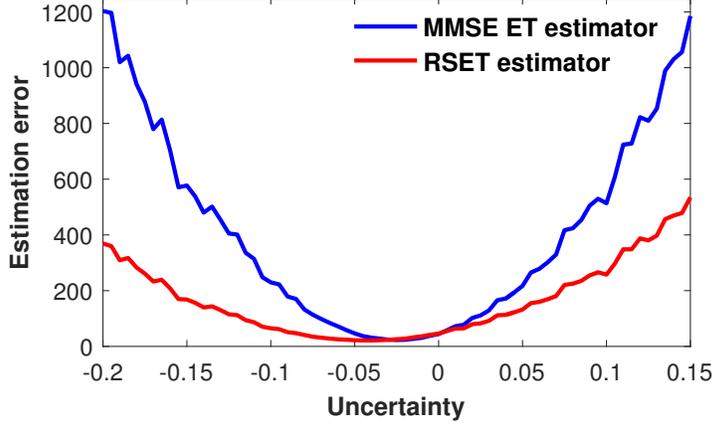
Figure 4.3: MSE comparison between the RSET and MMSE ET prior estimators with respect to system uncertainty.

tends to zero, the MMSE ET estimator and RSET estimator have similar estimation performance, sometimes the MMSE ET estimator even achieves a lower value of MSE; while for a larger absolute value of ς (larger system uncertainty), the RSET estimator outperforms the MMSE ET estimator. This indicates that the proposed estimator has good estimation performance under nominal scenarios, and acceptable estimation performance that degrades less rapidly than the MMSE ET estimator under non-nominal scenarios, so we conclude that the proposed estimator is more robust to system uncertainty compared with the MMSE ET estimator. In addition, we notice that Fig. 4.3 is not centered exactly around 0. The explanation of this phenomenon is that as the uncertainty parameter ς in ΔA ranges in value from -0.2 to 0.15, the system matrix $(A + \Delta A)$ varies accordingly whereas the nominal model A utilized by the RSET estimator and the MMSE ET estimator does not change. In other words, the true models of the system are different when ς takes different values, therefore only the estimation errors of the two estimators for a certain ς are comparable, not the estimation errors for different ς. The specific system model and parameters considered will determine the center of the figure and at what value of ς the smallest estimation error will be achieved.

Next, the tradeoff between communication rates and estimation performances for the RSET prior estimator and MMSE ET prior estimator under different event-triggering conditions is shown in Fig. 4.4, where $\theta = 0.00005$, $\varsigma = -0.04$, the value of Π in (4.3) ranges from 0.0002 to 0.362 and the other parameters remain the same as that in Fig. 4.1. The system runs for finite horizon $T = 160000$ for each point on the figure

87
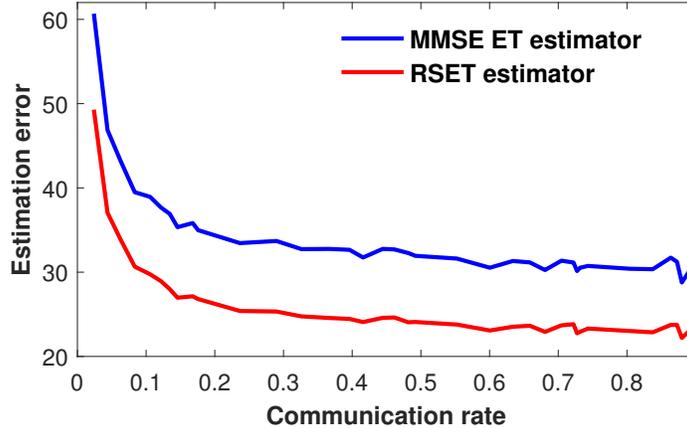
Figure 4.4: Tradeoff between communication rates and estimation performances for the proposed RSET and the MMSE ET prior estimators under different event-triggering conditions.

with the stability condition in part (a) of Theorem 4.2 verified to be satisfied. The average communication rate is defined as $\frac{1}{T}\sum_{k=1}^{T} 1_{\{\gamma_k=1\}}$. We notice that in Fig. 4.4, the proposed RSET estimator always outperforms the MMSE ET estimator no matter what the event-triggering conditions are. In addition, Fig. 4.4 shows that when the communication rate is close to 0, the MSEs for both estimators are large and tend to converge since no information is received at the remote estimator if the communication rate is 0. When the communication rate is close to 1, the MMSE ET estimator simplifies to the Kalman filter and the RSET estimator simplifies to the time-triggered risk-sensitive filter, which outperforms the Kalman filter under non-nominal scenario. This results in the gap between the RSET estimator and the MMSE ET estimator at the right end of Fig. 4.4.

Finally, the prior range of the risk-sensitive parameter which guarantees the proposed estimators to be stable is obtained based on part (b) of Theorem 4.2. We choose $\varsigma = -0.05$, $\Pi = 0.005$ and other parameters as same as that in Fig. 4.1. First, we choose $N = 3$ and obtain $\theta_N = 0.7674$ from (4.37). Then, we obtain $\tau_N = 1.6890 \times 10^{-7}$, which satisfies $\tau_N < \theta_N$ and is the smallest value of θ for which Ω_N^θ in (4.36) is singular. Next, to evaluate β_ρ, we observe that the matrix $G = [1415; 1387]$ can assign all the eigenvalues of the matrix $A - GDC$ to zero, which is likely to yield a satisfactory upper bound β_ρ. In this case, we select $\rho = 1.5$, the solution of (4.39) is $\Sigma_3 = [6.3540 \times 10^6 \ 6.2709 \times 10^6; 6.2709 \times 10^6 \ 6.1889 \times 10^6]$, where the initial covariance P_0 satisfies $P_0 < \Sigma_3$. From (4.38), we obtain $\beta_{1.5} = 4.4292 \times 10^{-8}$. Next, we verify that Σ_k is bounded and $\Sigma_k^{-1} - \theta Q > 0$ for all $k > 0$ with $\theta = 4.42 \times 10^{-8}$, which is chosen

based on part (b) of Theorem 4.2. The cumulative normalized histograms of the eigenvalues of Σ_k and $\Sigma_k^{-1} - \theta Q$ with $\beta_{1.5} = 4.42 \times 10^{-8}$ for runtime $T = 1000000$ are shown in Fig. 4.5, where a cumulative normalized histogram is a mapping that first counts the cumulative number of observations that is smaller than or equal to a specific value, then normalizes the counted value by the total number of observations. In Fig. 4.5, for the eigenvalues $\lambda_1(\Sigma_k)$ and $\lambda_2(\Sigma_k)$, we observe that the normalized cumulative values are zero for $\lambda_1(\Sigma_k) \leq 0$ and $\lambda_2(\Sigma_k) \leq 0$, indicating that the eigenvalues of Σ_k are positive; thus $\Sigma_k > 0$ for $0 < k \leq 1000000$. Similarly, from the two subfigures at the bottom of Fig. 4.5, we have $\Sigma_k^{-1} - \theta Q_k > 0$ for $0 < k \leq 1000000$. Based on Theorem 4.2, we conclude that the proposed RSET prior estimator remains stable; so does the RSET posterior estimator.
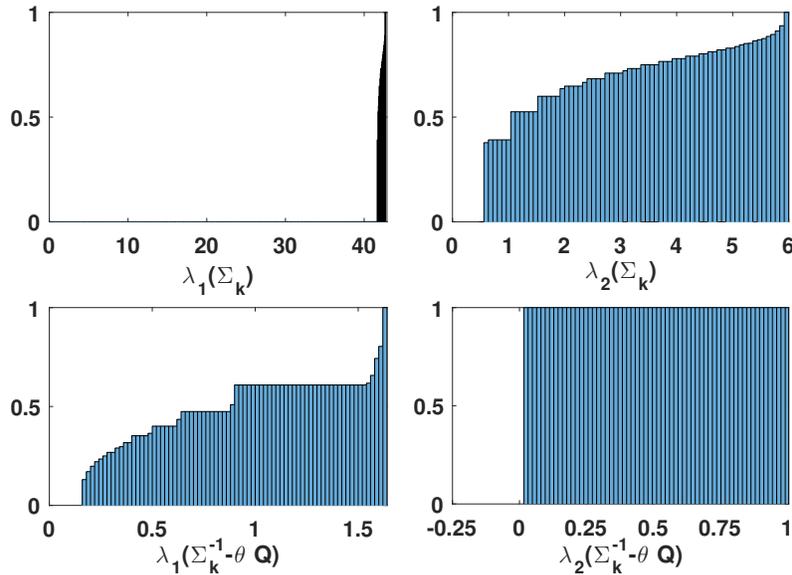
Figure 4.5: Cumulative normalized histograms of the eigenvalues of Riccati solution Σ_k and $\Sigma_k^{-1} - \theta Q$ for finite horizon $T = 1000000$ with $\theta = 4.42 \times 10^{-8}$.

4.1.7 Conclusion

In this section, a robust event-triggered state estimation problem with a stochastic event-triggering condition for linear Gaussian systems is investigated by minimizing a risk-sensitive cost function. Using the reference measure approach, we obtain the closed-form expressions of the risk-sensitive event-triggered prior and posterior state estimates, which evolve in recursive Kalman-like forms. It is shown that the open-loop MMSE event-triggered estimators proposed in [23] are recovered from the proposed

RSTE estimators as a special case when the risk-sensitive parameter tends to zero. Furthermore, we investigate the stability issue of the proposed state estimator and propose sufficient stability conditions, under which, the range of values of the risk-sensitive parameter and covariance of the initial state for which the proposed estimators are stable can be estimated in advance. Finally, comparative simulation results show that the proposed RSET estimator has good estimation performance under nominal scenarios, and acceptable estimation performance that degrades less rapidly than the MMSE ET estimator under non-nominal scenarios, indicating that the proposed estimator is more robust to system uncertainty compared with the MMSE ET estimator. Future work includes risk-sensitive event-triggered estimation for systems with multiple sensors and unreliable communication channels.

4.2 Robust state estimator design for systems with unknown exogenous inputs: A risk-sensitive approach[*]

In this section, a robust state estimation problem for stochastic discrete-time systems with exogenous unknown inputs is formulated by minimizing the risk-sensitive error criterion. The risk-sensitive state estimator is obtained by utilizing the reference measure approach.

4.2.1 Introduction

The unknown exogenous inputs can represent the impact of failure of actuators or plant components, connecting external inputs, as well as unknown and unpredictable cyber attacks in CPSs. Recently, an interesting investigation by Li [40] showed that by treating the unknown inputs as a process with non-informative prior, the Bayesian inference approach was successfully applied to derive the optimal MMSE estimate for systems with partially observed inputs, which was proved to reduce to those obtained by the unified minimum variance (UMV) approaches for the unknown input case. Besides, similar non-informative prior for unknown inputs was considered in [57], where an event-based MMSE estimator for linear-time varying systems with unknown inputs was obtained in a recursive form. The risk-sensitive filtering also formulates the problem

---

from a Bayesian perspective and is closely related to the MMSE estimator, which motivates us to investigate the risk-sensitive estimation problem while considering similar non-informative prior in [40] for the unknown inputs.

Different from the MMSE estimator, the risk-sensitive estimate cannot be simply obtained by minimizing the probability distribution of the state conditioned on the measured information up to current time instant due to the exponential form of the criterion. To overcome this difficulty, the reference measure approach is utilized to decouple the measurement and system state under the reference measure, so that the derivations of the reformulated cost criterion under the reference measure is simplified due to the favorable independence between the measurement and state.

The main contributions of the work are summarized as follows: (1) By proposing a Radon-Nikodym derivative, we propose a reference measure under which the measurement and system state become independent and the measurement process itself is i.i.d.; (2) we define a so-called information state under the reference measure, then derive its closed-form expression, which evolves in a recursive form with a Riccati equation; (3) with the help of the obtained information state, we obtain a recursive algorithm for the risk-sensitive estimate of the system state and the theoretical results are further validated by simulation results.

Notation: $\mathbb{R}$ denotes the set of real numbers. $\mathbb{N}$ denotes the set of nonnegative 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 $z, \xi \in \mathbb{R}^m$, z^i is the ith element of z and the event $\{\xi \leq z\} = \{\xi^i \leq z^i, i = 1, \ldots, m\}$. For a probability measure P ($\bar{\text{P}}$), we use E ($\bar{\text{E}}$) to represent the expectation operator and use $f(\cdot)(\bar{f}(\cdot))$ to represent the probability density function of a random variable. $\mathbf{1}\{\boldsymbol{A}\}$ is the indicator function of set $\boldsymbol{A}$. I_n is the $n \times n$ identity matrix.

4.2.2 Problem formulation

In this work, we consider a discrete-time linear time-invariant (LTI) system with unknown exogenous inputs:

$$x_{k+1} = Ax_k + Gd_k + w_k, \tag{4.43}$$

where $x_k \in \mathbb{R}^n$ is the state, w_k is a zero-mean Gaussian noise with covariance $\sigma > 0$, and $d_k \in \mathbb{R}^p$ is the unknown input. The probability distribution of w_k is described by $\Phi_w(\cdot)$. The initial value x_0 of the state is Gaussian with mean μ_0 and covariance P_0. The

probability distribution function of the initial state x_0 is denoted as $\pi_0(\cdot)$. Let $\mathcal{F}_k^x$ be the completion of the σ-field on Ω generated by $x_0, x_1, \ldots, x_k$. We assume rank $G = p$, which implies $p < n$ and the number of states is larger than that of disturbances. This assumption is intuitive since if there are more disturbances than the states, it will be unlikely to give a reasonable estimate based on the measurement information provided. The state information is measured by a sensor and the output equation is

$$y_k = Cx_k + v_k, \tag{4.44}$$

where $v_k \in \mathbb{R}^m$ is a zero-mean Gaussian noise with covariance $\varphi > 0$. The probability distribution of v_k is described by $\Phi_v(\cdot)$. We assume that x_0, w_k, v_k and d_k are independent of each other and the pair (C, A) is observable. Let $\mathcal{F}_k^y$ be the completion of the σ-field on Ω generated by $y_0, y_1, \ldots, y_k$. Let $\mathcal{I}_k = \{y_0, y_1, \cdots, y_k\}$ denote the available measurement information to the remote estimator up to time instant k. The optimal state estimation problem for systems with unknown exogenous inputs was investigated in the sense of MMSE in [40]. In this section, we consider the risk-sensitive generalization of the MMSE estimation problem. Our objective is to find the risk-sensitive estimate $\hat{x}_k$ of x_k governed by an unknown input term d_k, given the measurement information $\mathcal{I}_k$ such that

$$\hat{x}_k := \arg \min_{\xi \in \mathbb{R}^n} J_k(\xi), \tag{4.45}$$

where the risk-sensitive cost function is defined as

$$J_k(\xi) = \mathrm{E}\left[\exp[\theta \Psi_{0,k}(\xi)]|\mathcal{I}_k\right], \tag{4.46}$$

where $\theta > 0$ is the risk-sensitive parameter and $\Psi_{0,k}(\xi)$ is defined as

$$\Psi_{0,k}(\xi) = \hat{\Psi}_{0,k-1} + \frac{1}{2}(x_k - \xi)^\top Q_k(x_k - \xi), \tag{4.47}$$

with symmetric weighting matrix $Q_k > 0$ and

$$\hat{\Psi}_{0,k-1} = \frac{1}{2} \sum_{t=0}^{k-1} (x_t - \hat{x}_t)^\top Q_t(x_t - \hat{x}_t).$$

The consideration of exponential-form cost function penalizes all the higher order moments of the estimation error energy, so that the risk-sensitive estimator is more robust to system uncertainties compared with the MMSE estimator. In this work, we investigate the state estimation problem utilizing a Bayesian inference approach. As no

information is available for the unknown input d_k by the remote estimator, we model it with a non-informative prior distribution (improper distribution), i.e.,

$$f(d_k|\mathcal{I}_k) \propto 1. \tag{4.48}$$

Let $\mathcal{F}_k^d$ be the completion of the σ-field on Ω generated by d_0, d_1, ..., d_k. The intuition of this model is that all possible values of the unknown input d_k are equally likely to occur since the estimator has no clue of the value of d_k. In addition, the optimal MMSE estimate for systems with partially observed inputs [40], which is obtained by utilizing the Bayesian inference approach and treating the unknown input as a process with non-informative prior, was shown to reduce to those obtained by the UMV approach [33]. This result justifies the rationale of modeling the unknown input d_k with the non-informative prior. For further discussion on the property of improper prior distribution, see Remark 1 of [57].

One main difference between the MMSE and risk-sensitive estimators is that the MMSE estimator can be obtained based on the conditional probability distribution of the current state on the information available at the estimator, thus avoiding the complicated derivations of the mean of the summed squared estimation error from time instant 0 to k; whereas similar simplification does not work for the risk-sensitive case due to its exponential form of the cost function. Motivated by this, the reference measure approach is utilized to simplify the estimation problem and makes it possible for us to derive the recursive algorithm of calculating the risk-sensitive state estimate.

4.2.3 Reference measure approach

To obtain the risk-sensitive estimate of the state x_k governed by an unknown input d_k, we utilize the Bayesian inference approach and the reference measure approach. Firstly, we propose a reference measure $\bar{P}$ under which the variables satisfy the following relationships:

$$x_{k+1} = Ax_k + Gd_k + w_k, \tag{4.49}$$

where $w_k \in \mathbb{R}^n$ is i.i.d. and $w_k \sim N(0, \sigma)$. The measurement y_k is an i.i.d. variable under the reference measure $\bar{P}$ and $y_k \sim N(0, \varphi)$. The reason of proposing a reference measure is that the independency of the measurement y_k under the reference measure $\bar{P}$ will help us simplify the derivations of the risk-sensitive cost function under the reference measure. We assume the following relationship under $\bar{P}$

$$v_k = y_k - Cx_k. \tag{4.50}$$

We assume the unknown input follows a non-informative prior distribution under the reference measure, i.e.,

$$\bar{f}(d_k|\mathcal{I}_k) \propto 1. \tag{4.51}$$

To link the reference measure with the original "real-world" measure, we define a map from $\bar{P}$ to P:

$$\left.\frac{dP}{d\bar{P}}\right|_{\mathcal{G}_k} = \bar{\Lambda}_k, \tag{4.52}$$

where $\mathcal{G}_k$ is $\mathcal{F}_k^x \cup \mathcal{F}_k^y \cup \mathcal{F}_k^d$, and the Radon-Nikodym derivative is defined as

$$\bar{\Lambda}_k = \prod_{l=0}^{k} \bar{\lambda}_l, \quad \bar{\lambda}_k = \frac{\Phi_v(y_k - Cx_k)}{\Phi_v(y_k)}. \tag{4.53}$$

Based on the conditional Bayes theorem (Theorem 3.3 in Chapter 2 of [17]), the probability distribution under the reference measure $\bar{P}$ can be mapped back to the "real-world" measure P. The properties of the map defined in (4.52)-(4.53) from the reference measure to the original probability measure are summarized in the following result.

Lemma 4.9. *If the model in (4.49)-(4.51) is mapped from the reference measure $\bar{P}$ back to the "real-world" measure P via the Radon-Nikodym derivative in (4.52)-(4.53), then the obtained model satisfies the properties under measure P as stated in (4.43), (4.44) and (4.48).*

Proof. According to the definition of λ_{k+1} and equation (4.49), we have

$$\begin{aligned}
&\bar{E}\left[\bar{\lambda}_{k+1}|\mathcal{G}_k\right] \\
=&\bar{E}\left[\frac{\Phi_v(y_{k+1} - C(Ax_k + Gd_k + w_k))}{\Phi_v(y_{k+1})}\middle|\mathcal{G}_k\right].
\end{aligned}$$

Since under the reference measure $\bar{P}$, y_{k+1} and x_k, d_k are independent and y_{k+1} is i.i.d. with probability distribution $N(0, \varphi)$, we have

$$\begin{aligned}
&\bar{E}\left[\bar{\lambda}_{k+1}|\mathcal{G}_k\right] \\
=&\int_{\mathbb{R}^n}\int_{\mathbb{R}^m}\frac{\Phi_v(y_{k+1}-C(Ax_k+Gd_k+w_k))}{\Phi_v(y_{k+1})}\Phi_v(y_{k+1})\Phi_w(w_k)dy_{k+1}dw_k \\
=&\int_{\mathbb{R}^n}\int_{\mathbb{R}^m}\Phi_v(y_{k+1}-C(Ax_k+Gd_k+w_k)) \\
&\cdot d(y_{k+1}-C(Ax_k+Gd_k+w_k))\Phi_w(w_k)dw_k \\
=&1.
\end{aligned}$$

Following a similar procedure, we have

$$\bar{\mathrm{E}}\left[\bar{\lambda}_k|\mathcal{F}_k^x \cup \mathcal{F}_{k-1}^y \cup \mathcal{F}_{k-1}^d\right] = 1.$$

According to the conditional Bayes theorem, we have

$$\begin{aligned}
&\mathrm{P}[x_{k+1} - Ax_k - Gd_k \leq z|\mathcal{G}_k] \\
&= \frac{\bar{\Lambda}_k \bar{\mathrm{E}}\left[\bar{\lambda}_{k+1}\mathbf{1}\{x_{k+1} - Ax_k - Gd_k \leq z\}|\mathcal{G}_k\right]}{\bar{\Lambda}_k \bar{\mathrm{E}}\left[\bar{\lambda}_{k+1}|\mathcal{G}_k\right]} \\
&= \bar{\mathrm{E}}\left[\bar{\lambda}_{k+1}\mathbf{1}\{x_{k+1} - Ax_k - Gd_k \leq z\}|\mathcal{G}_k\right],
\end{aligned}$$

where we use the fact that $\bar{\mathrm{E}}\left[\bar{\lambda}_{k+1}|\mathcal{G}_k\right] = 1$. According to (4.49) and the fact that $w_k \in \mathbb{R}^n$ is i.i.d. and $w_k \sim N(0, \sigma)$ under the reference measure $\bar{\mathrm{P}}$, we further have

$$\begin{aligned}
&\mathrm{P}[x_{k+1} - Ax_k - Gd_k \leq z|\mathcal{G}_k] \\
&= \bar{\mathrm{E}}\left[\frac{\Phi_v[y_{k+1} - C(Ax_k + Gd_k + w_k)]}{\Phi_v(y_{k+1})}\mathbf{1}\{w_k \leq z\}\Big|\mathcal{G}_k\right] \\
&= \int_{\mathbb{R}^n}\int_{\mathbb{R}^m} \frac{\Phi_v[y_{k+1} - C(Ax_k + Gd_k + w_k)]}{\Phi_v(y_{k+1})}\mathbf{1}\{w_k \leq z\} \\
&\quad \cdot \Phi_v(y_{k+1})\Phi_w(w_k) \, \mathrm{d}y_{k+1} \, \mathrm{d}w_k \\
&= \int_{-\infty}^{z^1} \cdots \int_{-\infty}^{z^n} \Phi_w(w_k) \, \mathrm{d}w_k,
\end{aligned}$$

which recovers the relationship described in (4.43) under the "real-world" measure P. Furthermore, for the measurement process, according to the conditional Bayes theorem and (4.50), we have

$$\begin{aligned}
&\mathrm{P}[y_k - Cx_k \leq z|\mathcal{F}_k^x \cup \mathcal{F}_{k-1}^y \cup \mathcal{F}_{k-1}^y] \\
&= \bar{\mathrm{E}}\left[\frac{\Phi_v(y_k - Cx_k)}{\Phi_v(y_k)}\mathbf{1}\{y_k - Cx_k \leq z\}\Big|\mathcal{F}_k^x \cup \mathcal{F}_{k-1}^y \cup \mathcal{F}_{k-1}^y\right] \\
&= \int_{\mathbb{R}^m} \Phi_v(y_k - Cx_k)\mathbf{1}\{y_k - Cx_k \leq z\} \, \mathrm{d}(y_k - Cx_k) \\
&= \int_{-\infty}^{z^1} \cdots \int_{-\infty}^{z^n} \Phi_v(v_k) \, \mathrm{d}v_k.
\end{aligned}$$

According to repeated conditioning, we have

$$\begin{aligned}
&\mathrm{P}[y_k - Cx_k \leq z|\mathcal{F}_k^x] \\
&= \mathrm{P}[\mathrm{P}[y_k - Cx_k \leq z|\mathcal{F}_k^x \cup \mathcal{F}_{k-1}^y \cup \mathcal{F}_{k-1}^y]|\mathcal{F}_k^x] \\
&= \int_{-\infty}^{z^1} \cdots \int_{-\infty}^{z^n} \Phi_v(v_k) \, \mathrm{d}v_k,
\end{aligned}$$

which recovers the measurement process described in (4.44) under P. Following similar arguments as above, we can prove that the model in (4.48) can be obtained by mapping the relationship under $\bar{\mathrm{P}}$ in (4.51) back to the "real-world" measure P through the Radon-Nikodym derivative in (4.52) and (4.53), which completes the proof. $\qquad\square$

Based on the conditional Bayes theorem (Theorem 3.3 in Chapter 2 of [17]) and the fact that $\mathcal{I}_k$ is a sub-σ-field of $\mathcal{G}_k$, the risk-sensitive cost function in (4.46) can be reformulated as

$$
\begin{aligned}
J_k(\xi) &= \mathrm{E}\left[\exp[\theta\Psi_{0,k}(\xi)]|\mathcal{I}_k\right] \\
&= \frac{\bar{\mathrm{E}}[\bar{\Lambda}_k \exp[\theta\Psi_{0,k}(\xi)]|\mathcal{I}_k]}{\bar{\mathrm{E}}[\bar{\Lambda}_k|\mathcal{I}_k]}.
\end{aligned}
\tag{4.54}
$$

Noticing that the denominator in (4.54) is independent of ξ, we have

$$
\hat{x}_k = \arg\min_{\xi\in\mathbb{R}^n} J_k(\xi) = \arg\min_{\xi\in\mathbb{R}^n} \bar{\mathrm{E}}[\bar{\Lambda}_k \exp[\theta\Psi_{0,k}(\xi)]|\mathcal{I}_k],
\tag{4.55}
$$

which indicates that the problem of minimizing the risk-sensitive cost function $J_k(\xi)$ under the "real-world" measure P is equivalent to that of minimizing the reformulated cost function $\bar{\mathrm{E}}[\bar{\Lambda}_k \exp[\theta\Phi_{0,k}(\xi)]|\mathcal{I}_k]$ under the reference measure $\bar{\mathrm{P}}$. The independence of the measurement process y_k under $\bar{\mathrm{P}}$ can be utilized to further simplify the minimization problem under $\bar{\mathrm{P}}$.

4.2.4 Augmented variable and information state

Let $G^\perp$ denote a matrix such that $[G \ G^\perp] \in \mathbb{R}^{n\times n}$, $\mathrm{rank}[G \ G^\perp] = n$ and $G^\top G^\perp = 0$. In this section, we define a new augmented variable through a linear transformation

$$
z_k := Tx_k,
\tag{4.56}
$$

with $T := [G \ G^\perp]^{-1}$. By substituting (4.43) into the above equation, we obtain

$$
\begin{aligned}
z_k &= T(Ax_{k-1} + Gd_{k-1} + w_{k-1}) \\
&= \tilde{A}_{k-1}z_{k-1} + Dd_{k-1} + \tilde{w}_{k-1},
\end{aligned}
\tag{4.57}
$$

where $\tilde{A} = TAT^{-1}$, $D = TG = [I_p \ \mathbf{0}]^\top$, $\tilde{w}_{k-1} = Tw_{k-1}$ and $\tilde{w}_k$ is a Gaussian noise with zero mean and covariance $\tilde{\sigma} = T\sigma T^\top$. The probability distribution of $\tilde{w}_k$ is described by $\Phi_{\tilde{w}}(\cdot)$. Since T is an invertible square matrix, the weighting matrix Q_k in (4.47) can be written as

$$
Q_k = T^\top R_k T,
\tag{4.58}
$$

with $R_k = (T^\top)^{-1}Q_k T^{-1}$. By substituting (4.56) and (4.58) into (4.47) and reformulating $\Psi_{0,k}(\xi)$, we have

$$
\Psi_{0,k}(\xi) = \Gamma_{0,k}(\mu) = \hat{\Gamma}_{0,k-1} + \tfrac{1}{2}(z_k - \mu)^\top R_k(z_k - \mu),
$$

where

$$\hat{\Gamma}_{0,k-1} = \tfrac{1}{2} \sum_{t=0}^{k-1} (z_t - \hat{z}_t)^\top R_t (z_t - \hat{z}_t)$$

with $\hat{z}_k := T\hat{x}_k$ and $\mu := T\xi$. To solve the problem in (4.55), we may first work on

$$\hat{z}_k = \arg \min_{\mu \in \mathbb{R}^n} \bar{\mathrm{E}}[\bar{\Lambda}_k \exp[\theta\Gamma_{0,k}(\mu)]|\mathcal{I}_k], \qquad (4.59)$$

and then obtain the state estimate through $\hat{x}_k = T^{-1}\hat{z}_k$.

Definition 4.2 (Information State). *The information state $\alpha_k(z)$ is defined as the unnormalized density function such that*

$$\alpha_k(z)\mathrm{d}z = \bar{\mathrm{E}}[\bar{\Lambda}_{k-1} \exp\{\theta\hat{\Gamma}_{0,k-1}\}\mathbf{1}(z_k \in \mathrm{d}z)|\mathcal{I}_{k-1}]. \qquad (4.60)$$

The physical meaning of $\alpha_k(z)\mathrm{d}z$ is the information state of an augmented plant where the state includes the augmented variable z_k and part of the risk-sensitive cost in (4.59) under the reference measure $\bar{\mathrm{P}}$. Before continuing, we introduce some notations. Write

$$\bar{D} = [0 \ I_{n-p}]^\top \in \mathbb{R}^{n \times (n-p)}$$

and

$$\bar{\sigma} = \bar{D}(\bar{D}^\top \tilde{\sigma} \bar{D})^{-1} \bar{D}^\top.$$

The following result provides the recursive form of the information state $\alpha_k(z)$ for $k \geq 1$.

Lemma 4.10. *The information state $\alpha_k(z)$ is unnormalized Gaussian density given by*

$$\alpha_k(z) = \rho_k \exp\left[-\tfrac{1}{2}(z^\top \Sigma_k z - 2\zeta_k^\top z)\right], \qquad (4.61)$$

where ρ_k is a constant which is independent of z; Σ_k and ζ_k evolve according to the following form:

$$\Sigma_{k+1} = \bar{\sigma} - \bar{\sigma}\tilde{A}[\tilde{A}^\top \bar{\sigma}\tilde{A} - \theta R_k + \Sigma_k + \tilde{\varphi}]^{-1}\tilde{A}^\top \bar{\sigma}^\top, \qquad (4.62)$$

$$\zeta_{k+1}^\top = (\zeta_k^\top - \theta\hat{z}_k^\top R_k + y_k^\top \varphi^{-1}CT^{-1})$$

$$\cdot [\tilde{A}^\top \bar{\sigma}\tilde{A} - \theta R_k + \Sigma_k + \tilde{\varphi}]^{-1}\tilde{A}^\top \bar{\sigma}^\top, \qquad (4.63)$$

$$\tilde{\varphi} = (T^{-1})^\top C^\top \varphi^{-1}CT^{-1}. \qquad (4.64)$$

with initial conditions $\Sigma_0 = (TP_0T^\top)^{-1}$ and $\zeta_0^\top = (T\mu_0)^\top(TP_0T^\top)^{-1}$.

Proof. To prove this lemma, we start with the general case, then the case of $k = 0$ can be easily obtained, which will be shown at the end of the proof. We first assume that the information state at time instant k follows the form in (4.61) and prove this form holds for $k+1$. Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is any Borel test function. Based on the definition of $\bar{\Lambda}_k$ and $\hat{\Gamma}_{0,k}$, we have

$$
\begin{aligned}
&\bar{\text{E}}[f(z_{k+1})\bar{\Lambda}_k \exp(\theta\hat{\Gamma}_{0,k})|\mathcal{I}_k] \\
=&\bar{\text{E}}\Big[f(\tilde{A}x_k + Dd_k + \tilde{w}_k)\tfrac{\Phi_v(y_k - CT^{-1}z_k)}{\Phi_v(y_k)}\bar{\Lambda}_{k-1}\exp(\theta\hat{\Gamma}_{0,k-1}) \\
&\cdot \exp\big[\tfrac{\theta}{2}(z_k - \hat{z}_k)^\top R_k(z_k - \hat{z}_k)\big] \big|\mathcal{I}_k\Big].
\end{aligned}
$$

According to the definiton of the information state $\alpha_k(z)$, we have

$$
\begin{aligned}
&\bar{\text{E}}[f(z_{k+1})\bar{\Lambda}_k \exp(\theta\hat{\Gamma}_{0,k})|\mathcal{I}_k] \\
=& \int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^p} f(\tilde{A}x + Dd_k + \tilde{w}_k)\Phi_{\tilde{w}}(w)f(d|y_k)\,\mathrm{d}d\,\mathrm{d}w \\
&\cdot \tfrac{\Phi_v(y_k - CT^{-1}x)}{\Phi_v(y_k)}\exp\big[\tfrac{\theta}{2}(x - \hat{z}_k)^\top R_k(x - \hat{z}_k)\big]\alpha_k(x)\,\mathrm{d}x.
\end{aligned}
$$

Let $z = \tilde{A}x + Dd_k + \tilde{w}_k$ and the above equation becomes

$$
\begin{aligned}
&\bar{\text{E}}[f(z_{k+1})\bar{\Lambda}_k \exp(\theta\hat{\Gamma}_{0,k})|\mathcal{I}_k] \\
=& \int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^p} f(z)\Phi_{\tilde{w}}(z - \tilde{A}x - Dd)f(d|y_k)\,\mathrm{d}d\,\alpha_k(x) \\
&\cdot \exp\big[\tfrac{\theta}{2}(x - \hat{z}_k)^\top R_k(x - \hat{z}_k)\big]\tfrac{\Phi_v(y_k - CT^{-1}x)}{\Phi_v(y_k)}\,\mathrm{d}x\,\mathrm{d}z.
\end{aligned}
$$

Based on the definition of the information state, we have

$$
\bar{\text{E}}[f(z_{k+1})\bar{\Lambda}_k \exp(\theta\hat{\Gamma}_{0,k})|\mathcal{I}_k] = \int_{\mathbb{R}^n} f(z)\alpha_{k+1}(z)\,\mathrm{d}z.
$$

Combining the above results, we obtain

$$
\begin{aligned}
&\alpha_{k+1}(z) \\
=& \int_{\mathbb{R}^n}\int_{\mathbb{R}^p} \Phi_{\tilde{w}}(z - \tilde{A}x - Dd)f(d|y_k)\,\mathrm{d}d\,\alpha_k(x) \\
&\cdot \exp\big[\tfrac{\theta}{2}(x - \hat{z}_k)^\top R_k(x - \hat{z}_k)\big]\tfrac{\Phi_v(y_k - CT^{-1}x)}{\Phi_v(y_k)}\,\mathrm{d}x \\
\propto& \int_{\mathbb{R}^n}\int_{\mathbb{R}^p} \exp[-\tfrac{1}{2}(z - \tilde{A}x - Dd)^\top \tilde{\sigma}^{-1}(z - \tilde{A}x - Dd)]\,\mathrm{d}d \\
&\cdot \exp\big[\tfrac{\theta}{2}(x - \hat{z}_k)^\top R_k(x - \hat{z}_k)\big]\Phi_v(y_k - CT^{-1}x)\alpha_k(x)\mathrm{d}x,
\end{aligned}
$$

where (4.51) is used in the last equation. We denote $\tilde{z} = z - \tilde{A}x$ and have

$$
\tilde{z} = \begin{bmatrix} D^\top \\ \bar{D}^\top \end{bmatrix}\tilde{z} = \begin{bmatrix} \tilde{z}_1 \\ \tilde{z}_2 \end{bmatrix},
$$

where $\bar{D} = [\mathbf{0}\ I_{n-p}]^\top$. By basic manipulations, we have

$$\int_{\mathbb{R}^p} \exp\left[-\tfrac{1}{2}(z - \tilde{A}x - Dd)^\top \tilde{\sigma}^{-1}(z - \tilde{A}x - Dd)\right] \mathrm{d}d$$

$$= \int_{\mathbb{R}^p} \exp\left[-\frac{1}{2}\begin{bmatrix} d - \tilde{z}_1 \\ -\tilde{z}_2 \end{bmatrix}^\top \begin{bmatrix} D^\top \tilde{\sigma} D & D^\top \tilde{\sigma} \bar{D} \\ \bar{D}^\top \tilde{\sigma} D & \bar{D}^\top \tilde{\sigma} \bar{D} \end{bmatrix} \begin{bmatrix} d - \tilde{z}_1 \\ -\tilde{z}_2 \end{bmatrix}\right] \mathrm{d}d$$

$$\propto \exp\left[-\tfrac{1}{2}(z - \tilde{A}x)^\top \bar{\sigma}(z - \tilde{A}x)\right],$$

where $\bar{\sigma} = \bar{D}(\bar{D}^\top \tilde{\sigma} \bar{D})^{-1} \bar{D}^\top$ and the properties of the marginal distribution of a multivariate Gaussian random variable are used. Combining the above results and (4.61), we have

$$\alpha_{k+1}(z)$$
$$\propto \int_{\mathbb{R}^n} \exp\left[-\tfrac{1}{2}(z - \tilde{A}x)^\top \bar{D}(\bar{D}^\top \tilde{\sigma}^{-1} \bar{D})^{-1} \bar{D}^\top (z - \tilde{A}x)\right]$$
$$\cdot \exp\left[\tfrac{\theta}{2}(x - \hat{z}_k)^\top R_k(x - \hat{z}_k)\right] \exp\left[-\tfrac{1}{2}(y - CT^{-1}x)^\top \varphi^{-1}(y - CT^{-1}x)\right]$$
$$\cdot \exp\left[-\tfrac{1}{2}(z^\top \Sigma_k z - 2\zeta_k^\top z)\right] \mathrm{d}x.$$

Based on the properties of Gaussian distributions and completing square technique for vectors, we have

$$\alpha_{k+1}(z) \propto \exp\left[-\tfrac{1}{2}(z^\top \Sigma_{k+1} z - 2\zeta_{k+1}^\top z)\right],$$

where Σ_{k+1} and $\zeta_{k+1}^\top$ follow equations (4.62)-(4.64). For the case when $k = 0$, we can easily obtain that

$$\alpha_0(z) \propto \exp\left[-\tfrac{1}{2}(z - T\mu_0)^\top (TP_0 T^\top)^{-1}(z - T\mu_0)\right].$$

Thus, the initial conditions of the information state follows (4.61) with $\Sigma_0 = (TP_0 T^\top)^{-1}$ and $\zeta_0^\top = (T\mu_0)^\top (TP_0 T^\top)^{-1}$, which completes the proof. $\qquad\square$

Based on the above results of the information state, we derive the risk-sensitive estimate $\hat{z}_k$.

Lemma 4.11. *For $k \geq 1$, if $\Sigma_k - \theta R_k + \tilde{\varphi} > 0$, the risk-sensitive estimate of the augmented state $\hat{z}_k$ satisfies*

$$\hat{z}_k = (\tilde{\varphi} + \Sigma_k)^{-1}[(T^{-1})^\top C^\top (\varphi^{-1})^\top y_k + \zeta_k]. \tag{4.65}$$

Proof. According to (4.59), we need to calculate $\bar{\mathrm{E}}[\bar{\Lambda}_k \exp[\theta\Gamma_{0,k}(\mu)]|\mathcal{I}_k]$. Combining the definition of $\bar{\Lambda}_k$ and the above results on the information state α_k, we have

$$
\begin{aligned}
&\bar{\mathrm{E}}[\bar{\Lambda}_k \exp[\theta\Gamma_{0,k}(\mu)]|\mathcal{I}_k]\\
&= \int_{\mathbb{R}^n} \frac{\Phi_v(y_k - CT^{-1}z)}{\Phi_v(y_k)} \alpha_k(z) \exp\left[\frac{\theta}{2}(z_k - \mu)^\top R_k(z_k - \mu)\right] \mathrm{d}z\\
&\propto \int_{\mathbb{R}^n} \exp\left[-\frac{1}{2}(y - CT^{-1}z)^\top \varphi^{-1}(y - CT^{-1}z)\right.\\
&\quad \left.-\frac{1}{2}(z^\top \Sigma_k z - 2\zeta_k^\top z) + \frac{\theta}{2}(z_k - \mu)^\top R_k(z_k - \mu)\right]\mathrm{d}z.
\end{aligned}
$$

Using the properties of Gaussian distributions and completing square technique for vectors, we obtain

$$
\bar{\mathrm{E}}[\bar{\Lambda}_k \exp[\theta\Gamma_{0,k}(\mu)]|\mathcal{I}_k] \propto \exp[(\mu - \mu')^\top M(\mu - \mu')],
$$

where $M = \Sigma_k - \theta R_k + \tilde{\varphi}$ is independent of μ and

$$
\mu' = (\tilde{\varphi} + \Sigma_k)^{-1}[(T_{k-1}^{-1})^\top C^\top (\varphi^{-1})^\top y_k + \zeta_k].
$$

According to the definition of $\hat{z}_k$ in (4.59), if $M > 0$, the minimum of $\bar{\mathrm{E}}[\bar{\Lambda}_k \exp[\theta\Gamma_{0,k}(\mu)]|\mathcal{I}_k]$ is achieved when $\mu = \mu'$, thus, the risk-sensitive estimate of the augmented state follows (4.65), which completes the proof. □

4.2.5 Recursive risk-sensitive state estimate

Combining Lemmas 4.10 and 4.11, we have the following results on the recursive form of the risk-sensitive state estimate $\hat{x}_k$.

Theorem 4.3. *For the linear systems with unknown exogenous inputs in (4.43) and (4.44), if $\Sigma_k - \theta R_k + \tilde{\varphi} > 0$, for $k \geq 1$, the risk-sensitive estimate of the system state $\hat{x}_k$ evolves according to the following recursive form:*

$$
\begin{aligned}
&\hat{x}_{k+1}\\
&= T^{-1}(\tilde{\varphi}^{-1} + \Sigma_k)^{-1}\big[(T^{-1})^\top C^\top (\varphi^{-1})^\top y_{k+1}\\
&\quad + \bar{\sigma}\tilde{A}(\tilde{A}^\top \bar{\sigma}\tilde{A} - \theta R_k + \Sigma_k + \tilde{\varphi})^{-1}(-\theta R_k + \Sigma_k + \tilde{\varphi})T\hat{x}_k\big], \quad (4.66)
\end{aligned}
$$

where $\hat{x}_0 = T^{-1}(\tilde{\varphi} + \Sigma_z)^{-1}[(T^{-1})^\top C^\top (\varphi^{-1})^\top y_k + ((TP_0T^\top)^{-1})^\top T\mu_0]$ and Σ_k follows (4.62).

Proof. From (4.65), we have

$$\zeta_k = (\tilde{\varphi} + \Sigma_k)\hat{z}_k - (T^{-1})^\top C^\top (\varphi^{-1})^\top y_k. \tag{4.67}$$

Substituting (4.67) into (4.63), we obtain

$$\zeta_{k+1} = \bar{\sigma}\tilde{A}[\tilde{A}^\top \bar{\sigma}\tilde{A} - \theta R_k + \Sigma_k + \tilde{\varphi}]^{-1}$$
$$\cdot (\Sigma_k - \theta R_k + \tilde{\varphi})\hat{z}_k. \tag{4.68}$$

We combine (4.68), the risk-sensitive estimate $\hat{z}_{k+1}$ in (4.65) and the linear transformation $\hat{x}_k = T^{-1}\hat{z}_k$, then it follows that the risk-sensitive estimate $\hat{x}_{k+1}$ evolves according to (4.66). For the initial condition of $k = 0$, by substituting ζ_0 in Lemma 4.10 into (4.65) and due to that $\hat{x}_k = T^{-1}\hat{z}_k$, we obtain $\hat{x}_0 = T^{-1}(\tilde{\varphi} + \Sigma_z)^{-1}[(T^{-1})^\top C^\top (\varphi^{-1})^\top y_k + ((TP_0T^\top)^{-1})^\top T\mu_0]$. $\square$

In Theorem 4.3, a sufficient condition $\Sigma_k - \theta R_k + \tilde{\varphi} > 0$ is employed to guarantee that the reformulated risk-senstive cost function $\bar{\mathbb{E}}[\bar{\Lambda}_k \exp[\theta\Gamma_{0,k}(\mu)]|\mathcal{I}_k]$ is non-degenerate, so that we can obtain the risk-sensitive estimate of the state.

4.2.6 Numerical example

To demonstrate the analytical results, we present a simulation example in this subsection. We consider a stable LTI process with unknown inputs in (4.43) with nominal model $A = [0.5887 \ 0.3251; 0.5028 \ 0.3245]$, $\Delta A = [0 \ 0; \delta \ 0]$ representing the model uncertainty. The process noise w_k is a zero-mean Gaussian noise with covariance $\sigma = [1.1778 \ 0; 0 \ 1.1778]$. For this case, the unknown signal d_k utilized is chosen to have the form shown in Fig. 4.6 with $G = [0.5722 \ 0.0090]^\top$. We consider the measurement process in (4.44) with $C = [0.5666 \ -0.9671; -0.2248 \ -0.7100]$ and noise covariance $\varphi = [2.0605 \ 0; 0 \ 2.0605]$. For the risk-sensitive estimation problem in (4.45), we choose the weighting matrix $Q_k = [1 \ 0; 0 \ 1]$ and the risk-sensitive parameter $\theta = 0.0819$. The performance comparison between the estimates from the Kalman filter (KF), MMSE estimator for systems with unknown inputs and the proposed risk-sensitive (RS) estimator for $\delta = -0.3$ are shown in Fig. 4.6. The vertical axis of Fig. 4.6 represents the first element of the system state, denoted as x_k^1. The resultant MSE for the KF, MMSE estimator and RS estimator are 56.04, 18.62 and 6.80, respectively, indicating that the proposed RS estimator outperforms the MMSE estimator when system parameter uncertainty exists. Though the MMSE estimator serves as the best estimator in the mean

101
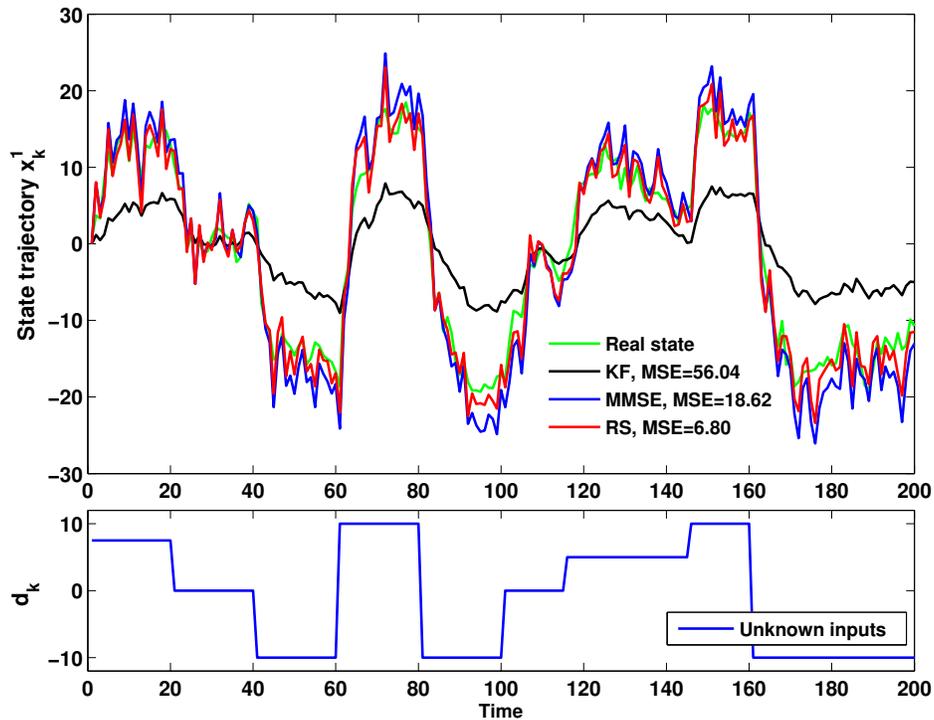
Figure 4.6: Performance comparison between the KF, MMSE estimator for systems with unknown inputs and proposed RS estimator.

square error sense under the nominal model case, the merits of the proposed RS estimator will unfold in the scenario with system parameter uncertainties. The risk-sensitive parameter can be tuned by the user to balance the tradeoff between the estimator optimality under the nominal case and the robustness to the system uncertainty. A larger risk-sensitive parameter can result in a relatively degraded but still acceptable performance under the nominal case; while this larger choice of parameter can lead to noticeable performance improvement for the scenario with system uncertainty.

4.2.7 Conclusion

In this work, the risk-sensitive state estimation problem for discrete-time LTI systems with unknown exogenous inputs is investigated by treating the unknown input as a process modeled by a non-informative prior. We propose a reference measure under which the system state and measurement become independent and then derive a recursive form of an information state under the reference measure, based on which, we propose a recursive algorithm for the risk-sensitive state estimate. Simulation examples are included to show the effectiveness of the proposed method, where comparative

estimation results of the proposed RS estimator and the MMSE estimator for systems with unknown inputs indicate that the former outperforms the latter under the scenario subject to system parameter uncertainties. In the current estimation framework, the measurement transmission protocol is in a time-triggered fashion; one extension is to take the event-triggered scheduling into consideration. Besides, the stability analysis of the RS estimator is another interesting direction to be explored in our future work.

Chapter 5

Distributed robust state estimation for sensor networks: A risk-sensitive approach*

5.1 Introduction

In this chapter, we investigate the robust distributed state estimation problem utilizing the risk-sensitive filtering approach. The motivation stems from that for wireless sensor networks, the sensors involved in the networks are usually intelligent nodes with limited computation capability and constrained power supply. For large-scale sensor networks, sensors typically exchange information only with their neighbors due to their limited power and communication bandwidth. Even if the all-to-all links can be achieved, the capability of local filter would not afford the heavy computational burden if the information from the entire network is utilized. As a result, distributed state estimation problem of a dynamical system which is measured by a sensor network is of vital importance.

In this work, we consider the risk-sensitive filter, which counteracts the uncertainty by minimizing a cost criterion in exponential form. We aim to develop a distributed risk-sensitive filtering algorithm for a linear Gaussian system that is measured by a sensor network, where the nodes can communicate only with their neighbors and each node runs a local filter to estimate the state based on the measurements from its neighbors. Based on the fact that risk-sensitive filters for linear Gaussian systems evolve in a recursive Kalman-like forms [70], we obtain a decentralized risk-sensitive

---

*Parts of the results in this section were submitted to the 57th IEEE Conference on Decision and Control, Miami Beach, FL, USA, December 17-19, 2018.

filter, which involves the two average terms of the dynamics from the entire sensor network. By utilizing the consensus filter proposed in [69] to achieve data fusion, we propose a distributed risk-sensitive filtering algorithm, where the local filter at each sensor updates its state estimate in a Riccati-based linear recursive form. We analyze the asymptotic properties of the proposed distributed risk-sensitive filter for LTI systems. The effectiveness of the proposed distributed risk-sensitive filtering algorithm is illustrated by simulation examples.

Notation: $\mathbb{R}$ denotes the set of real numbers. $\mathbb{N}$ denotes the set of nonnegative 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 $\mathbb{R}^m := \mathbb{R}^{m \times 1}$. Let $\mathbb{N}$ denote the set of nonnegative integers. Write $\mathbb{N}_{1:M} := \{1, 2, \ldots, M\}$. I_n is the $n \times n$ identity matrix. $0_{n \times m}$ is the $n \times m$ zero matrix. $\mathrm{col}(C_1, C_2, \ldots, C_N)$ denotes $[C_1^\top \ C_2^\top \ \ldots \ C_N^\top]^\top$. $\mathrm{diag}(A_1, A_2, \ldots, A_3)$ stands for a block-diagonal matrix.

5.2   Problem formulation

In this chapter, we consider a discrete-time linear process driven by white noise:

$$x(k + 1) = A(k)x(k) + B(k)w(k), \tag{5.1}$$

where $x(k) \in \mathbb{R}^n$ is the state, $w(k)$ is a zero-mean Gaussian noise with covariance $Q(k) > 0$. The initial value $x(0)$ of the state is Gaussian with mean μ_0 and covariance P_0. The state information is measured by a sensor network consisting of N nodes. The interaction topology among the nodes is represented by a graph $G = (V, E)$ where $V = \{1, 2, \ldots, N\}$ is the set of nodes and $E \subset V \times V$ is the set of links. We assume the graph G is undirected and connected (there is a path between every pair of nodes). For $i \in V$, the output equation of the ith sensor is

$$y_i(k) = C_i(k)x(k) + v_i(k), \tag{5.2}$$

where $v_i(k) \in \mathbb{R}^m$ is a zero-mean Gaussian noise with covariance $R_i(k) > 0$. We assume that $x(0)$, $w(k)$ and $v_i(k)$ for $i \in V$ are independent of each other. The output matrix C_i's can be different across the sensor network. Let $\mathbf{y}(k) = \mathrm{col}(y_1(k), y_2(k), \ldots, y_N(k)) \in \mathbb{R}^{mN}$ be the collective sensor measurement of the sensor network at time instant k. Defining an output matrix $C(k) = \mathrm{col}(C_1(k), C_2(k), ..., C_N(k))$, we have the sensing model of the entire sensor network:

$$\mathbf{y}(k) = C(k)x(k) + \mathbf{v}(k), \tag{5.3}$$

where $\mathbf{v}(k) = \mathrm{col}(v_1(k), v_2(k), \ldots, v_N(k)) \in \mathbb{R}^{mN}$ is the collective measurement noise with covariance $R(k) = \mathrm{diag}(R_1(k), R_2(k), \ldots, R_N(k))$. We assume that the pair $(A(k), C(k))$ is observable, namely, the system is collectively observable by the entire sensor network. Let $\mathcal{I}_k = \{\mathbf{y}(0), \mathbf{y}(1), \cdots, \mathbf{y}(k)\}$ denote all the available measurement information up to time instant k.

In this chapter, we consider the risk-sensitive generalization of the MMSE estimation problem. The centralized risk-sensitive estimate $\hat{x}_k$ of x_k given the measurement information $\mathcal{I}_k$ is defined as

$$\hat{x}(k) := \arg \min_{\xi \in \mathbb{R}^n} \mathrm{E}\left[\exp[\theta \Psi_{0,k}(\xi)] | \mathcal{I}_k\right], \tag{5.4}$$

where $\theta > 0$ is the risk-sensitive parameter and $\Psi_{0,k}(\xi)$ is defined as

$$\Psi_{0,k}(\xi) = \hat{\Psi}_{0,k-1} + \tfrac{1}{2}[x(k) - \xi]^\top \pi(k)[x(k) - \xi], \tag{5.5}$$

with symmetric weighting matrix $\pi(k) > 0$ and

$$\hat{\Psi}_{0,k-1} = \tfrac{1}{2}\sum_{t=0}^{k-1}[x(t) - \hat{x}(t)]^\top \pi(t)[x(t) - \hat{x}(t)].$$

The consideration of an exponential-form cost function penalizes all the higher order moments of the estimation error energy, so that the risk-sensitive estimator enjoys robustness characteristics, the precise meaning of which is interpreted in [7]. The centralized risk-sensitive state estimation problem in (5.4) for linear Gaussian systems was solved in [13] utilizing the reference measure approach, where the obtained risk-sensitive filter evolves according to the following recursive form:

$$M(k)^{-1} = P(k)^{-1} + C(k)^\top R(k)^{-1} C(k), \tag{5.6}$$

$$K(k) = M(k)C(k)^\top R(k)^{-1}, \tag{5.7}$$

$$\hat{x}(k) = \bar{x}(k) + K(k)[\mathbf{y}(k) - C(k)\bar{x}(k)], \tag{5.8}$$

$$P(k+1) = A(k)[M(k)^{-1} - \theta\pi(k)]^{-1}A(k)^\top + B(k)Q(k)B(k)^\top, \tag{5.9}$$

$$\bar{x}(k+1) = A(k)\hat{x}(k), \tag{5.10}$$

with $P(0) = P_0$ and $\bar{x}(0) = \mu_0$ if $M(k)^{-1} - \theta\pi(k) > 0$ for $k \geq 0$. The risk-sensitive filter in (5.6)-(5.10) is a centralized filter since the risk-sensitive estimate $\hat{x}_k$ in (5.8) is obtained by minimizing the average of exponential criterion in (5.4) given the output matrix $C(k)$, the covariance $R(k)$ and information $\mathcal{I}_k$ from the entire network. The centralized risk-sensitive filter is linear and in a Kalman-like form, with the difference

in the Riccati equation in (5.9), which is similar to the one obtained in [66] for discrete-time H_∞ filtering. In this work, we answer the following question: how to design a fully distributed risk-sensitive filter for systems that are measured by a sensor network, such that the asymptotic performance of the local estimators can be guaranteed in a certain sense.

5.3   Distributed risk-sensitive filter

Before continuing, we define two network-wide aggregate quantities: the fused inverse-covariance matrices:

$$S(k) = \tfrac{1}{N} C(k)^\top R(k)^{-1} C(k)$$
$$= \tfrac{1}{N} \sum_{i=1}^{M} C_i(k)^\top R_i(k)^{-1} C_i(k), \tag{5.11}$$

and the fused sensor data:

$$d(k) = \tfrac{1}{N} C(k)^\top R(k)^{-1} \mathbf{y}(k) \tag{5.12}$$
$$= \tfrac{1}{N} \sum_{i=1}^{N} C_i(k)^\top R_i(k)^{-1} y_i(k). \tag{5.13}$$

Now we are at the position to present the following decentralized risk-sensitive filter for the sensor network.

Proposition 5.1. *Consider the discrete-time linear Gaussian process in (5.1) measured by a sensor network with N nodes, where the sensing model is described in (5.2). Suppose node $i \in \mathbb{N}_{1:N}$ in the sensor network calculates its local risk-sensitive state estimate using the following update equations if $M_i(k)^{-1} - \frac{\theta}{N}\pi(k) > 0$ for $k \geq 0$*

$$M_i(k)^{-1} = P_i(k)^{-1} + S(k), \tag{5.14}$$
$$\hat{x}_i(k) = \bar{x}_i(k) + M_i(k)[d(k) - S(k)\bar{x}_i(k)], \tag{5.15}$$
$$P_i(k+1) = A(k)[M_i(k)^{-1} - \tfrac{\theta}{N}\pi(k)]^{-1}A(k)^\top + B(k)Q_i(k)B(k)^\top, \tag{5.16}$$
$$\bar{x}_i(k+1) = A(k)\hat{x}_i(k), \tag{5.17}$$

where $\theta > 0$, $Q_i(k) = NQ(k)$, $P_i(0) = NP(0)$ and $\bar{x}_i(0) = \mu_0$. The local risk-sensitive estimate obtained by each sensor is the same as the centralized risk-sensitive estimate in (5.4), namely, $\hat{x}_i(k) = \hat{x}(k)$ for any $i \in \mathbb{N}_{1:N}$.

Proof. For the centralized risk-sensitive estimator, substituting (5.6) into (5.9), we have

$$P(k+1)$$
$$= A(k)[P(k)^{-1} + C(k)^\top R(k)^{-1} C(k) - \theta\pi(k)]^{-1}A(k)^\top + B(k)Q(k)B(k)^\top. \tag{5.18}$$

Substituting (5.7) and (5.10) into (5.8), we have

$$\hat{x}(k) = A(k-1)\hat{x}(k-1) + M(k)C(k)^\top R(k)^{-1}[\mathbf{y}(k) - C(k)A(k-1)\hat{x}(k-1)]. \quad (5.19)$$

We prove the above proposition by mathematical induction. For each sensor node i, we substitute (5.14) into (5.16) and combine $Q_i(k) = NQ(k)$ and the definition of S_k in (5.11), then we have

$$P_i(k+1)$$
$$= A(k)[P_i(k)^{-1} + \tfrac{1}{N}C(k)^\top R(k)^{-1}C(k) - \tfrac{\theta}{N}\pi(k)]^{-1}A(k)^\top + B(k)NQ(k)B(k)^\top. \quad (5.20)$$

Based on the results in (5.18) and (5.20), if $P_i(k) = NP(k)$, we have $P_i(k+1) = NP(k+1)$. By mathematical induction, since the basis $P_i(0) = NP(0)$, we have $P_i(k) = NP(k)$ for $k \geq 0$, thus $M_i(k) = NM(k)$. Combining (5.15), (5.17) and $M_i(k) = NM(k)$, we obtain

$$\hat{x}_i(k)$$
$$= \bar{x}_i(k) + M_i(k)[d(k) - S(k)\bar{x}_i(k)]$$
$$= A(k-1)\hat{x}_i(k-1) + M(k)C(k)^\top R(k)^{-1}[\mathbf{y}(k) - C(k)A(k-1)\hat{x}_i(k-1)], \quad (5.21)$$

where the definitions of $d(k)$ and $S(k)$ are used in the last equation. Since $\bar{x}_i(0) = \bar{x}(0)$, we have $\hat{x}_i(0) = \hat{x}(0)$. By mathematical induction, combining (5.19), (5.21) and the basis $\hat{x}_i(0) = \hat{x}(0)$, we have $\hat{x}_i(k) = \hat{x}(k)$ for any $i \in \mathbb{N}_{1:N}$, which completes the proof. $\qquad\square$

The above proposition presents a decentralized risk-sensitive filter, where each sensor calculates $S(k)$ and $d(k)$. In other words, each sensor is required to have access to the measurement information of the entire sensor network, resulting in the communication complexity of $O(n^2)$. To further develop a distributed risk-sensitive filter, we consider the scenario that each node only communicates information with its neighbors in the network. Let $N_i = \{j : (i,j) \in E\}$ be the set of neighbors of node i on graph G. Similar to the micro-Kalman filter iterations proposed in [47], the results in Proposition 5.1 would hold regardless what data fusion method is used to obtain the fused terms $S(k)$ and $d(k)$. Compared with the decentralized Kalman filter algorithm, the decentralized risk-sensitive filter has different local Riccati update equations and will reduce to the decentralized Kalman filter if the risk-sensitive parameter $\theta = 0$. In addition, the two decentralized filters share the same attempt to obtain the fused terms $S(k)$ and $d(k)$,

which motivates us to utilize the data fusion methods considered in the distributed Kalman filtering.

To develop a distributed risk-sensitive filtering algorithm, we utilize the dynamic consensus data fusion filter proposed in [69] to perform the averaging task and approximately compute $S(k)$ and $d(k)$ in Proposition 5.1. For any sensor node $i \in \mathbb{N}_{1:N}$, the dynamics of the continuous-time high-gain version of the consensus filter is given by

$$\dot{q}_i = \beta \sum_{j \in N_i}(q_j - q_i) + \beta \sum_{j \in N_i}(u_j - u_i),$$
$$p_i = q_i + u_i,$$

where the gain $\beta \sim O(1/\lambda_2)$ and λ_2 is the second-smallest eigenvalue of the Laplacian matrix L of the graph G, which denotes the algebraic connectivity of G. If the topology of the graph G is sparse, the gain β would be relatively large. u_i is the input of the consensus filter of node i, q_i is the state of the consensus filter of node i and p_i is its output. It is proved in [69] that for a connected graph G, the dynamic average consensus of the input $u_i(t)$ can be aymptotically reached, namely, $p_i(t)$ asymptotically converges to $\frac{1}{M}\sum_{i \in \mathbb{N}_{1:M}} u_i(t)$ as $t \to \infty$. By discretization, we obtain the discrete-time version of the above consensus filter in the following form

$$q_i(k) = q_i(k-1) + \epsilon\beta \sum_{j \in N_i}[q_j(k-1) - q_i(k-1)] + \epsilon\beta \sum_{j \in N_i}[u_j(k) - u_i(k)], \quad (5.22)$$
$$p_i(k) = q_i(k) + u_i(k), \quad (5.23)$$

where ϵ is the step-size. By utilizing the discrete-time version of the consensus filter in (5.22) and (5.23) to compute $S(k)$ and $d(k)$, we obtain the distributed risk-sensitive filtering algorithm in Algorithm 1.

Algorithm 1 Distributed risk-sensitive filtering algorithm for node i with consensus-based data fusion

1: Initialization: $q_i(-1) = 0_{n \times 1}$, $X_i(-1) = 0_{n \times n}$ and $P_i(0) = NP_0$, $\bar{x}_i(0) = \mu_0$.
2: For time instant $k \geq 0$, the following steps are implemented at node $i \in V$.
3: Calculate the local data

$$u_i(k) = C_i(k)^\top R_i(k)^{-1} y_i(k), \ U_i(k) = C_i(k)^\top R_i(k)^{-1} C_i(k).$$

4: Node i sends the data packet $(u_i(k), q_i(k-1), U_i(k), X_i(k-1))$ to its neighbors and receives data $(u_j(k), q_j(k-1), U_j(k), X_j(k-1))$ for $j \in N_i$ from its neighbors.
5: Update the consensus filter for the fused measurement data:

$$q_i(k) = q_i(k-1) + \epsilon\beta \sum_{j \in N_i} [(q_j(k-1) - q_i(k-1)) + (u_j(k) - u_i(k))],$$
$$d_i(k) = q_i(k) + u_i(k).$$

6: Update the consensus filter for the fused inverse-covariance:

$$X_i(k) = X_i(k-1) + \epsilon\beta \sum_{j \in N_i} [(X_j(k-1) - X_i(k-1)) + (U_j(k) - U_i(k))], \quad (5.24)$$
$$S_i(k) = X_i(k) + U_i(k). \quad (5.25)$$

7: Estimate the state using the local risk-sensitive filter

$$M_i(k) = (P_i(k)^{-1} + S_i(k))^{-1}, \quad (5.26)$$
$$\hat{x}_i(k) = \bar{x}_i(k) + M_i(k)[d_i(k) - S_i(k)\bar{x}_i(k)]. \quad (5.27)$$

8: Update the local risk-sensitive filter

$$P_i(k+1) = A(k)[M_i(k)^{-1} - \tfrac{\theta}{N}\pi(k)]^{-1} A(k)^\top + B(k)NQ_i(k)B(k)^\top, \quad (5.28)$$
$$\bar{x}_i(k+1) = A(k)\hat{x}_i(k). \quad (5.29)$$

For node $i \in V$, the consensus filters are initialized with $q_i(-1) = 0_{n \times 1}$, $X_i(-1) = 0_{n \times n}$ and the local risk-sensitive filter is initialized with $P_i(0) = NP_0$ and $\bar{x}_i(0) = \mu_0$. At each time instant k, node i sends the data packet $(u_i(k), q_i(k-1), U_i(k), X_i(k-1))$ to its neighbors and receives data $(u_j(k), q_j(k-1), U_j(k), X_j(k-1))$ for $j \in N_i$ from its neighbors, implying that the communicational complexity is scalable for large-scale sensor networks. For node i, at time instant k, the inputs of the consensus filter for the fused measurement data are $u_i(k)$ and $(u_j(k), q_j(k-1))$ for $j \in N_i$ and the inputs of the consensus filter for the fused inverse-covariance are $U_i(k)$ and $(U_j(k), X_j(k-1))$ for $j \in N_i$. The outputs $d_i(k)$ and $S_i(k)$ of the consensus filters at node i are approximations of $d(k)$ and $S(k)$, which are utilized by the local risk-sensitive filter to calculate the state

estimate. In Algorithm 1, ϵ is the step-size of discretization and we refer to Section II.C in [48] for hints on the right choice of the step-size. The proposed distributed risk-sensitive filtering is a generalization of the distributed consensus-based Kalman filtering algorithm (Algorithm 1 in [47]), where the difference lies in the update of $P_i(k)$ at the local filter; the former reduces to the latter if the risk-sensitive parameter $\theta = 0$. Different from the existing results on robust distributed state estimation which are typically formulated as H_∞ filtering problems and involve LMIs in the local filter design, the local risk-sensitive filter of the proposed distributed algorithm is updated in a Riccati-based linear recursive form.

Next, we analyze the asymptotic properties of the local estimators in the proposed distributed risk-sensitive filtering algorithm. To simplify the analysis, we focus on the LTI case where $\pi(k) = \pi$, $A(k) = A$, $Q(k) = Q$, $B(k) = B$, $C_i(k) = C_i$ and $R_i(k) = R_i$ for $i \in V$. For this case, we present the stability results on the distributed risk-sensitive filter proposed in Algorithm 1.

Theorem 5.1. *For the LTI system described in (5.1), which is measured by a sensor network with sensing model in (5.2), each local filter of the distributed risk-sensitive filter proposed in Algorithm 1 is stable as $k \to \infty$ if we choose a risk-sensitive parameter θ such that the centralized risk-sensitive filter is stable as $k \to \infty$.*

Proof. For the centralized filter, substituting (5.8) into (5.10), we have

$$
\begin{aligned}
\bar{x}(k+1) &= A[\bar{x}(k) + K(k)(\mathbf{y}(k) - C\bar{x}(k))] \\
&= [A - AK(k)C]\bar{x}(k) + AK(k)\mathbf{y}(k) \\
&= [A - AM(k)C^\top R^{-1}C]\bar{x}(k) + AK(k)\mathbf{y}(k). \quad (5.30)
\end{aligned}
$$

Since the system is time-invariant, we have $S(k) = S = \frac{1}{N}C_i^\top R_i^{-1}C_i$. Next, we focus on the distributed risk-sensitive filter. According to Algorithm 1, for each node $i \in V$, the consensus filter for the fused inverse-covariance evolves according to (5.24) and (5.25). Due to the fact that the system is time-invariant, we have $U_i(k) = U_i = C_i^\top R_i^{-1}C_i$, thus $S_i(k) = X_i(k) + U_i$. Combining (5.24) and $S_i(k) = X_i(k) + U_i$, we have

$$
S_i(k) = S_i(k-1) + \epsilon\beta \sum_{j \in N_i}[(S_j(k-1) - S_i(k-1))],
$$

where the initial states $S_i(-1) = X_i(-1) + U_i = C_i^\top R_i^{-1}C_i$ since $X_i(-1) = 0_{n \times n}$ for $i \in V$. According to Theorem 2 of [48], the static average consensus of the initial states $S_i(-1)$ can be asymptotically reached, implying that the output $S_i(k)$ of the consensus

111

filter tracks $\frac{1}{N}\sum_{i=1}^{N} S_i(-1) = \frac{1}{N}\sum_{i=1}^{N} C_i^\top R_i^{-1} C_i$ with zero steady-state error, namely,

$$\lim_{k\to\infty} S_i(k) = \frac{1}{N}\sum_{i=1}^{N} C_i^\top R_i^{-1} C_i \qquad (5.31)$$

for $i \in V$. Combining (5.26) and (5.28), we have the Riccati equation of local filter

$$P_i(k+1) = A[P_i(k)^{-1} + S_i(k) - \frac{\theta}{N}\pi]^{-1} A^\top + BNQB^\top. \qquad (5.32)$$

Based on (5.31), (5.32) and the Riccati equation of the centralized filter in (5.18), we can easily obtain $\lim_{k\to\infty} P_i(k) = NP(k)$, thus $\lim_{k\to\infty} M_i(k) = NM(k)$ for the local risk-sensitive filter at node i. Substituting (5.27) into (5.29), we have

$$\begin{aligned}
\bar{x}_i(k+1) &= A[\bar{x}_i(k) + M_i(k)[d_i(k) - S_i(k)\bar{x}_i(k)]] \\
&= [A - AM_i(k)S_i(k)]\bar{x}_i(k) + AM_i(k)d_i(k).
\end{aligned}$$

Based on the facts that $\lim_{k\to\infty} M_i(k) = NM(k)$ and $\lim_{k\to\infty} S_i(k) = \frac{1}{N}\sum_{i=1}^{N} C_i^\top R_i^{-1} C_i$, we have as $k \to \infty$,

$$\begin{aligned}
&\bar{x}_i(k+1) \\
&= [A - AM(k)\sum_{i=1}^{N} C_i^\top R_i^{-1} C_i]\bar{x}_i(k) + ANM(k)d_i(k) \\
&= [A - AM(k)C^\top R^{-1}C]\bar{x}_i(k) + ANM(k)d_i(k). \qquad (5.33)
\end{aligned}$$

From (5.30) and (5.33), we notice the fact that for $i \in V$, the centralized risk-sensitive filter and the local risk-sensitive filters share the same term $[A - AM(k)C^\top R^{-1}C]$ as $k \to \infty$. Therefore, we conclude that each local filter of the distributed risk-sensitive filter proposed in Algorithm 1 is stable as $k \to \infty$ if we choose a risk-sensitive parameter θ such that the centralized risk-sensitive filter is stable as $k \to \infty$. This completes the proof. $\qquad\square$

Remark 5.1. The above theorem presents the stability results of the proposed distributed risk-sensitive filter for the LTI case. According to the results in [39], the centralized risk-sensitive filter in (5.6)-(5.10) is proved to be asymptotically stable if $M(k)^{-1} - \theta\pi(k) > 0$ for all $k > 0$. See [39] for the range of values of the risk-sensitivity parameter θ and the initial error covariance P_0 to ensure that $M(k)^{-1} - \theta\pi(k) > 0$ for all $k > 0$.

The risk-sensitive parameter θ adds one dimension of freedom in the estimator design, thus allows the user to achieve a desired balance between the optimality of the filter under the nominal model scenario and the robustness to the system uncertainty, which will be further illustrated by position tracking examples in Section 5.4.

5.4　Simulation examples

In this section, we show the robustness of the proposed filtering algorithm to system uncertainty. The proposed distributed risk-sensitive filtering algorithm is applied to the target tracking of an object in $\mathbb{R}^2$ considered in [47]. The dynamics of the moving target is given by

$$\dot{x}(t) = (A_0 + \Delta A)x(t) + B_0 w(t), \tag{5.34}$$

with known nominal model

$$A_0 = 2 \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$
$$B_0 = I_2,$$

and unknown model uncertainty

$$\Delta A = 2 \begin{bmatrix} 0 & \varsigma \\ 0 & 0 \end{bmatrix}.$$

The noise w is a zero-mean Gaussian with covariance $Q = 25I_2$. The nominal model (A_0, B_0) of the process describes a target that moves on noisy circular trajectories. The local filters at sensors only know the nominal model (A_0, B_0) of the process, but not the true model $(A_0 + \Delta A, B_0)$. By ZOH discretization, we obtain the true discrete-time model of this moving target

$$x(k+1) = Ax(k) + Bw(k),$$

with sampling period $\epsilon = 0.015$ (≈ 70 Hz) and parameters

$$A = I_2 + \epsilon(A_0 + \Delta A) + \frac{\epsilon^2}{2}(A_0 + \Delta A)^2 + \frac{\epsilon^3}{6}(A_0 + \Delta A)^3,$$
$$B = \epsilon B_0.$$

The parameters of the discrete-time nominal process model (A_N, B_N) can be obtained by following a similar discretization procedure on the nominal continuous-time model (A_0, B_0). The first element and second element of the state $x(k)$ represent the values of the moving point along x-axis and y-axis correspondingly in a Cartesian coordinate system. The initial conditions are

$$\mu_0 = [15, -10], \ P_0 = 10I_2.$$
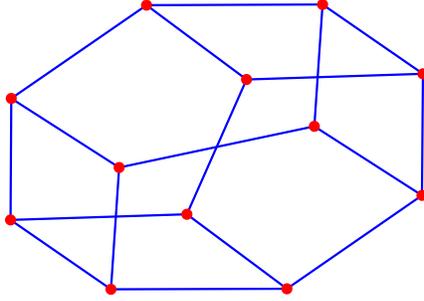
Figure 5.1: Sensor network G_1 with 12 nodes and 18 edges.

To make it easier to distinguish the estimates of different nodes in figures, we first consider a small sensor network in Fig. 5.1, which has $N_1 = 12$ nodes and 18 edges and its topology $G_1 = (V_1, E_1)$ is undirected with $V_1 = \{1, 2, \ldots, N_1\}$ and $E_1 \subset V_1 \times V_1$. Each sensor can only exchange messages with its neighbors. We randomly choose half of the nodes to sense along the horizontal axis (x-axis) and the other half sense along the vertical axis (y-axis), e.g., the measurement model of sensor $i \in V_1$ can be described by

$$y_i(k) = C_i x(k) + v_i(k), \tag{5.35}$$

where either $C_i = C_x = [1 \ 0]$ or $C_i = C_y = [0 \ 1]$ and $v_i(k)$ is a zero-mean Gaussian noise with $R_i = c_v^2 \sqrt{i}$ and $c_v = 30$. The sensor network knows the nominal model of the system, based on which, distributed filtering algorithms would be utilized to estimate the position of the target. We consider the scenario with model uncertainty $\varsigma = -0.15$, where the true trajectory of the target is in a noisy elliptical motion. We set the gain $\beta = 7$, $\pi(k) = I_2$, the risk-sensitive parameter $\theta = 0.002$ and it is verified that $M_i(k)^{-1} - \frac{\theta}{N_1}\pi(k) > 0$ for $k \geq 0$ and $i \in V_1$. For $\varsigma = -0.15$, the comparison of the state estimates of all nodes obtained by the proposed distributed risk-sensitive (distributed RS) filtering algorithm and the distributed Kalman filtering (distributed KF) algorithm with high-pass consensus filtering of the sensed data in [47] is demonstrated in Fig. 5.2. It is shown in Fig. 5.2 that the distributed risk-sensitive filter (marked as green squares) achieves better tracking of the target's true position at time $t = \epsilon k$ (marked as a blue star) compared with the distributd Kalman filter (marked as red circles) when the model uncertainty exists and $\varsigma = -0.15$. To further show the merits of
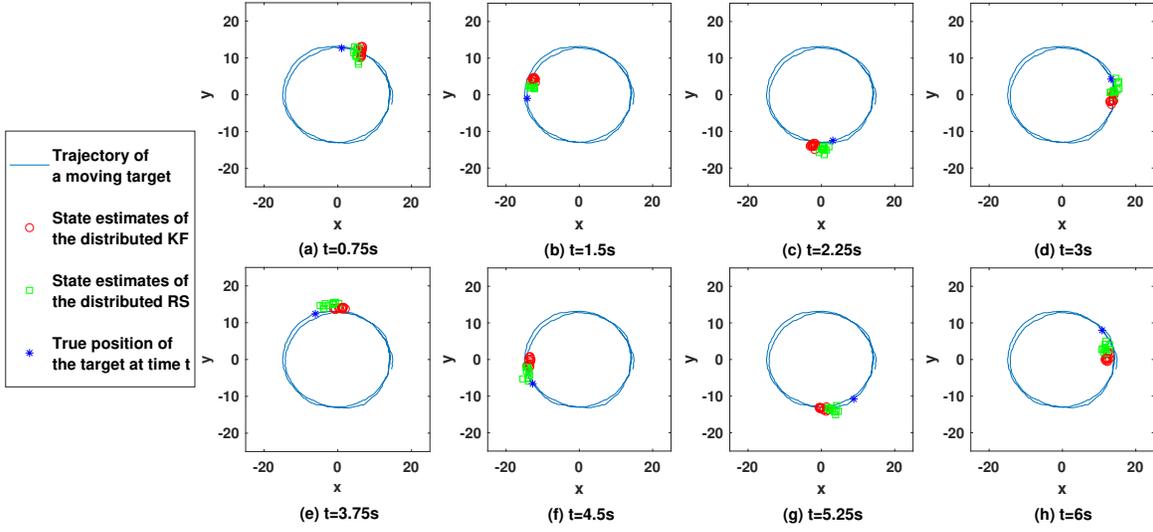
Figure 5.2: Comparison between the state estimates of all nodes in sensor network G_1 for $\varsigma = -0.15$.

the proposed method, the variation of the resultant root mean square errors (RMSEs) of all nodes for the distributed risk-sensitive filter and the distributed Kalman filter as the uncertainty parameter ς in ΔA ranges in value from -0.25 to 0.25, where the other parameters remain the same is shown in Fig. 5.3. Each point in Fig. 5.3 is obtained by averaging 30 random runs for the first 15 seconds, during which period, the condition $M_i(k)^{-1} - \frac{\theta}{N_1}\pi(k) > 0$ for $k \geq 0$ and $i \in V_1$ is verified to hold. From Fig. 5.3, we notice that for the case that the uncertainty parameter ς tends to zero, the distributed Kalman filter achieves a slightly lower value of RMSE compared with the distributed risk-sensitive filter since the Kalman filter is proved to be the optimal estimtor with minimum mean square estimation error for linear Gaussian systems under the nominal scenario; while for larger absolute value of ς (larger system uncertainty), the distributed risk-sensitive filter outperforms the distributed Kalman filter. This indicates that the proposed distributed risk-sensitive filter has good estimation performance under nominal scenarios and acceptable estimation performance that degrades less rapidly than the distributed Kalman filter under non-nominal scenarios; so we conclude that the proposed filter is more robust to system uncertainty compared with the distributed Kalman filter.

Next, we consider a larger-size sensor network in Fig. 5.4, which is randomly generated with $N_2 = 50$ nodes and 250 edges, whose topology $G_2 = (V_2, E_2)$ is undirected and connected with $V_2 = \{1, 2, \ldots, N_2\}$ and $E \subset V_2 \times V_2$. The sensing model of G_2
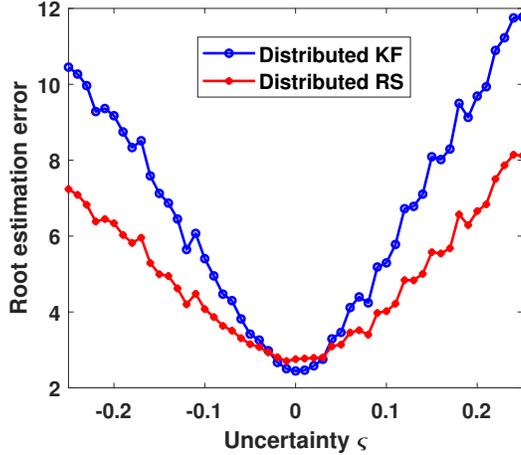
Figure 5.3: RMSE comparison of the proposed ditributed risk-sensitive filter and the distributed Kalman filter with respect to system uncertainty for network G_1.

is similar as G_1, e.g., the measurement model of sensor $i \in V_2$ is described in (5.35) with $R_i = c_v^2 \sqrt{i}$ and $c_v = 30$. We set $\beta = 7$, $\pi(k) = I_2$, $\theta = 0.002$ and verify that $M_i(k)^{-1} - \frac{\theta}{N_2}\pi(k) > 0$ for $k \geq 0$ and $i \in V_2$. For $\varsigma = -0.15$, the estimation performance comparison of the two distributed filtering algorithm is shown in Fig. 5.5, where the curve is obtained by averaging 10 random runs of each algorithm. The RMSEs for the distributed risk-sensitive filter and the distributed Kalman filter from 0 to 15 seconds are 4.9374 and 5.5924, respectively, indicating that the proposed filter outperforms the distributed Kalman filter when $\varsigma = -0.15$. It is shown in Fig. 5.5 that the estimation performance of the two filters are close for the first 3 seconds since the model mismatches of the system matrix is not yet fully reflected in the system dynamics; for $t > 3$, the proposed distributed risk-sensitive filter always outperforms the distributed Kalman filter though the two filters both have relatively large RMSEs due to the existence of system parameter uncertainty.

5.5 Conclusion

In this chapter, we develop a distributed risk-sensitive filtering algorithm for linear Gaussian systems measured by a sensor network. The consensus filters are utilized to obtain the average of the sensor data and the inverse-covariance matrices, based on which, the filter at each sensor node obtains its local state estimate. The stability analysis of the proposed distributed risk-sensitive filter for LTI systems is included. Simulation results of tracking a moving target in $\mathbb{R}^2$ show that the proposed distributed
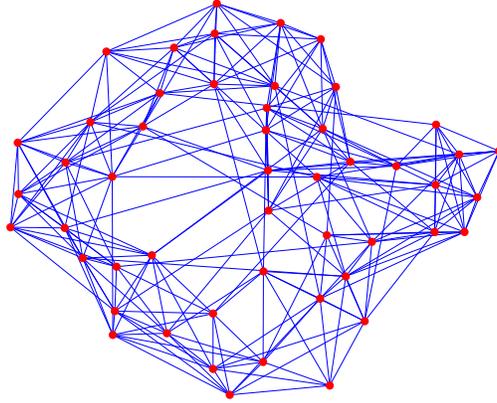
Figure 5.4: Sensor network G_2 with 50 nodes and 250 edges.
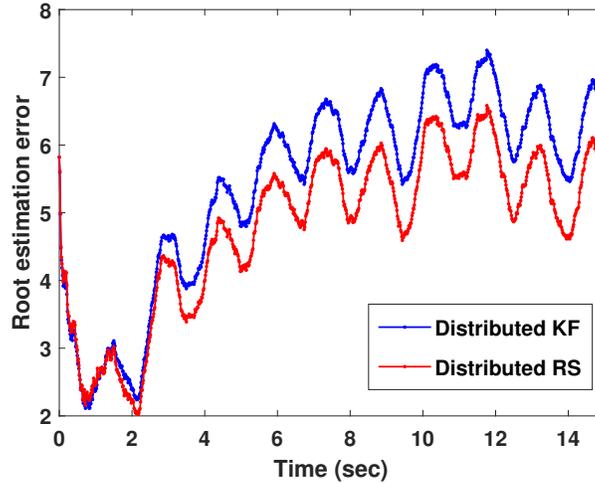
Figure 5.5: Comparison of RMSEs of the proposed distributed risk-sensitive filter and the distributed Kalman filter for sensor network G_2 with $\varsigma = -0.15$.

risk-sensitive filter has good estimation performance under nominal scenarios, and acceptable estimation performance that degrades less rapidly than the distributed Kalman filter under non-nominal scenarios, implying better robustness of the proposed filter to system uncertainty than that of the distributed Kalman filter. The consensus data fusion algorithm considered in this chapter is an approximate algorithm, which is just one option for the data fusion task. In the upcoming research, we will focus on finding the upper bound of the estimation errors of the proposed distributed filtering algorithm and explore various consensus-based data fusion methods to achieve better estimation performance.

Chapter 6

Conclusions and future work

This chapter concludes the thesis. A summary of the main findings is presented in Section 6.1, while Section 6.2 contains suggestions for future research.

6.1 Conclusions

This thesis investigates optimal state estimation problems for CPS mainly from two perspectives: the first aims at maintaining system performance with limited resources; the second targets at robust state estimation for systems with model uncertainties. Estimators are derived by minimizing a certain cost criterion, either the MSE cost or the exponential form cost. The obtained results are mostly in recursive forms, which are easy to implement and asymptotic properties of some proposed estimators are analyzed. The outcomes of the studies in this thesis are summarized as follows:

1. The MMSE estimator for hidden Markov models with energy harvesting sensors equipped with stochastic energy-based event-schedulers is derived, where the simulation examples show that the utilization of the energy-based event-triggering condition leads to smarter sensor energy management in reducing the occurrence rate of the sensor's sleeping mode. In addition, the structure of the state estimator for hidden Markov models with dynamic event-triggers is given for unreliable communication channels.

2. An event-triggered MMSE state estimator is proposed for a linear Gaussian system with an energy harvesting sensor, where a stochastic energy-dependent event-triggering transmission protocol is proposed to balance the communication rate and estimation performance according to the sensor's battery energy. Also, the

relationship between the average communication rate and energy harvesting rate is discussed.

3. Closed-form expressions of the risk-sensitive event-triggered posterior and prior estimates for linear Gaussian systems with a stochastic event-triggering condition are presented, which are shown to evolve in recursive Kalman-like structures. Two sufficient stability conditions for the proposed risk-sensitive event-triggered estimators are given. Moreover, by treating the unknown inputs as a process modeled by a non-informative prior, we obtain the closed-form expression of the risk-sensitive state estimate for discrete-time systems with unknown exogenous inputs.

4. A distributed robust state estimation algorithm is proposed for a linear Gaussian system measured by a sensor network, where the high-gain dynamic consensus filter is utilized to compute the fused data. It is shown that the local filter at each sensor is updated in a Riccati-based linear recursive form. Furthermore, the asymptotic stability of local estimators for linear time-invariant systems are discussed.

The effectiveness of the proposed methods is validated by simulation examples.

6.2   Future work

The results achieved so far in this thesis are of importance for improving the estimation performance by exploiting the information contained in event-triggering conditions and counteracting against system uncertainties. However, there remain many problems to be studied in the general area of optimal state estimation of CPS. The following promising directions deserve efforts for future work.

6.2.1   Event-triggered state estimation for linear Gaussian systems with a lossy communication channel

The effect of packet dropout is considered in the event-triggered state estimation for hidden Markov models in Section 2.2, but not for linear Gaussian systems. The existence of packet dropout would ruin the Gaussianity of the conditional probability distribution of the state, which makes it challenging to derive the closed-form expression of the optimal event-triggered state estimate. The reference measure approach may be

119

utilized to solve this problem and the main difficulty lies in the construction of a map from the "real-world" to a new probability measure and the derivation of the conditional probability density function of the state under the new probability measure.

6.2.2 Risk-sensitive event-triggered estimation for systems with multiple sensors

As the developments of applications on sensor networks, it is of vital importance to extend the results in Chapter 4 to systems with multiple sensors, where each sensor has its own event-triggering condition. The main goal is to obtain the optimal fusion algorithm of hybrid event-triggered measurement information and the structure of the risk-sensitive event-triggered state estimator.

6.2.3 Estimation performance analysis of the distributed risk-sensitive filtering algorithm

For the proposed distributed risk-sensitive filtering algorithm in Chapter 5, one future research direction is to analyze its asymptotic properties, e.g., find the upper bound of the estimation errors of local filters in the proposed distributed algorithm. Also, we may explore various consensus-based data fusion methods to achieve better estimation performance.

6.2.4 Risk-sensitive control of systems with event-triggered scheduling

The risk-sensitive approach is utilized to solve the robust event-triggered state estimation problems in this thesis. We may further consider the risk-sensitive closed-loop control of systems with event-triggered scheduling, where the control input is transmitted to the actuator through a communication channel only if a predefined event-triggering condition is violated. Utilizing the reference measure approach, we target at obtaining the event-triggered risk-sensitive control law, which minimizes the cost criterion in an exponential form.

Bibliography

[1] J. H. Ahrens and H. K. Khalil. High-gain observers in the presence of measurement noise: A switched-gain approach. *Automatica*, 45(4):936–943, 2009.

[2] T. Al-ani. *Hidden Markov Models in Dynamic System Modelling and Diagnosis*. INTECH Open Access Publisher, 2011.

[3] B. Anderson and J. B. Moore. Detectability and stabilizability of time-varying discrete-time linear systems. *SIAM Journal on Control and Optimization*, 19(1):20–32, 1981.

[4] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang. What will 5G be? *IEEE Journal on Selected Areas in Communications*, 32(6):1065–1082, 2014.

[5] E. Arslan, R. Vadivel, M. S. Ali, and S. Arik. Event-triggered H_∞ filtering for delayed neural networks via sampled-data. *Neural Networks*, 91:11–21, 2017.

[6] P. Baronti, P. Pillai, V. W. Chook, S. Chessa, A. Gotta, and Y. F. Hu. Wireless sensor networks: A survey on the state of the art and the 802.15.4 and zigbee standards. *Computer Communications*, 30(7):1655–1695, 2007.

[7] R. K. Boel, M. R. James, and I. R. Petersen. Robustness and risk-sensitive filtering. *IEEE Transactions on Automatic Control*, 47(3):451–461, 2002.

[8] W. Chen, D. Shi, J. Wang, and L. Shi. Event-triggered state estimation: Experimental performance assessment and comparative study. *IEEE Transactions on Control Systems Technology*, 25(5):1865–1872, 2017.

[9] P. Corke, T. Wark, R. Jurdak, W. Hu, P. Valencia, and D. Moore. Environmental wireless sensor networks. *Proceedings of the IEEE*, 98(11):1903–1917, 2010.

[10] E. Dahlman, S. Parkvall, and J. Skold. *4G: LTE/LTE-Advanced for Mobile Broadband*. Academic press, 2013.

[11] S. Dey and C. D. Charalambous. Discrete-time risk-sensitive filters with non-gaussian initial conditions and their ergodic properties. *Asian Journal of Control*, 3(4):262–271, 2001.

[12] S. Dey, R. Elliott, and J. B. Moore. Finite-dimensional risk-sensitive filtering for continuous-time nonlinear systems. In *Proceedings of the European Control Conference*, pages 617–622, 1997.

[13] S. Dey and J. B. Moore. Risk-sensitive filtering and smoothing via reference probability methods. In *Proceedings of the American Control Conference*, volume 1, pages 129–133, 1995.

[14] S. Dey and J. B. Moore. Risk-sensitive filtering and smoothing via reference probability methods. *IEEE Transactions on Automatic Control*, 42(11):1587–1591, 1997.

[15] H. S. Dhillon, Y. Li, P. Nuggehalli, Z. Pi, and J. G. Andrews. Fundamentals of base station availability in cellular networks with energy harvesting. In *Proceedings of the IEEE Global Communications Conference*, pages 4110–4115, 2013.

[16] H. Dong, Z. Wang, S. X. Ding, and H. Gao. Event-based H_∞ filter design for a class of nonlinear time-varying systems with fading channels and multiplicative noises. *IEEE Transactions on Signal Processing*, 63(13):3387–3395, 2015.

[17] R. J. Elliott, L. Aggoun, and J. B. Moore. *Hidden Markov Models: Estimation and Control*. Springer, 1995.

[18] A. C. Fu, E. Modiano, and J. N. Tsitsiklis. Optimal energy allocation and admission control for communications satellites. *IEEE/ACM Transactions on Networking*, 11(3):488–500, 2003.

[19] 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.

[20] M. Gatzianas, L. Georgiadis, and L. Tassiulas. Control of wireless networks with rechargeable batteries. *IEEE Transactions on Wireless Communications*, 9(2):581–593, 2010.

[21] X. Ge and Q.-L. Han. Distributed event-triggered H_∞ filtering over sensor networks with communication delays. *Information Sciences*, 291:128–142, 2015.

[22] P. Gyorke and B. Pataki. Energy-aware measurement scheduling in WSNs used in AAL applications. *IEEE Transactions on Instrumentation and Measurement*, 62(5):1318–1325, 2013.

[23] D. Han, Y. Mo, J. Wu, S. Weerakkody, B. Sinopoli, and L. Shi. Stochastic event-triggered sensor schedule for remote state estimation. *IEEE Transactions on Automatic Control*, 60(10):2661–2675, 2015.

[24] 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.

[25] S. Hirche and M. Buss. Human-oriented control for haptic teleoperation. *Proceedings of the IEEE*, 100(3):623–647, 2012.

[26] S. Hu and D. Yue. Event-based H_∞ filtering for networked system with communication delay. *Signal Processing*, 92(9):2029–2039, 2012.

[27] N. Jaggi, K. Kar, and A. Krishnamurthy. Rechargeable sensor activation under temporally correlated events. *Wireless Networks*, 15(5):619–635, 2009.

[28] M. R. James, J. S. Baras, and R. J. Elliott. Risk-sensitive control and dynamic games for partially observed discrete-time nonlinear systems. *IEEE transactions on Automatic Control*, 39(4):780–792, 1994.

[29] A. H. Jazwinski. *Stochastic processes and filtering theory*. Courier Corporation, 2007.

[30] S. J. Julier and J. K. Uhlmann. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92(3):401–422, 2004.

[31] T. Kailath, A. H. Sayed, and B. Hassibi. *Linear Estimation*. Prentice Hall Upper Saddle River, NJ, 2000.

[32] R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1):35–45, 1960.

[33] P. K. Kitanidis. Unbiased minimum-variance linear state estimation. *Automatica*, 23(6):775–778, 1987.

[34] P. R. Kumar and P. Varaiya. *Stochastic Systems: Estimation, Identification and Adaptive Control*. Prentice Hall, 1986.

[35] E. A. Lee. Cyber physical systems: Design challenges. In *Proceedings of the 11th IEEE International Symposium on Object and Component-Oriented Real-Time Distributed Computing*, pages 363–369, 2008.

[36] I. Lee, O. Sokolsky, S. Chen, J. Hatcliff, E. Jee, B. Kim, A. King, M. Mullen-Fortino, S. Park, A. Roederer, et al. Challenges and research directions in medical cyber–physical systems. *Proceedings of the IEEE*, 100(1):75–90, 2012.

[37] J.-S. Lee, Y.-W. Su, and C.-C. Shen. A comparative study of wireless protocols: Bluetooth, UWB, ZigBee, and Wi-Fi. In *Proceedings of the 33rd Annual Conference of the IEEE Industrial Electronics Society*, pages 46–51, 2007.

[38] S. Lee, W. Liu, and I. Hwang. Markov chain approximation algorithm for event-based state estimation. *IEEE Transactions on Control Systems Technology*, 23(3):1123–1130, 2015.

[39] B. C. Levy and M. Zorzi. A contraction analysis of the convergence of risk-sensitive filters. *SIAM Journal on Control and Optimization*, 54(4):2154–2173, 2016.

[40] B. Li. State estimation with partially observed inputs: A unified Kalman filtering approach. *Automatica*, 49(3):816–820, 2013.

[41] P. Marwedel. *Embedded System Design: Embedded Systems Foundations of Cyber-Physical Systems*. Springer Science & Business Media, 2010.

[42] X. Meng and T. Chen. Event triggered robust filter design for discrete-time systems. *IET Control Theory & Applications*, 8(2):104–113, 2014.

[43] M. Mitici, J. Goseling, M. de Graaf, and R. J. Boucherie. Data retrieval time for energy harvesting wireless sensor networks. *Ad Hoc Networks*, 53:32–40, 2016.

[44] Y. Mo, T. H.-J. Kim, K. Brancik, D. Dickinson, H. Lee, A. Perrig, and B. Sinopoli. Cyber–physical security of a smart grid infrastructure. *Proceedings of the IEEE*, 100(1):195–209, 2012.

[45] A. Molin and S. Hirche. An iterative algorithm for optimal event-triggered estimation. *IFAC Proceedings Volumes*, 45(9):64–69, 2012.

[46] A. Nayyar, T. Basar, D. Teneketzis, and V. V. Veeravalli. Optimal strategies for communication and remote estimation with an energy harvesting sensor. *IEEE Transactions on Automatic Control*, 58(9):2246–2260, 2013.

[47] R. Olfati-Saber. Distributed Kalman filtering for sensor networks. In *Proceedings of the 46th IEEE Conference on Decision and Control*, pages 5492–5498, 2007.

[48] R. Olfati-Saber, J. A. Fax, and R. M. Murray. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95(1):215–233, 2007.

[49] U. Orguner and M. Demirekler. Risk-sensitive filtering for jump Markov linear systems. *Automatica*, 44(1):109–118, 2008.

[50] M. Rabi, G. Moustakides, and J. Baras. Multiple sampling for estimation on a finite horizon. In *Proceedings of the 45th IEEE Conference on Decision and Control*, pages 1351–1357, 2006.

[51] M. Rabi, G. Moustakides, and J. Baras. Adaptive sampling for linear state estimation. *SIAM Journal on Control and Optimization*, 50(2):672–702, 2012.

[52] R. R. Rajkumar, I. Lee, L. Sha, and J. Stankovic. Cyber-physical systems: the next computing revolution. In *Proceedings of the 47th Design Automation Conference*, pages 731–736. ACM, 2010.

[53] V. R. Ramezani and S. I. Marcus. Estimation of hidden Markov models: Risk-sensitive filter banks and qualitative analysis of their sample paths. *IEEE Transactions on Automatic Control*, 47(12):1999–2009, 2002.

[54] C. M. Roberts. Radio frequency identification (RFID). *Computers & Security*, 25(1):18–26, 2006.

[55] D. S. Rosenberg. *Completing the Square.* Available at `https://davidrosenberg.github.io/mlcourse/Archive/2017/Notes/completing-the-square.pdf`, 2017.

[56] S. Sadhu, S. Bhaumik, A. Doucet, and T. K. Ghoshal. Particle-method-based formulation of risk-sensitive filter. *Signal Processing*, 89(3):314–319, 2009.

[57] D. Shi, T. Chen, and M. Darouach. Event-based state estimation of linear dynamic systems with unknown exogenous inputs. *Automatica*, 69:275–288, 2016.

[58] D. Shi, T. Chen, and L. Shi. An event-triggered approach to state estimation with multiple point-and set-valued measurements. *Automatica*, 50(6):1641–1648, 2014.

[59] D. Shi, T. Chen, and L. Shi. Event-triggered maximum likelihood state estimation. *Automatica*, 50(1):247–254, 2014.

[60] D. Shi, R. J. Elliott, and T. Chen. Event-based state estimation of discrete-state hidden Markov models. *Automatica*, 65:12–26, 2016.

[61] D. Shi, R. J. Elliott, and T. Chen. On finite-state stochastic modeling and secure estimation of cyber-physical systems. *IEEE Transactions on Automatic Control*, 62(1):65–80, 2017.

[62] D. Shi, L. Shi, and T. Chen. *Event-Based State Estimation: A Stochastic Perspective*. Springer, 2016.

[63] J. Sijs and M. Lazar. Event based state estimation with time synchronous updates. *IEEE Transactions on Automatic Control*, 57(10):2650 –2655, 2012.

[64] J. Sijs, B. Noack, and U. Hanebeck. Event-based state estimation with negative information. In *Proceedings of the 16th International Conference on Information Fusion*, pages 2192–2199, 2013.

[65] P. Sikka, P. Corke, and L. Overs. Wireless sensor devices for animal tracking and control. In *Proceedings of the 29th Annual IEEE International Conference on Local Computer Networks*, pages 446–454, 2004.

[66] D. Simon. *Optimal State Estimation: Kalman, H_∞, and Nonlinear Approaches*. John Wiley & Sons, 2006.

[67] 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, 2004.

[68] Y. Song, M. Zhao, W. Zhou, and H. Han. Throughput-optimal user association in energy harvesting relay-assisted cellular networks. In *Proceedings of the IEEE 6th International Conference on Wireless Communications and Signal Processing*, pages 1–6, 2014.

[69] D. P. Spanos, R. Olfati-Saber, and R. M. Murray. Dynamic consensus on mobile networks. In *Proceedings of IFAC World Congress*, pages 1–6, 2005.

[70] J. L. Speyer, C. Fan, and R. N. Banavar. Optimal stochastic estimation with exponential cost criteria. In *Proceedings of the 31st IEEE Conference on Decision and Control*, pages 2293–2299, 1992.

[71] A. Sultan. Sensing and transmit energy optimization for an energy harvesting cognitive radio. *IEEE Wireless Communications Letters*, 1(5):500–503, 2012.

[72] S. Trimpe. Event-based state estimation with switching static-gain observers. *IFAC Proceedings Volumes*, 45(26):91–96, 2012.

[73] S. Trimpe. Stability analysis of distributed event-based state estimation. In *Proceedings of the IEEE 53rd Annual Conference on Decision and Control*, pages 2013–2019, 2014.

[74] S. Trimpe and R. D'Andrea. An experimental demonstration of a distributed and event-based state estimation algorithm. *IFAC Proceedings Volumes*, 44(1):8811–8818, 2011.

[75] J. Weimer, J. Araújo, and K. H. Johansson. Distributed event-triggered estimation in networked systems. *IFAC Proceedings Volumes*, 45(9):178–185, 2012.

[76] J. L. Willems. *Stability Theory of Dynamical Systems*. Nelson, 1970.

[77] J. Wu, Q.-S. Jia, K. H. Johansson, and L. Shi. Event-based sensor data scheduling: trade-off between communication rate and estimation quality. *IEEE Transactions on Automatic Control*, 58(4):1041–1046, 2013.

[78] J. Yick, B. Mukherjee, and D. Ghosal. Wireless sensor network survey. *Computer networks*, 52(12):2292–2330, 2008.

[79] S. S. H. Zaidi, S. Aviyente, M. Salman, K.-K. 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, 2011.

[80] H. Zhang, L. Xie, and Y. C. Soh. Risk-sensitive filtering, prediction and smoothing for discrete-time singular systems. *Automatica*, 39(1):57–66, 2003.

[81] S. Zhang, N. Zhang, S. Zhou, Z. Niu, and X. S. Shen. *Wireless Traffic Steering for Green Cellular Networks.* Springer, 2016.

[82] X.-M. Zhang and Q.-L. Han. Event-based H_∞ filtering for sampled-data systems. *Automatica*, 51:55–69, 2015.

[83] 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, 2012.

[84] L. Zou, Z. Wang, H. Gao, and X. Liu. Event-triggered state estimation for complex networks with mixed time delays via sampled data information: The continuous-time case. *IEEE Transactions on Cybernetics*, 45(12):2804–2815, 2015.