論文 / 著書情報
Article / Book Information

題目(和文)	外れ値環境下におけるロバスト推定および制御に関する研究
Title(English)	A Study of Robust Estimation and Control under Outliers
著者(和文)	金田泰昌
Author(English)	Yasuaki Kaneda
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第9452号, 授与年月日:2014年3月26日, 学位の種別:課程博士, 審査員:山北　昌毅,三平　満司,大山　真司,倉林　大輔,早川　朋久
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第9452号, Conferred date:2014/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

A Study of Robust Estimation and Control under Outliers

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Engineering

by

Yasuaki Kaneda

Supervisor

Associated Professor Masaki Yamakita

Graduate School of Science and Engineering
Department of Mechanical and Control Engineering
Tokyo Institute of Technology

February, 2014

Acknowledgments

I would like to express my deepest gratitude to my supervisor Associated Professor Masaki Yamakita for his guidances and supports in many things. I would like to thank Sasebo National College of Technology Lecture Teruyoshi Sadahiro who guided me in my entrance to graduate school of Tokyo Institute of Technology and have advised many things. I would also like to thank to committees, Professor Mitsuji Sampei, Associated Professor Shinji Ohyama, Associated Professor Daisuke Kurabayashi, and Associated Professor Tomohisa Hayakawa, who kindly spent time checking and approving this dissertation.

I would like to thank Tokyo Metropolitan Industrial Technology Research Institute and all members of Information Technology group. President Masatoshi Kataoka has provided me with an opportunity to apply for the graduate school for working adults program. Director Hirofumi Sawachika and Management Researcher Kazumi Sakamaki have consulted my convenience to take a degree. Senior Researcher Yasuharu Irizuki has been discussing about my research and supporting me. Special thanks to Dr. Yuji Takeda, Dr. Mamoru Ohara, Mr. Yoshitsugu Nakagawa, Mr. Norihiro Ohira, Mr. Tadashi Okabe, Dr. Takashi Yamaguchi, Dr. Masayuki Murakami, and Mr. Shinichi Tomiyama.

I am deeply grateful to all members of the research meeting "ATACS". Tokyo Denki University President Katsuhisa Furuta, Emeritus Professor of Tokyo Institute of Technology, gave me very important comments and suggestions though he was very busy. Special thanks to Tokyo Denki Univ. Prof. Shoshiro Hatakeyama, Prof. Tetsuo Shiotsuki, Prof. Jun Ishikawa, Waseda Univ. Prof. Harutoshi Ogai, Tokyo Univ. of Tech. Prof. Yasuhiro Ohyama, Prof. Jin-Hua She, Kumamoto Univ. Prof. Nobutomo Matsunaga, Tokyo City Univ. Prof. Kenichiro Nonaka, Kyushu Inst. of Tech. Prof. Masanobu Koga, Osaka Univ. Associated Prof. Masato Ishikawa, Tokyo Denki Univ. Associated Prof. Satoshi Suzuki, Associated Prof. Masami Iwase, Associated Prof. Norihiro Kamamichi, Univ. of Yamanashi Assistant Prof. Koji Makino, Tokyo Univ. of Tech. Assistant Prof. Kaoru Mitsuhashi, Kumamoto Univ. Assistant Prof. Hiroshi Okajima, Tokyo Denki Univ. Assistant Prof. Masaki Izutsu, Tokyo City Univ. Assistant Prof. Kazuma Sekiguchi, and Tokyo Inst. of Tech. Assistant Prof. Tatsuya Ibuki. They provided me with many comments and allowed me to grow.

I would like to thank everyone in Yamakita laboratory for supporting and helping me succeed my degree. Ms. Mari Kobayasi, very kind secretary, has helped me a lot in every kinds of documents. Kyushu Inst. of Tech. Assistant Prof. Yuta Hanazawa has been discussing about my research and suggesting new ideas for me. Dr. Sirichai Pornsarayouth and Mr. Srang Sarot have also been discussing many things in English, so they are my English teachers. Mr. Kazuhiro Tanaka and Mr. Tadashi Sumioka have stimulated my interests, and Mr. Yusuke Yashiro has influenced my research. Mr. Masahiro Kawaguchi and Mr. Hiroyuki Suda were good pair in Yamakita laboratory and have made me pleasure.

Ms. Suthira Limkul has provided relaxed atmosphere in the laboratory. Mr. Kazuyoshi Odachi , Mr. Hiroyuki Oyama, Mr. Terumitsu Hayashi, and Mr. Hiroaki Ishiyama have been talking about many things in my public and private life, and they have made me relaxed in the laboratory. Mr. Mikiya Hara, Mr. Yuto Noda, Mr. Yu Iemura, Mr. Umihiko Amanuma, Mr. Yongjae Kim, Mr. Hiroki Fujii, Mr. Yusuke Yaginuma, and Mr. Mamoru Watanabe have made me happy in the laboratory. Not to mention many friends who have been supporting me.

Finally, I want to thank my family. Especially, I would like to offer my special thanks to my wife and son, Kaori Yamamoto and Koushi Yamamoto. They have encouraged me to study more and made me happy in my life.

To them, I am eternally grateful. Again, thank you.

Yasuaki Kaneda

Oh-okayama, Tokyo
December, 2013

Abstract

Outliers are a kind of non-Gaussian measurement noise generated by heavier tailed distributions than a normal distribution. Hence, abnormal values, which are distant so much from mean values of distributions, are unusually occurred in a time domain. In other words, the outliers are contained in measurements infrequently and their values can usually be considered as zero, so it can be said that the outliers tend to be sparse. They are happened in many applications, and they provide negative effects on various fields. In control engineering, these outliers deteriorate state estimates and control performances. For example, target tracking systems using radar measurement, visual feedback systems, wireless sensor network systems, networked control systems, and so on. Therefore, control systems require robust estimators and controllers under outliers. We propose a practical robust estimation method and control strategy under outliers based on robust Kalman filter (RKF) via l_1 regression. In addition, we analyze performances of the proposed methods, and the effectiveness is demonstrated by some numerical simulations.

RKF via l_1 regression is one of the most attractive reduction methods of effects of outliers due to an easy structure and implementation. Additionally, Since the RKF truncates outliers by some thresholds, it has less delay than the other RKF. However, regularization parameters of the RKF need to be tuned by some heuristic design methods. First of all, we propose a new design method of the RKF. Both primal and dual problems can derive a condition of the proposed parameters, and it is shown that statistics of Gaussian noise determine the parameters of the RKF. This means that the proposed design method provides the parameters with physical meanings, and we can design the parameters systematically. It is also shown that a covariance matrix of an innovation of the RKF is bounded by that of normal Kalman filter (KF) without outliers. The covariance matrix of the innovation of the RKF comes close to an ideal one under outliers. The RKF with the proposed design method is applied to a target tracking system under clutters and two-wheeled vehicle control with outliers.

The RKF can be formulated as a l_1 optimization problem. In general, the optimization problem cannot be solved analytically, and some numerical iterative methods are needed. A convergence rate and accuracy of the solutions of the RKF depend on conditions of the iterations. Secondly, we propose a closed form solution of the RKF by an approximation of its optimal solution, and it gives a fast algorithm. The approximated solution can be calculated by upper and lower bounds of the optimal solution. In addition, an estimation error of the approximated solution is analyzed. It is shown that the proposed algorithm has almost same performances as KF without outliers under some conditions.

Moreover, in order to construct a robust controller under outliers, we apply an idea of the RKF to self-tuning controller (STC), and we propose a robust STC (RSTC) under outliers. A parameter update law of the conventional STC can be written as a recursive least squares

(RLS) estimation, and RLS estimation can be given by a solution of a minimization problem of estimated errors. Therefore, the proposed method estimates parameters and outliers explicitly by addition of a l_1 regression term to the minimization problem, and the estimated outliers are removed from measurement outputs in a controller. The proposed method is solved in a closed form because of a l_1 optimization problem with a single variable, so the algorithm is very efficient. In order to guarantee a stability of the controller, it is required not only to reduce effects of the outliers, but also to analyze performances of the reduction method. We analyze control performances of the proposed method under outliers, and it is shown that steady state errors in the proposed RSTC are nearly equal to ones in the conventional STC without outliers.

For nonlinear systems, extended KF (EKF) is often used to extend the aforementioned methods. However, EKF needs to compute Jacobians of the nonlinear systems and yields unstable solutions numerically. Gaussian sum filter and Particle filter are other famous KF for nonlinear systems and non-Gaussian measurement noise including outliers. They can approximate arbitrary distributions and can provide global optimal estimates. However, it takes so long time to compute the algorithms, and it is unsuitable for real time applications. Finally, we extend the RKF to nonlinear systems by using unscented KF (UKF), and we propose a robust UKF (RUKF). We also propose a new design method of its regularization parameters. Similarly to linear systems, it is shown that statistics of Gaussian measurement noise determine the parameters of RUKF, and we can design the parameters systematically. And also, the proposed design method provides the parameters with physical meanings. Moreover, the regularization parameters make performances of RUKF come close to ones of UKF without outliers. Since RUKF is based on UKF and l_1 optimization problem, it can be computed more efficiently than Gaussian sum filter and particle filter.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background and Motivation

Hawkins defined outliers in [1], and he said "*An outlier is an observation which deviates so much from the other observations as to arouse suspicions that it was generated by a different mechanism*". In more detail, outliers are values which are so far from a rest of a group of values in a sample space, and they are generated by a different probability density function (PDF) from the rest of the group. In the Hawkins' definition of outliers, observations containing some measurement noise are called outliers, and a difference of sets of observations defines outliers. However, in this dissertation, term "outliers" means the only measurement noise, and two sets of measurement noise are considered, i.e., sets of "normal" and "abnormal" measurement noise. The normal measurement noise can usually be described as stochastic variables generated by a normal distribution. On the other hand, the outliers are a kind of non-Gaussian measurement noise generated by heavier tailed distributions than a normal distribution, and abnormal values, which are distant so much from mean values of distributions, are unusually occurred in a time domain. In other words, the outliers are contained in measurements infrequently and their values can usually be considered as zero, so it can be said that the outliers tend to be sparse.

The outliers give negative effects on various fields [2, 3]. For example, fraud detection refers to detection of criminal activities occurring in commercial organizations such as banks, credit card companies, insurance agencies, cell phone companies, stock market, and so on. These frauds can be regarded as outliers. Intrusion detection refers to detection of malicious activity in a computer related system, and it is also one of outlier detection methodologies. Fault diagnosis refers to industrial damage detection, and it is also one of very important problems of outlier detection.

Also in control engineering, outliers are often happened and provide control systems with negative effects. For example, external environments introduce outliers into control systems using non-contact sensors, e.g., radar measurements, global positioning system (GPS), ultrasonic wave sensors, image measurements, and so on. In target tracking systems using radar measurements [4, 5, 6], outliers are occurred on distance and angle information due to reflection noise, and the outliers are called clutter. In unmanned aerial vehicle (UAV) using visual feedback [7], temporary change of image contrast in background causes outliers of position data. In unmanned ground vehicle (UGV) using GPS [8], radio disturbances due to some obstacles provide position data with outliers. Moreover, wireless sensor network systems include outliers, and detection methods have been proposed [9]. Networked control systems (NCSs) [10] refer to control systems wherein the control loops are closed through a network. When network is unstable or has variable delay to send and receive commands, NCSs have intermittent observations and the observations can be regarded as outliers. Since Kalman filter (KF) is well known as a linear minimum variance estimator for linear systems with Gaussian assumptions, performances of KF cannot be guaranteed under the situations. In [11], in order to use KF in NCSs, a stability of KF with intermittent observations was analyzed. System identifications also deal with missing data [12] and nonuniform sampling data [13], and these data can also be regarded as outliers. These outliers deteriorate state estimates and control performances, so these systems require robust estimators and controllers under outliers.

For state estimations under outliers, many methods have been proposed [14]–[20], [34]–[36],[49]. For example, in [14], for a system identification, multiplication of a median of past N-th data and some coefficients gives a threshold of outliers, and the outliers are detected and removed. The method requires many past data, and the threshold is determined by trial and error. In [15], a reduction method of outliers using sliding mode has been proposed. However, parameters of the method are designed by heuristic methods, and chattering is introduced in digital implementations.

KF for non-Gaussian measurement noise including outliers has been proposed, and the methods are called robust KF (RKF) [16]–[20]. For linear systems, in [16, 17], Bayesian model is introduced to KF, and expectation maximization (EM) algorithm is used. In [18], use of variational Bayesian method gives approximations of joint posterior distributions of state and noise variances to realize a low computational cost. In [19], the method also learns the covariance matrix of measurement noise by iterations to compute a Kalman gain. These methods learn their parameters automatically and adjust their parameters for outliers. However, these methods use some weighted averages to reduce effects of the outliers, so these methods have time delay. On the other hand, RKF in [20] applies l_1 regression to KF. The method uses a property that outliers tend to be sparse and estimates the outliers explicitly. l_1 regression [21] provides some thresholds of solutions and can gives sparse solutions. Therefore, estimates of the outliers using l_1 regression are truncated by the thresholds, and the method has less delay than the other RKF. In addition, the method is easier to implement and compute more efficiently than the other RKF due to a simple structure and convex optimization problem, so the method attracts many attentions. However, parameters of the method are designed by heuristic methods. Moreover, the RKF requires some iterative algorithms to solve the optimization, so a convergence rate and accuracy of the solutions of the RKF depend on conditions of the iterations. There is no practical RKF without heuristic designs.

For state estimations under outliers in nonlinear systems, extended KF (EKF) is often used to extend the aforementioned methods. However, EKF requires Jacobians of the nonlinear systems and yields unstable solutions numerically. Strong nonlinearity deteriorates performances of EKF. Gaussian sum filter (GSF) is other famous KF for nonlinear systems and non-Gaussian measurement noise including outliers [34, 35]. Since GSF can approximate arbitrary distributions by using a Gaussian mixture distribution, GSF can provide a global optimal estimate. However, the complex distributions need to be known, and computational costs are very high. Particle filter [36] can also approximate arbitrary distributions by using a Monte Carlo method, and no prior information of distributions is needed. However, computation time becomes very large, and it is unsuitable for real time applications. In other approaches, robust filter is realized by maximum a posterior (MAP) estimation using a Laplace distribution as a prior distribution [49]. However, it can consider only output equations of nonlinear systems, not dynamics of the nonlinear systems. There are few robust nonlinear estimators under outliers in practice.

In control designs under outliers, there are no methods considering outliers explicitly. The most basic strategy to construct control systems under outliers separates control designs and estimation problems, and the estimation problems deal with the outliers. However, in order to guarantee a stability of the control system, it is required not only to reduce effects of the outliers, but also to analyze performances of the reduction method. Moreover, the separation principle is not satisfied for nonlinear systems in general. For such systems, It is important to guarantee a stability of a total system including a controller and observer.

1.2 Purpose of this dissertation

This dissertation proposes a robust estimation under outliers without heuristic designs and analyzes performances of the proposed estimation. We propose an efficient algorithm of the proposed estimation, and a robust nonlinear estimation is also proposed. Additionally, this dissertation also proposes a robust controller under outliers by applying the proposed estimation, and we analyze a stability of the proposed controller under outliers. Concretely, we focus on RKF via l_1 regression, and we propose a new design method and efficient algorithm of the RKF. Moreover, we propose a robust self-tuning controller and robust nonlinear estimator by applying an idea of the RKF.

1.3 Outline

This dissertation comprises seven chapters and three appendixes, and it is organized as follows.

Chapter 1: Introduction

Firstly, in this chapter, we introduce background, motivation, and purpose for this dissertation.

Chapter 2: Preliminaries

We provide some preliminaries for this dissertation. In this chapter, KF and RKF via l_1 regression, which are origins of this dissertation, are also described.

Chapter 3: Design Method of Robust Kalman Filter Based on Statistics and Its Application

In this chapter, we propose a new design method of RKF via l_1 regression. We show that statistics of Gaussian noise determine the parameters of the RKF, and we can design the parameters systematically. We apply the method to a target tracking system under clutter and a control problem of a two-wheeled vehicle under outliers. The situations are often observed in many cases, e.g., a position control of a vehicle using GPS, using ultra sonic sensors, using image measurements, and so on.

Chapter 4: Fast Algorithm of Robust Kalman Filter via l_1 Regression by a Closed Form Solution

In this chapter, we propose an efficient algorithm of RKF via l_1 regression. We derive upper and lower bounds of an optimal estimate of the RKF and compute an approximated estimate using the both bounds. In addition, the approximated estimate is given by a closed form, so no iteration is needed and it gives a fast computation.

Chapter 5: Robust Self-Tuning Controller under Outliers

In this chapter, we extend a self-tuning controller (STC) in [47] and propose a robust STC (RSTC) under outliers. A parameter update law of the conventional STC is given by a recursive least squares (RLS) estimation. This means that the parameter update law is equivalent to a solution of a minimization problem which consists of a quadratic form of estimated errors. We apply an idea of RKF via l_1 regression to the minimization problem, and we estimate outliers explicitly by adding l_1 regression term to the minimization problem. We construct the RSTC by removing the estimated outliers from the STC. Moreover, we analyze control performances of the proposed RSTC.

Chapter 6: Robust Nonlinear State Estimation and Design Method of Its Parameters

In this chapter, we expand RKF via l_1 regression to nonlinear systems under outliers by using unscented KF (UKF) [50]–[53], and a robust UKF (RUKF) is proposed. In addition, we also derive a design method of its parameters by using a framework of a MAP estimation, and we show that the parameters can be determined systematically.

Chapter 7: Conclusion

This chapter summarizes the contributions of each chapter and provides our future works.

Chapter 2

Preliminaries

2.1 Notation about Vector and Matrix

A vector is represented as a bold character and its element is as a subscript i. For example, i-th element of a vector $\boldsymbol{x}$ is represented as x_i.

In this dissertation, time is expressed as k. When variables and functions depend on time k, the time dependent variables and functions are also denoted as a subscript k. For example, when a vector $\boldsymbol{x}$ depends on time k, the vector and its i-th element are expressed as $\boldsymbol{x}_k$ and $x_{k,i}$, respectively.

$||\cdot||_1$ and $||\cdot||_\infty$ are represented as l_1 and l_∞ norms, respectively. Let $\boldsymbol{x} = [x_1, \cdots, x_n]^T$, then they are defined as

$$\begin{aligned} ||\boldsymbol{x}||_1 &:= \sum_{i=1}^{n} |x_i|, \\ ||\boldsymbol{x}||_\infty &:= \max_i |x_i|. \end{aligned}$$

A matrix is also represented as a bold character and its (i,j)-th element is as a subscript ij. For example, (i,j)-th element of a matrix $\boldsymbol{A}$ is represented as A_{ij} or a_{ij}. If a matrix $\boldsymbol{A}$ depends on time k, the matrix and its (i,j)-th element are expressed as $\boldsymbol{A}_k$ and $A_{k,ij}$ (or $a_{k,ij}$), respectively.

$\boldsymbol{I}$ is an identity matrix with an appropriate dimension.

diag$(\cdot)$ represents a diagonal matrix. For example, $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \cdots, \lambda_m)$ indicates the following diagonal matrix:

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{bmatrix}.$$

$\mathbb{E}[\boldsymbol{x}]$ is an expectation of $\boldsymbol{x}$, and it represents a mean of $\boldsymbol{x}$. Let $\bar{\boldsymbol{x}} = \mathbb{E}[\boldsymbol{x}]$, then a second moment (covariance matrix) of $\boldsymbol{x}$ is written as $\mathbb{E}[(\boldsymbol{x} - \bar{\boldsymbol{x}})(\boldsymbol{x} - \bar{\boldsymbol{x}})^T]$.

2.2 Definition of Derivatives of Absolute Functions

A derivative of an absolute function is defined as the following sub-gradient:

$$\frac{\partial |x_i|}{\partial x_i} :\in \begin{cases} \{1\} & x_i > 0 \\ [-1,\ 1] & x_i = 0 \\ \{-1\} & x_i < 0 \end{cases}.$$

2.3 Notation of Probability Density Function

A normal distribution, whose mean is $\boldsymbol{\mu}$ and covariance matrix is $\boldsymbol{\Sigma}^2$, is described as $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}^2)$, or it is denoted as $p_n(\boldsymbol{x})$ briefly.

PDFs of Cauchy and Gaussian mixture distributions are denoted as $p_c(x)$ and $p_g(\boldsymbol{x})$, respectively, and they are given by the following equations:

$$\begin{aligned} p_c(x) &= \frac{1}{\pi}\frac{\delta}{\delta^2 + (x - x_0)^2}, \\ p_g(\boldsymbol{x}) &= (1-\varepsilon)\mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1^2) + \varepsilon\mathcal{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2^2), \end{aligned}$$

where x_0 is a center and δ is a width of the Cauchy distribution. ε is a random variable distributed by a Bernoulli distribution whose probability is p.

For a stochastic variable $\boldsymbol{x} \in \mathbb{R}^m$, let $p_l(\boldsymbol{x})$ be a PDF of a Laplace distribution, and it is defined as

$$p_l(\boldsymbol{x}) = 2^{-m/2}\det(\boldsymbol{S})^{-1/2} \exp\left[-\sqrt{2}\left|\left|\boldsymbol{S}^{-1/2}(\boldsymbol{x} - \boldsymbol{\mu})\right|\right|_1\right], \tag{2.1}$$

where $\boldsymbol{\mu}$ is a mean and $\boldsymbol{S}$ is a covariance matrix. Several kinds of multivariate Laplace distributions have been used [22, 49], and Eq. (2.1) is a one of the multivariate Laplace distributions. A derivation of the PDF is written in Appendix A.

2.4 Kalman Filter

Let $\boldsymbol{x}_k \in \mathbb{R}^n$ and $\boldsymbol{y}_k \in \mathbb{R}^m$ be a state and measurement at time k, respectively. We consider the following linear time-invariant (LTI) system:

$$\begin{aligned} \boldsymbol{x}_k &= \boldsymbol{A}\boldsymbol{x}_{k-1} + \boldsymbol{w}_k, \\ \boldsymbol{y}_k &= \boldsymbol{C}\boldsymbol{x}_k + \boldsymbol{v}_k, \end{aligned} \tag{2.2}$$

where $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is a system matrix and $\boldsymbol{C} \in \mathbb{R}^{m\times n}$ is an observation matrix. $\boldsymbol{w}_k \in \mathbb{R}^n$ is a system Gaussian noise at time k and $\boldsymbol{v}_k \in \mathbb{R}^m$ is a Gaussian measurement noise at time k.

We assume that $\boldsymbol{w}_k$ is independent of $\boldsymbol{v}_k$. Let $\boldsymbol{P} \in \mathbb{R}^n$ be a covariance matrix of a state estimation error, and let $\boldsymbol{Q} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{R} \in \mathbb{R}^{m\times m}$ denote covariance matrices of $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$, respectively.

Considering the LIT system (2.2), predict and update laws of KF are expressed as

$$\text{Predict law:} \quad \begin{cases} \hat{\boldsymbol{x}}_{k|k-1} &= \boldsymbol{A}\hat{\boldsymbol{x}}_{k-1|k-1}, \\ \boldsymbol{P}_{k|k-1} &= \boldsymbol{A}\boldsymbol{P}_{k-1|k-1}\boldsymbol{A}^T + \boldsymbol{Q}, \end{cases} \tag{2.3}$$

$$\text{Update law:} \quad \begin{cases} \boldsymbol{L} &= \boldsymbol{P}_{k|k-1}\boldsymbol{C}^T(\boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{R})^{-1}, \\ \boldsymbol{e}_k &= \boldsymbol{y}_k - \boldsymbol{C}\hat{\boldsymbol{x}}_{k|k-1}, \\ \hat{\boldsymbol{x}}_{k|k} &= \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{L}\boldsymbol{e}_k, \\ \boldsymbol{P}_{k|k} &= (\boldsymbol{I} - \boldsymbol{L}\boldsymbol{C})\boldsymbol{P}_{k|k-1}, \end{cases} \tag{2.4}$$

where $\hat{\boldsymbol{x}}$ is an estimate of $\boldsymbol{x}$.

A solution of the following optimization problem can derive the update law of KF (2.4):

$$\hat{\boldsymbol{x}}_{k|k} = \arg\min_{\boldsymbol{x}_k} \boldsymbol{v}_k^T\boldsymbol{R}^{-1}\boldsymbol{v}_k + \left(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}\right)^T \boldsymbol{P}_{k|k-1}^{-1} \left(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}\right). \tag{2.5}$$

Actually, a derivation of the solution is written in Appendix B.

2.5 Robust Kalman Filter via l_1 Regression

In this section, RKF in [20] is described.

We consider the following LTI system:

$$\begin{aligned} \boldsymbol{x}_k &= \boldsymbol{A}\boldsymbol{x}_{k-1} + \boldsymbol{w}_k, \\ \boldsymbol{y}_k &= \boldsymbol{C}\boldsymbol{x}_k + \boldsymbol{v}_k + \boldsymbol{z}_k, \end{aligned} \tag{2.6}$$

where $\boldsymbol{z}_k \in \mathbb{R}^m$ is an outlier in a measurement at time k, and it is a different point from Eq. (2.2).

A predict law of RKF via l_1 regression is given by a same form as KF:

$$\begin{aligned} \hat{\boldsymbol{x}}_{k|k-1} &= \boldsymbol{A}\hat{\boldsymbol{x}}_{k-1|k-1}, \\ \boldsymbol{P}_{k|k-1} &= \boldsymbol{A}\boldsymbol{P}_{k-1|k-1}\boldsymbol{A}^T + \boldsymbol{Q}. \end{aligned} \tag{2.7}$$

Adding l_1 regularization term to Eq. (2.5) results in an update law of the RKF:

$$\begin{aligned} \{\hat{\boldsymbol{x}}_{k|k}, \hat{\boldsymbol{z}}_k\} = \arg\min\nolimits_{\boldsymbol{x}_k,\boldsymbol{z}_k} & \boldsymbol{v}_k^T\boldsymbol{R}^{-1}\boldsymbol{v}_k \\ & + \left(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}\right)^T \boldsymbol{P}_{k|k-1}^{-1} \left(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}\right) + \lambda||\boldsymbol{z}_k||_1, \\ \text{subject to} \quad & \boldsymbol{L} = \boldsymbol{P}_{k|k-1}\boldsymbol{C}^T(\boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{R})^{-1}, \\ & \boldsymbol{e}_k = \boldsymbol{y}_k - \boldsymbol{C}\hat{\boldsymbol{x}}_{k|k-1}, \\ & \boldsymbol{x}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{L}(\boldsymbol{e}_k - \boldsymbol{z}_k), \\ & \boldsymbol{P}_{k|k} = (\boldsymbol{I} - \boldsymbol{L}\boldsymbol{C})\boldsymbol{P}_{k|k-1}, \end{aligned} \tag{2.8}$$

where λ is a regularization parameter. In KF, the solution of the optimization problem (2.5) yields a linear form, i.e., $\hat{\boldsymbol{x}}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{L}\boldsymbol{e}_k$. However, it is assumed that an estimate of the RKF is given by the linear form and an innovation is represented by a prediction error except outliers. Since a criteria using a quadratic form is very sensitive to outliers, l_1 norm is adopted to evaluate the outliers in Eq. (2.8). Additionally, by using l_1 regularization term, solutions tend to be sparse, i.e., the solutions can contain many zero values. Hence, l_1 regularization term can provide solutions containing many zero values more frequently than l_2 regularization term. The outliers are infrequent and their values can usually become zero, and l_1 regularization term provides estimates of $\boldsymbol{z}_k$ with a property that solutions may contain many zero values.

Eq. (2.8) can be rewritten as

$$\begin{aligned} \boldsymbol{L} &= \boldsymbol{P}_{k|k-1}\boldsymbol{C}^T(\boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{R})^{-1}, \\ \boldsymbol{e}_k &= \boldsymbol{y}_k - \boldsymbol{C}\hat{\boldsymbol{x}}_{k|k-1}, \\ \hat{\boldsymbol{x}}_{k|k} &= \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{L}(\boldsymbol{e}_k - \hat{\boldsymbol{z}}_k), \\ \boldsymbol{P}_{k|k} &= (\boldsymbol{I} - \boldsymbol{L}\boldsymbol{C})\boldsymbol{P}_{k|k-1}, \end{aligned} \tag{2.9}$$

and $\hat{\boldsymbol{z}}_k$ is given by a solution of the following optimization problem via l_1 regression:

$$\hat{\boldsymbol{z}}_k = \arg\min_{\boldsymbol{z}_k} (\boldsymbol{e}_k - \boldsymbol{z}_k)^T \boldsymbol{W} (\boldsymbol{e}_k - \boldsymbol{z}_k) + \lambda ||\boldsymbol{z}_k||_1, \tag{2.10}$$

where $\boldsymbol{W}$ is a positive definite matrix and given by

$$\begin{aligned} \boldsymbol{W} &= (\boldsymbol{I} - \boldsymbol{C}\boldsymbol{L})^T\boldsymbol{R}^{-1}(\boldsymbol{I} - \boldsymbol{C}\boldsymbol{L}) + \boldsymbol{L}^T\boldsymbol{P}_{k|k-1}^{-1}\boldsymbol{L} \\ &= \left(\boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{R}\right)^{-1}. \end{aligned} \tag{2.11}$$

In multi output systems, using multi regularization parameters, i.e., $\lambda_i, i = 1, \cdots, m$, the optimization problem (2.10) can be generalized as

$$\hat{\boldsymbol{z}}_k = \arg\min_{\boldsymbol{z}_k} \ (\boldsymbol{e}_k - \boldsymbol{z}_k)^T \boldsymbol{W} (\boldsymbol{e}_k - \boldsymbol{z}_k) + \sum_{i=1}^{m} \lambda_i |z_{k,i}|. \tag{2.12}$$

where $\boldsymbol{\lambda} = [\lambda_1, \cdots, \lambda_m]^T \in \mathbb{R}^m$.

The regularization parameters need to be tuned suitably for each application, and it is often determined by heuristic methods. This dissertation proposes a new systematic design method of the parameter by using the covariance matrices, $\boldsymbol{Q}$ and $\boldsymbol{R}$.

Chapter 3

Design Method of Robust Kalman Filter via l_1 Regression Based on Statistics and Its Application

3.1 Introduction

In automobiles, industrial robots, medical machines, and so on, computerization gives the machines high functionalities. However, external noise has negative effects on the computerized machines, especially, outliers are contained in sensor signals of the machines. In other cases, non-contact sensors, e.g., radar measurements, GPS, ultrasonic wave sensors, image measurements, and so on, are often used in control systems. However, external environments introduce outliers into these sensor signals. For example, in target tracking systems [4, 5, 6], outliers are occurred on distance and angle information due to reflection noise, and the outliers are called clutter. Additionally, in UAV using visual feedback [7], temporary change of image contrast in background causes outliers of position data. In UGV using GPS [8], radio disturbances due to some obstacles provide position data with outliers. These outliers deteriorate state estimates and control performances.

In order to reduce effects of the outliers, many methods have been proposed [14]–[20]. For example, for a system identification, multiplication of a median of past N-th data and some coefficients give a threshold of outliers, and the outliers are detected and removed by the threshold [14]. The method requires many past data, and the threshold is determined by trial and error. In [15], a reduction method of outliers using sliding mode has been proposed. However, parameters of the method are designed by heuristic methods, and chattering is introduced in digital implementations.

Kalman filter (KF) for outliers has also been proposed in [16]–[20], and the methods are called robust KF (RKF). In [16, 17], Bayesian model is introduced to KF, and EM algorithm is used. In [18], use of variational Bayesian method gives approximations of joint posterior distributions of state and noise variances to realize a low computational cost. In [19], the method also learns the covariance matrix of measurement noise by iterations to compute a Kalman gain. These methods learn their parameters automatically and adjust their parameters for the outliers. But, these methods use some weighted averages to reduce effects of the outliers, so these methods have time delay. On the other hand, RKF in [20] applies l_1 regression to KF. l_1 regression [21] provides some thresholds of solutions and can gives sparse solutions, so estimates of the outliers using l_1 regression are truncated by the thresholds. In addition, the method is easy to implement and compute due to a simple structure and convex optimization problem, so the method attracts many attentions. However, parameters of the method are designed by heuristic methods.

In this chapter, we propose a new design method of RKF via l_1 regression. We show that statistics of Gaussian noise determine the parameters of the RKF, and we can design the parameters systematically.

The organization of this chapter is as follows. In section 2, we propose a new design method of the RKF. In section 3, we derive the proposed method from a Lagrange dual problem. In section 4, we analyze performances of the RKF with the proposed design method. In section 5 and 6, we apply the RKF with the proposed design method to a radar tracking system with clutters and a two-wheeled vehicle under outliers, respectively. We demonstrate its effectiveness by some numerical simulations in these sections. Conclusion is given in section 7.

Hereafter, we denote RKF via l_1 regression as just "RKF", and we use the notation without distinction from the other RKF.

3.2 Proposed Design Method of RKF Based on Statistics

3.2.1 Design Method for Single Output Systems

First, to improve understanding, a design method for single output systems is explained.

In the case of single output systems, an optimal solution $\hat{z}_k$ of Eq. (2.10) is given by the following equation analytically [21, 23]:

$$\hat{z}_k = \begin{cases} e_k - \frac{\lambda}{2W} & (z_k \geq 0) \\ e_k + \frac{\lambda}{2W} & (z_k < 0) \end{cases} .$$

Therefore,

$$\hat{z}_k = \begin{cases} e_k - \frac{\lambda}{2W} & \left(e_k \geq \frac{\lambda}{2W}\right) \\ 0 & \left(-\frac{\lambda}{2W} \leq e_k < \frac{\lambda}{2W}\right) \\ e_k + \frac{\lambda}{2W} & \left(e_k < -\frac{\lambda}{2W}\right) \end{cases} . \tag{3.1}$$

For example, Fig. 3.1 shows a graph of Eq. (3.1), where $W = 1$ and $\lambda = 10$. It can be seen that $\hat{z}_k = 0$ in the range of $-\frac{\lambda}{2W} \leq e_k \leq \frac{\lambda}{2W}$, so Eq. (2.9) is equal to an update law of standard KF in the range. On the other hand, $\hat{z}_k$ is non-zero if an absolute value of the output error e_k is greater than $\frac{\lambda}{2W}$. This means that $\frac{\lambda}{2W}$ is a threshold of outlier $\hat{z}_k$, and $\frac{\lambda}{2W}$ can be interpreted as a possible range of e_k without outlier $\hat{z}_k$.

Let e_k^* be a prediction error without outlier z_k:

$$\begin{aligned} e_k^* &:= e_k|_{z_k=0} \\ &= \boldsymbol{C}\left(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}\right) + v_k. \end{aligned} \tag{3.2}$$

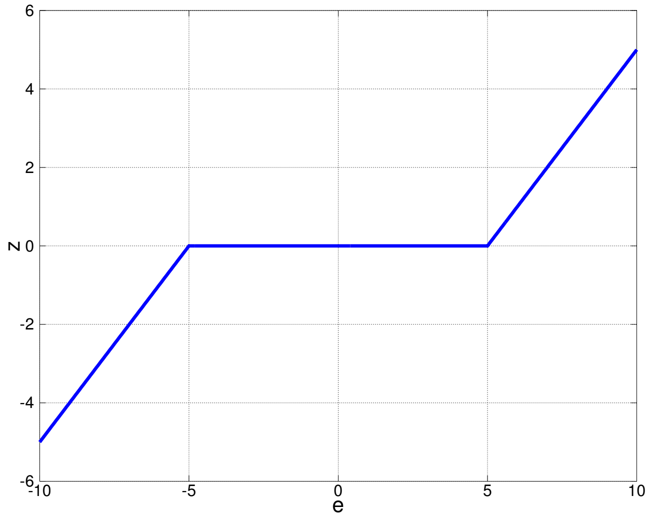


Figure 3.1: Solution of robust Kalman filter for single output systems, where $W = 1$ and $\lambda = 10$, therefore $\lambda/(2W) = 5$.

The range of e_k^* depends on distributions of an estimation error $\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}$ and measurement Gaussian noise v_k, i.e., e_k^* also depends on a Gaussian distribution, so we use a standard deviation (STD) of e_k^* as the possible range of e_k^*. Let σ_{e^*} be a STD of e_k^*, then a variance $\sigma_{e^*}^2$ is given by

$$\begin{aligned} \sigma_{e^*}^2 &= \mathbb{E}\left[(e_k^*)^2\right] \\ &= \boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + R. \end{aligned}$$

Therefore, the parameter of RKF for single output systems satisfies the following equation:

$$\begin{aligned} \frac{\lambda}{2W} &= \sigma_{e^*}, \\ \therefore \lambda &= 2W\sigma_{e^*}. \end{aligned} \tag{3.3}$$

3.2.2 Design Method for Multi Output Systems

A same procedure as the design for single output systems derives parameters of RKF for multi output systems. A first order necessary condition of an optimality for Eq. (2.12) derives the following inclusion:

$$\left(2\boldsymbol{W}(\boldsymbol{e}_k - \hat{\boldsymbol{z}}_k)\right)_i \in \begin{cases} \{\lambda_i\} & \hat{z}_{k,i} > 0 \\ [-\lambda_i, \lambda_i] & \hat{z}_{k,i} = 0 \\ \{-\lambda_i\} & \hat{z}_{k,i} < 0 \end{cases}, \tag{3.4}$$

where $(\cdot)_i$ is represented as a i-th element of a vector. Eq. (2.12) is a convex optimization problem, so the condition is also sufficient. Note that, in general, the optimization problem for multi output systems cannot be solved analytically.

Assuming that measurements include no outliers in the same way as the design for single output systems, i.e., $\boldsymbol{z}_k = \boldsymbol{0}$, Eq. (3.2) and (3.4) give the following inclusion:

$$\left(2\boldsymbol{W}\boldsymbol{e}_k^*\right)_i \in [-\lambda_i, \lambda_i]. \tag{3.5}$$

Eq. (3.5) indicates that λ_i is a some upper bound for $e_{k,i}^*$. In an opposite manner, if we know a prior information of an upper bound of $\boldsymbol{e}_k^*$ for all time, we can calculate λ_i by the upper bound, i.e., parameters of RKF are given by

$$\boldsymbol{\varepsilon} = 2\boldsymbol{W} \sup_k \boldsymbol{e}_k^*, \tag{3.6}$$

$$\lambda_i = |\varepsilon_i|, \tag{3.7}$$

where $\sup_k \boldsymbol{e}_k^* := [\sup_k e_{k,1}^*, \cdots, \sup_k e_{k,m}^*]^T$, and $\boldsymbol{\varepsilon} = [\varepsilon_1, \cdots, \varepsilon_m]^T \in \mathbb{R}^m$.

However, we cannot calculate $\sup_k \boldsymbol{e}_k^*$ in advance. In this chapter, we replace $\sup_k \boldsymbol{e}_k^*$ with another value. Note that $\boldsymbol{e}_k^*$ is the output error considering only a Gaussian distribution as measurement noise. Then, Eq. (3.5) is expressed as

$$\begin{bmatrix} \lambda_1\eta_1 \\ \vdots \\ \lambda_m\eta_m \end{bmatrix} = 2\boldsymbol{W}\boldsymbol{e}_k^*, \tag{3.8}$$

where η_i can be chosen randomly in $[-1,\ 1]$, so η_i can be regarded as a stochastic variable without loss of generality. Moreover, η_i is independent of the other stochastic variables because η_i can be chosen independently. Note that $\boldsymbol{e}_k^*$ is a prediction error considering only Gaussian noise as measurement noise. A covariance matrix of $\boldsymbol{e}_k^*$, i.e., $\boldsymbol{\Sigma}_{\boldsymbol{e}^*}$, is given by

$$\begin{aligned} \boldsymbol{\Sigma}_{\boldsymbol{e}^*} &= \mathbb{E}[\boldsymbol{e}_k^*\boldsymbol{e}_k^{*T}] \\ &= \boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{R}. \end{aligned} \tag{3.9}$$

Note that Eq. (2.11), $\mathbb{E}[\eta_i^2] \le 1$, and $\mathbb{E}[\eta_i\eta_j] = 0$ $(i \neq j)$ are satisfied. Eq. (3.8) yields the following inequality:

$$\begin{aligned} \begin{bmatrix} \lambda_1^2 & & \\ & \ddots & \\ & & \lambda_m^2 \end{bmatrix} &\ge \begin{bmatrix} \lambda_1^2\mathbb{E}[\eta_1^2] & & \\ & \ddots & \\ & & \lambda_m^2\mathbb{E}[\eta_m^2] \end{bmatrix} \\ &= \mathbb{E}\left[\begin{bmatrix} \lambda_1\eta_1 \\ \vdots \\ \lambda_m\eta_m \end{bmatrix}\begin{bmatrix} \lambda_1\eta_1 & \cdots & \lambda_m\eta_m \end{bmatrix}\right] \\ &= 4\boldsymbol{W}\mathbb{E}\left[\boldsymbol{e}_k^*\boldsymbol{e}_k^{*T}\right]\boldsymbol{W} \\ &= 4\boldsymbol{W}. \end{aligned} \tag{3.10}$$

This means that the parameter $\boldsymbol{\lambda}$ should be determined to satisfy Eq. (3.10).

Assume that $\boldsymbol{\Sigma}_{\boldsymbol{e}^*} = \boldsymbol{W}^{-1}$ is a diagonal matrix and given by $\boldsymbol{\Sigma}_{\boldsymbol{e}^*} = \mathrm{diag}(\sigma_1^2, \cdots, \sigma_m^2)$. The parameter $\boldsymbol{\lambda}$ can be selected as $\mathrm{diag}(\lambda_1^2, \cdots, \lambda_m^2) = 4\boldsymbol{W}$. Let $\boldsymbol{\sigma}_{\boldsymbol{e}^*} = [\sigma_1, \cdots, \sigma_m]^T \in \mathbb{R}^m$, then the parameter $\boldsymbol{\lambda}$ can be rewritten as

$$\begin{aligned} \boldsymbol{\lambda} &= 2 \begin{bmatrix} 1/\sigma_1 \\ \vdots \\ 1/\sigma_m \end{bmatrix} \\ &= 2\boldsymbol{W}\boldsymbol{\sigma}_{e^*}. \end{aligned} \tag{3.11}$$

In single output systems, Eq. (3.11) coincides with Eq. (3.3)

In the case of a single regularization parameter, Eq. (3.5) is replaced as the following equation:

$$\left|\left|2\boldsymbol{W}\boldsymbol{e}_k^*\right|\right|_\infty \le \lambda.$$

This means that λ should be chosen by $\left|\left|2\boldsymbol{W}\sup_k \boldsymbol{e}_k^*\right|\right|_\infty$, i.e., $||\boldsymbol{\lambda}||_\infty$ satisfying Eq. (3.10).

Some remarks are given for the proposed design method of the regularization parameter.

Remark 3.1 Eq. (3.10) and (3.11) mean that a STD is used instead of $\sup_k \boldsymbol{e}_k^*$ in Eq. (3.6). It is interpreted that $\boldsymbol{e}_k^*$ is in a set except outliers, i.e., in a "normal value", then $\sup_k \boldsymbol{e}_k^*$ is replaced with the STD in practice. Use of the STD to determine the regularization parameter can simplify computations. Additionally, a covariance matrix of an innovation of RKF can be bounded by a covariance matrix of normal KF without outliers, and it is proven in section 3.4. □

Remark 3.2 The covariance matrix $\boldsymbol{P}$, which is used in Eq. (3.10) and (3.11), is predicted and updated, and its tuning parameters are $\boldsymbol{Q}$ and $\boldsymbol{R}$. This means that parameter designs of RKF are equivalent to ones of standard KF, and we need no prior information for outliers. □

Remark 3.3 Only designing of standard KF determines and updates the parameters of RKF automatically, so we can interpret RKF with the proposed design method as a "self-tuning" RKF. □

Remark 3.4 Traditionally, the parameters of RKF are determined by heuristic methods, so the conventional design methods might be inappropriate in physical meanings. However, the proposed design method can determine the parameters via statistics of noise, i.e., $\boldsymbol{Q}$ and $\boldsymbol{R}$, so the parameters have physical meanings. □

3.3 Derivation of Proposed Design Method by Lagrange Dual Problem

We consider a new variable $\boldsymbol{r}_k \in \mathbb{R}^m$. The optimization problem (2.12) can be rewritten as the following equation [24, 25]:

$$\begin{aligned} &\min_{\boldsymbol{r}_k, \boldsymbol{z}_k} \ \boldsymbol{r}_k^T \boldsymbol{W} \boldsymbol{r}_k + \textstyle\sum_{i=1}^m \lambda_i |z_{k,i}|, \\ &\text{subject to } \boldsymbol{r}_k = \boldsymbol{e}_k - \boldsymbol{z}_k. \end{aligned} \tag{3.12}$$

Let $\boldsymbol{\nu} \in \mathbb{R}^m$ be a Lagrange multiplier, then a Lagrangian can be defined as

$$L(\boldsymbol{r}_k, \boldsymbol{z}_k, \boldsymbol{\nu}) := \boldsymbol{r}_k^T \boldsymbol{W} \boldsymbol{r}_k + \sum_{i=1}^{m} \lambda_i |z_{k,i}| + \boldsymbol{\nu}^T \left(\boldsymbol{r}_k - \boldsymbol{e}_k + \boldsymbol{z}_k\right). \tag{3.13}$$

Let $g(\boldsymbol{\nu})$ be a Lagrange dual function, then a Lagrange dual problem is given by the following maximization problem:

$$\begin{aligned}&\max_{\boldsymbol{\nu}} g(\boldsymbol{\nu}),\\ &g(\boldsymbol{\nu}) := \inf_{\boldsymbol{r}_k, \boldsymbol{z}_k} L(\boldsymbol{r}_k, \boldsymbol{z}_k, \boldsymbol{\nu}).\end{aligned}$$

We derive the Lagrange dual function $g(\boldsymbol{\nu})$. Since the Lagrangian L is a quadratic form for $\boldsymbol{r}_k$ and $\boldsymbol{W}$ is a positive-definite matrix, L is bounded below in $\boldsymbol{r}_k$. On the other hand, L is linear for all $\boldsymbol{z}_k$, except $\boldsymbol{z}_k = \boldsymbol{0}$, so L is bounded below only if $\boldsymbol{z}_k = \boldsymbol{0}$. Therefore, if $\boldsymbol{z}_k \neq 0$, $\inf_{\boldsymbol{r}_k, \boldsymbol{z}_k} L(\boldsymbol{r}_k, \boldsymbol{z}_k, \boldsymbol{\nu}) = -\infty$.

If $\boldsymbol{z}_k = \boldsymbol{0}$ is satisfied, $\frac{\partial L}{\partial \boldsymbol{r}_k} = \boldsymbol{0}$ give the following equation:

$$\boldsymbol{r}_k = -\frac{1}{2}\boldsymbol{W}^{-1}\boldsymbol{\nu} \quad (|\nu_i| \leq \lambda_i, \quad i = 1 \cdots m).$$

Then, the Lagrangian L is represented as

$$L(\boldsymbol{\nu}, \boldsymbol{r}_k, \boldsymbol{z}_k) = -\frac{1}{4}\boldsymbol{\nu}^T \boldsymbol{W}^{-1} \boldsymbol{\nu} - \boldsymbol{\nu}^T \boldsymbol{e}_k, \ (|\nu_i| \leq \lambda_i, i = 1 \cdots m).$$

Therefore, the Lagrange dual function $g(\boldsymbol{\nu})$ is given by

$$\begin{aligned} g(\boldsymbol{\nu}) &= \inf_{\boldsymbol{r}_k, \boldsymbol{z}_k} L(\boldsymbol{r}_k, \boldsymbol{z}_k, \boldsymbol{\nu}) \\ &= \begin{cases} -\frac{1}{4}\boldsymbol{\nu}^T \boldsymbol{W}^{-1}\boldsymbol{\nu} - \boldsymbol{\nu}^T \boldsymbol{e}_k & |\nu_i| \leq \lambda_i, i = 1 \cdots m \\ -\infty & \text{otherwise} \end{cases}. \end{aligned}$$

If $|\nu_i| \leq \lambda_i$, the Lagrange dual problem has feasible solutions. This means that the regularization parameters provide upper bounds for feasible solutions in a dual space. In addition, since the primal problem (3.12) is convex and satisfies Slater's condition [24], optimal solutions of the primal and dual problem coincide with each other.

Note that $\boldsymbol{z}_k = \boldsymbol{0}$ if $|\nu_i| \leq \lambda_i$ is satisfied. An optimal solution of the Lagrange dual problem, i.e., $\boldsymbol{\nu}^*$, is given by

$$\begin{aligned} \boldsymbol{\nu}^* &= -2\boldsymbol{W}\boldsymbol{e}_k|_{\boldsymbol{z}_k = \boldsymbol{0}} \\ &= -2\boldsymbol{W}\boldsymbol{e}_k^* \quad (|\nu_i| \leq \lambda_i). \end{aligned}$$

The condition is equivalent to Eq. (3.5), and the same results as the previous section can be obtained in the dual space.

Remark 3.5 The dual problem derives Eq. (3.5) without the assumption that $\boldsymbol{z}_k = \boldsymbol{0}$ used in the previous section. □

3.4 Covariance Matrix under outliers

3.4.1 Covariance Matrices of an Innovation and Estimated Outliers and Its Upper Bounds

For the proposed design method, the following theorem is satisfied.

Theorem 3.1 Assuming that regularization parameters of RKF is given by Eq. (3.10), a covariance matrices of an innovation and estimation error of outliers are given by the following inequalities, respectively:

$$\mathbb{E}[(\boldsymbol{e}_k - \hat{\boldsymbol{z}}_k)(\boldsymbol{e}_k - \hat{\boldsymbol{z}}_k)^T] \leq \gamma \left(\boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{R}\right), \tag{3.14}$$

$$\mathbb{E}\left[(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k)(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k)^T\right] \leq (1+\gamma) \left(\boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{R}\right), \tag{3.15}$$

where γ is a real number, and $\gamma \geq 1$. □

Proof: Let a sub-gradient of $|z_{k,i}|$ be $\eta_i \in [-1,\ 1]$, then the necessary condition of Eq. (2.12) can be rewritten as

$$\forall \hat{\boldsymbol{z}}_k, \quad -2\boldsymbol{W}(\boldsymbol{e}_k - \hat{\boldsymbol{z}}_k) + \begin{bmatrix} \lambda_1\eta_1 \\ \vdots \\ \lambda_m\eta_m \end{bmatrix} = \boldsymbol{0}. \tag{3.16}$$

In similar to Eq. (3.8), it is assumed that η_i is mutually independent stochastic variable without loss of generality. Assuming that $\boldsymbol{\lambda}$ satisfies Eq. (3.10), then a real number $\gamma \geq 1$ exists such that

$$4\boldsymbol{W} \leq \begin{bmatrix} \lambda_1^2 & & \\ & \ddots & \\ & & \lambda_m^2 \end{bmatrix} \leq 4\gamma\boldsymbol{W}.$$

Eq. (3.16) gives the following inequality:

$$\begin{aligned} 4\boldsymbol{W}\ \mathbb{E}[(\boldsymbol{e}_k - \hat{\boldsymbol{z}}_k)(\boldsymbol{e}_k - \hat{\boldsymbol{z}}_k)^T]\ \boldsymbol{W}^T &\leq \begin{bmatrix} \lambda_1^2 & & \\ & \ddots & \\ & & \lambda_m^2 \end{bmatrix} \\ &\leq 4\gamma\boldsymbol{W}. \end{aligned}$$

$$\begin{aligned} \therefore \mathbb{E}[(\boldsymbol{e}_k - \hat{\boldsymbol{z}}_k)(\boldsymbol{e}_k - \hat{\boldsymbol{z}}_k)^T] &\leq \gamma\boldsymbol{W}^{-1} \\ &= \gamma\left(\boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{R}\right), \quad \gamma \geq 1. \end{aligned}$$

From Eq. (2.6) and (3.16),

$$\begin{aligned} \boldsymbol{z}_k - \hat{\boldsymbol{z}}_k &= \boldsymbol{y}_k - \boldsymbol{C}\boldsymbol{x}_k - \boldsymbol{v}_k - \hat{\boldsymbol{z}}_k \\ &= \boldsymbol{e}_k - \boldsymbol{C}(\boldsymbol{x}_k - \boldsymbol{x}_{k|k-1}) - \boldsymbol{v}_k - \hat{\boldsymbol{z}}_k \\ &= \frac{1}{2}\boldsymbol{W}^{-1} \begin{bmatrix} \lambda_1\eta_1 \\ \vdots \\ \lambda_m\eta_m \end{bmatrix} - \boldsymbol{C}(\boldsymbol{x}_k - \boldsymbol{x}_{k|k-1}) - \boldsymbol{v}_k. \end{aligned}$$

Since every stochastic variables are mutually independent, a covariance matrix of an estimation of outliers is bounded by

$$\mathbb{E}\left[(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k)(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k)^T\right] \leq (1+\gamma)\left(\boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{R}\right), \quad \gamma \geq 1.$$

■

Remark 3.6 Eq. (3.14) means that a covariance matrix of an innovation of RKF is bounded by that of normal KF without outliers. This means that the covariance matrix of RKF can be bounded by order of $\boldsymbol{R}$ if $\boldsymbol{P}_{k|k-1}$ converges to a small value. □

Remark 3.7 Single output systems satisfy $\gamma = 1$. Also in the case that the weight matrix (2.11) is diagonal matrix, γ becomes 1. An upper bound of the covariance matrix of the innovation comes close to an ideal one even under outliers. □

3.4.2 Covariance Matrix of a State Estimation Error under Outliers

Eq. (2.6) and (2.9) yield the following equation:

$$\begin{aligned}\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k} &= \boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1} - \boldsymbol{L}(\boldsymbol{e}_k - \hat{\boldsymbol{z}}_k) \\ &= (\boldsymbol{I} - \boldsymbol{L}\boldsymbol{C})(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}) - \boldsymbol{L}\boldsymbol{v}_k - \boldsymbol{L}\left(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k\right).\end{aligned}$$

Both in KF and RKF, Eq. (2.9) updates a covariance matrix of a state estimation error. However, an actual updated covariance matrix under outliers is given by

$$\begin{aligned}\boldsymbol{P}_{k|k} &= \mathbb{E}\left[(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k})(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k})^T\right] \\ &= (\boldsymbol{I} - \boldsymbol{L}\boldsymbol{C})\boldsymbol{P}_{k|k-1}(\boldsymbol{I} - \boldsymbol{L}\boldsymbol{C})^T + \boldsymbol{L}\boldsymbol{R}\boldsymbol{L}^T + \boldsymbol{L}\mathbb{E}\left[(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k)(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k)^T\right]\boldsymbol{L}^T \\ &= (\boldsymbol{I} - \boldsymbol{L}\boldsymbol{C})\boldsymbol{P}_{k|k-1} + \boldsymbol{L}\mathbb{E}\left[(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k)(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k)^T\right]\boldsymbol{L}^T. \end{aligned} \tag{3.17}$$

In standard KF, $\hat{\boldsymbol{z}}_k = \boldsymbol{0}$, so an actual $\boldsymbol{P}_{k|k}$ under outliers depends on a second moment of $\boldsymbol{z}_k$. For example, if $\boldsymbol{z}_k$ is distributed by a distribution whose second moment is infinite, like a Cauchy distribution [26], the updated covariance matrix $\boldsymbol{P}_{k|k}$ should be infinite in ideal. However, it results in no update of a state. On the other hand, in RKF using the proposed algorithm, Eq. (3.17) satisfies the following inequality and bounded:

$$\begin{aligned}\boldsymbol{P}_{k|k} &\leq (\boldsymbol{I} - \boldsymbol{L}\boldsymbol{C})\boldsymbol{P}_{k|k-1} + (1+\gamma)\boldsymbol{L}\boldsymbol{C}\boldsymbol{P}_{k|k-1} \\ &= (\boldsymbol{I} + \gamma\boldsymbol{L}\boldsymbol{C})\boldsymbol{P}_{k|k-1}, \quad \gamma \geq 1. \end{aligned} \tag{3.18}$$

In RKF using the proposed algorithm, the updated covariance matrix of a state estimation error can be selected among solutions satisfying Eq. (3.18). The update law (2.9) is one of the solutions.

3.5 Application to Target Tracking Systems under Clutters

3.5.1 Target Tracking Using Kalman Filter

We apply the proposed method to a target tracking systems under clutters shown in [4].

Fig. 3.2 shows a coordinate of the target tracking systems. An orthogonal coordinate, whose x-axis is east, y-axis is north, and z-axis is a vertical direction, is called "north east coordinate". Assuming a polar coordinate consisting of a distance from a radar, elevation, and azimuth satisfies, the following relation between the north east and polar coordinates is satisfied:

$$\begin{bmatrix} x_k \\ y_k \\ z_k \end{bmatrix} = \begin{bmatrix} r_k \cos\theta_k \sin\varphi_k \\ r_k \cos\theta_k \cos\varphi_k \\ r_k \sin\theta_k \end{bmatrix}. \tag{3.19}$$

Let $\boldsymbol{x}_k = [x_k, \dot{x}_k, y_k, \dot{y}_k, z_k, \dot{z}_k]^T \in \mathbb{R}^6$ be a state of the target at time k, and a state equation of the target is given by the following constant acceleration model on the north east coordinate:

$$\boldsymbol{x}_k = \boldsymbol{A}_a \boldsymbol{x}_{k-1} + \boldsymbol{w}_k, \tag{3.20}$$

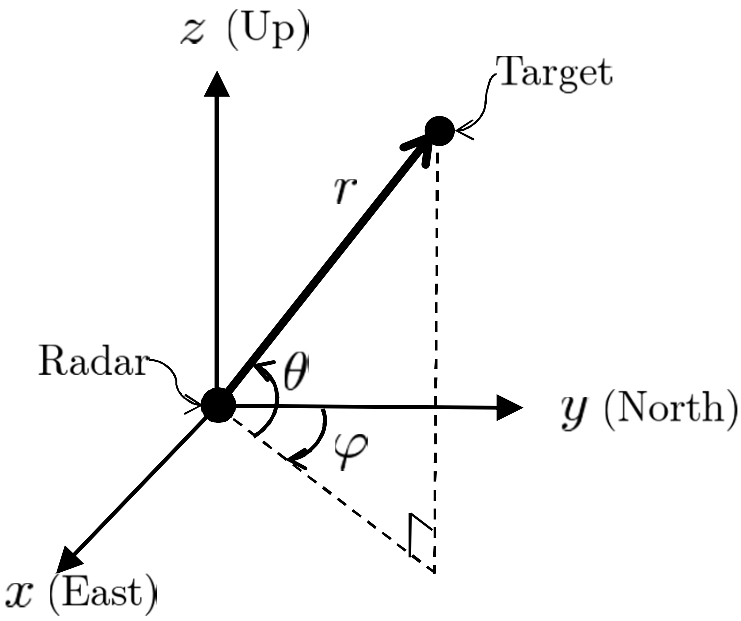


Figure 3.2: North east up coordinates and polar coordinates for target tracking systems.

where $\boldsymbol{A}_a$ and $\boldsymbol{w}_k$ are as follows:

$$\begin{aligned} \boldsymbol{A}_a &= \begin{bmatrix} \boldsymbol{A} & & \\ & \ddots & \\ & & \boldsymbol{A} \end{bmatrix}, \\ \boldsymbol{w}_k &= \begin{bmatrix} \boldsymbol{G}^T a_x \\ \boldsymbol{G}^T a_y \\ \boldsymbol{G}^T a_z \end{bmatrix}, \\ \boldsymbol{A} &= \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}, \\ \boldsymbol{G} &= \begin{bmatrix} \frac{\Delta t^2}{2} \\ \Delta t \end{bmatrix}, \end{aligned}$$

where Δt is a sampling time, and a_x, a_y, and a_z are accelerations of x_k, y_k, and z_k, respectively. Assume that a_x, a_y, and a_z are mutually independent. Let σ_{a_x}, σ_{a_y}, and σ_{a_z} be standard deviations of a_x, a_y, and a_z, respectively. A covariance matrix of $\boldsymbol{w}_k$, i.e., $\boldsymbol{Q}$ is given by

$$\begin{aligned} \boldsymbol{Q} &= E\left[\boldsymbol{w}_k \boldsymbol{w}_k^T\right] \\ &= \begin{bmatrix} \sigma_{a_x}^2 \boldsymbol{G}\boldsymbol{G}^T & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \sigma_{a_y}^2 \boldsymbol{G}\boldsymbol{G}^T & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \sigma_{a_z}^2 \boldsymbol{G}\boldsymbol{G}^T \end{bmatrix}. \end{aligned}$$

Suppose that an output is defined by $\boldsymbol{y}_k = [x_k, y_k, z_k]^T \in \mathbb{R}^3$. An output equation is given by

$$\boldsymbol{y}_k = \boldsymbol{C}_a \boldsymbol{x}_k + \boldsymbol{\Gamma} \boldsymbol{v}_k, \tag{3.21}$$

where $\boldsymbol{v}_k$ is a measurement noise including outliers on the polar coordinate. $\boldsymbol{C}_a$ is a measurement matrix and $\boldsymbol{\Gamma}$ is a matrix which transform a small measurement noise on the polar coordinate to a measurement noise on the north east coordinate, and they are given by

$$\begin{aligned} \boldsymbol{C}_a &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \\ \boldsymbol{\Gamma} &= \left. \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \varphi} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \varphi} \end{bmatrix} \right|_{\boldsymbol{r}=\boldsymbol{r}_k}, \end{aligned}$$

where $\boldsymbol{r} = [r,\ \theta,\ \varphi]^T \in \mathbb{R}^3$. Actual information obtained from the radar is only output on the polar coordinate. Eq. (3.19) transforms the output to north east coordinate's one, then Eq. (3.21) is applied.

3.5.2 Model of Clutters

In target tracking systems, outliers are occurred on distance and angle information due to reflection noise, and the outliers are called clutters. Moreover, the clutters can be modeled as Symmetric α-Stable (SαS) distribution, or Gaussian mixture model [5, 6]. In this section, we consider two types of outliers; one is Cauchy distribution which is SαS distribution with $\alpha = 1$, and the other is Gaussian mixture model.

3.5.3 Conditions

In reference to [4], the following scenario is considered. The radar is located in the origin, and the target starts from the position that an attitude is 9000m and surface distance 70km. After the target starts, the target moves in the direction of the origin along x-axis at 170m/s. Sampling time is 6 second, and a STD of an acceleration is $[\sigma_{a_x},\ \sigma_{a_y},\ \sigma_{a_z}] = [10\text{m/s}^2,\ 0.5\text{m/s}^2,\ 1.0\text{m/s}^2]$.

It is assumed that a nominal measurement noise is Gaussian measurement noise whose mean is $\mathbf{0}$ and STDs are 100m in distance and 7mrad in angle. In the use of Cauchy distribution as the clutters, Cauchy noise is added to the nominal noise. The parameters are $x_0 = 0$ and $\delta = 5 \times 10^{-1}$, and its gain is tuned to be proportional to the STD of the nominal noise. In the use of Gaussian mixture model as the clutters, $p = 0.1$ and its STD is 10 times larger than the nominal noise.

Under the aforementioned conditions, we compare performances of RKF by using between the proposed design method and heuristic design method. Performances of KF are also evaluated for comparison. As a heuristic design method, we search regularization parameters which minimizes a summation of root mean square errors (RMSEs) of each estimate. However, since it is difficult to determine multi regularization parameters by the heuristic method, we search a single regularization parameter by the heuristic method. Additionally, in the heuristic design method, a suitable parameter depends on noise, and it is changed at each simulation. In the demonstration, we search the parameter 10 times, and select the average. We use $\lambda = 0.0124$ under the Cauchy noise, and $\lambda = 0.0273$ under the Gaussian mixture noise.

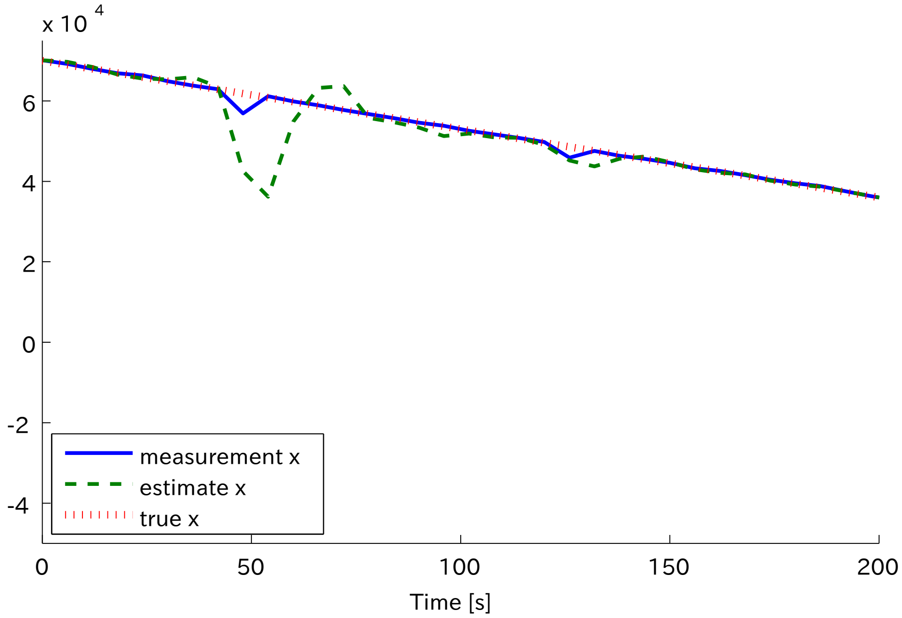
In this system, a structure of the system results in a diagonal matrix $\boldsymbol{W}$ in the optimization problem (2.12), so the regularization parameter is given by Eq. (3.11).

3.5.4 Results

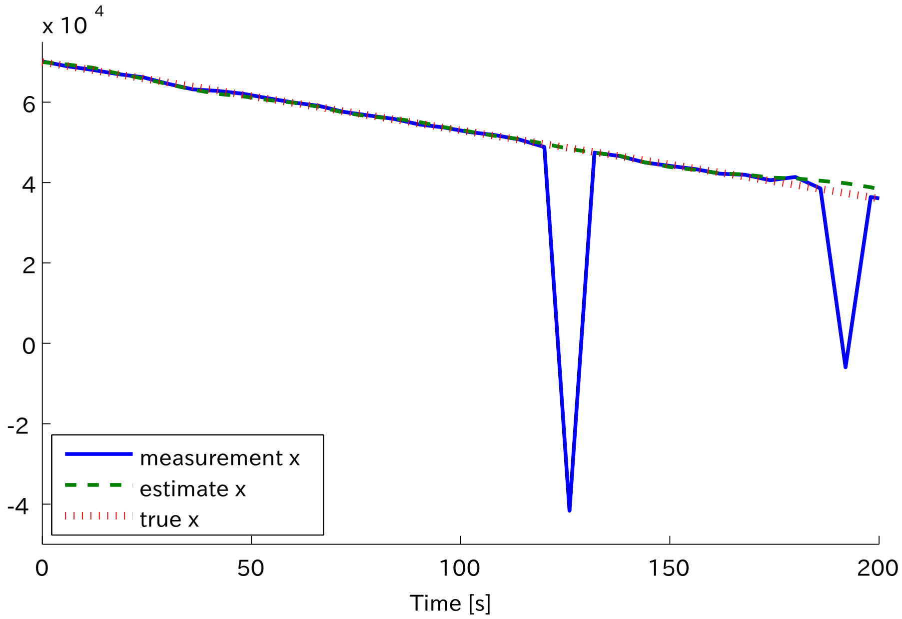
Fig. 3.3 – 3.8 show estimates of x, y, and z by using KF and RKF with the proposed design method. Fig. 3.3 – 3.5 show estimates under the Cauchy noise, and Fig. 3.6 – 3.8 show estimates under the Gaussian mixture noise. Under clutters, estimation errors of KF are very large. On the other hand, RKF can reduce effects of the clutters.

Fig. 3.9(a) shows RMSEs of estimates and average of all RMSEs under the Cauchy noise. Fig. 3.9(b) shows RMSEs under the Gaussian mixture noise. RMSEs are changed every simulations. In the demonstration, we average 10 times results.

Under the Cauchy noise, the proposed design method gives smaller means of estimation errors than the heuristic one except $\dot{z}$. An estimation error of $\dot{z}$ designed by the proposed method is only 4.35% larger than that designed by the heuristic one. And also, the proposed design method provides the smaller maximums of the errors than the heuristic one.

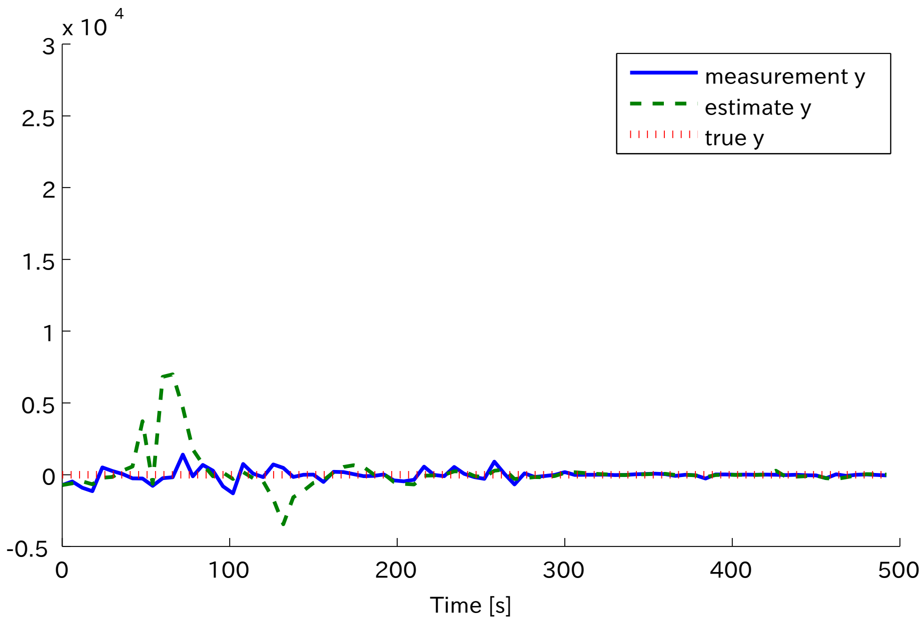


(a) using KF

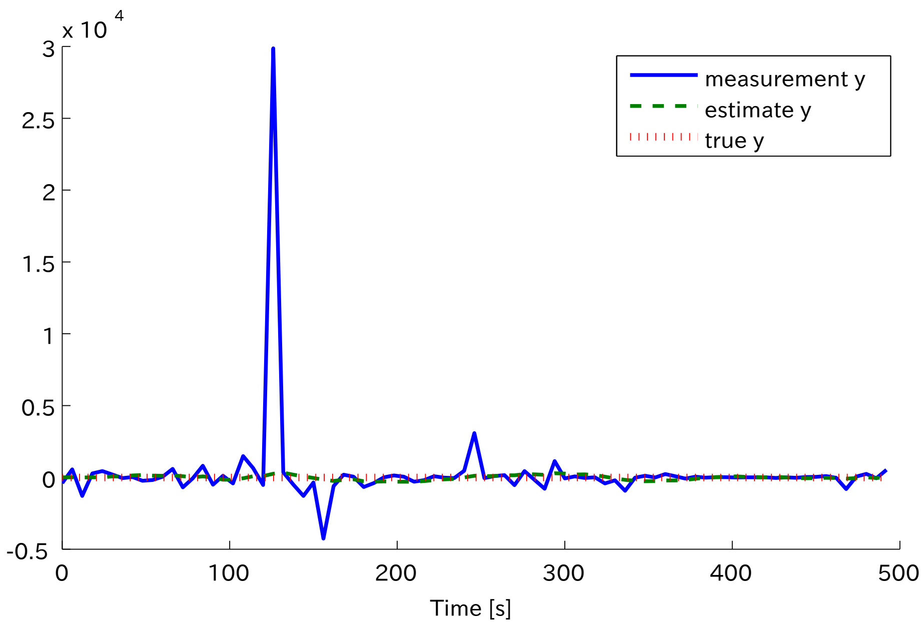


(b) using RKF designed by the proposed method

Figure 3.3: Estimates of x using KF and RKF designed by the proposed method under Cauchy distribution. The solid lines are measurements, dashed lines are estimates, and dotted lines are true signals.

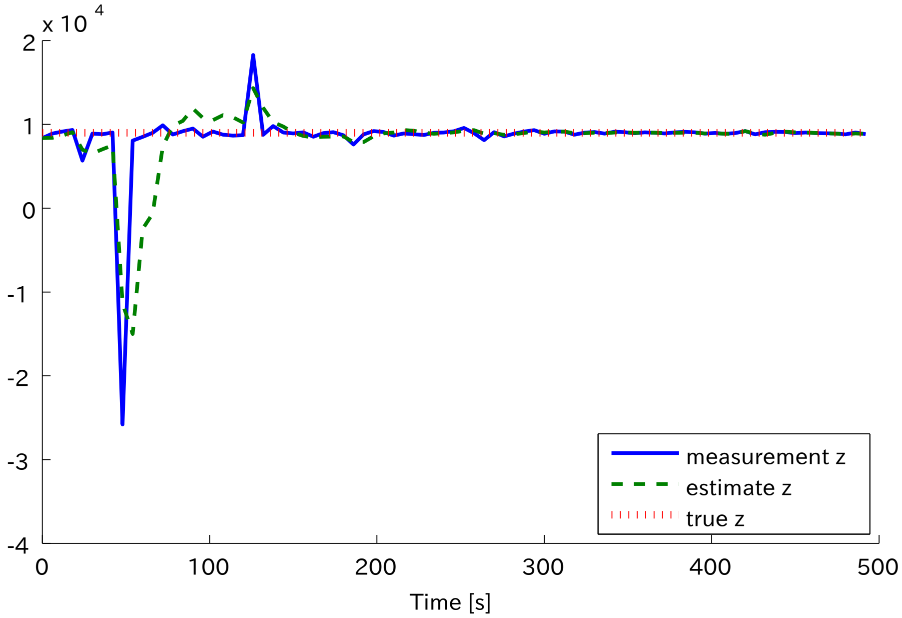


(a) using KF

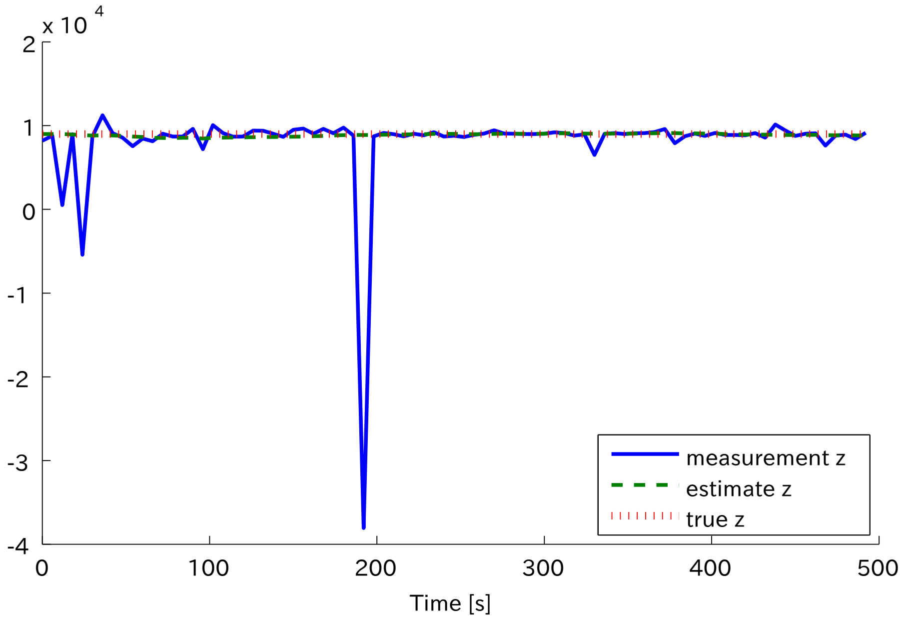


(b) using RKF designed by the proposed method

Figure 3.4: Estimates of y using KF and RKF designed by the proposed method under Cauchy distribution. The solid lines are measurements, dashed lines are estimates, and dotted lines are true signals.

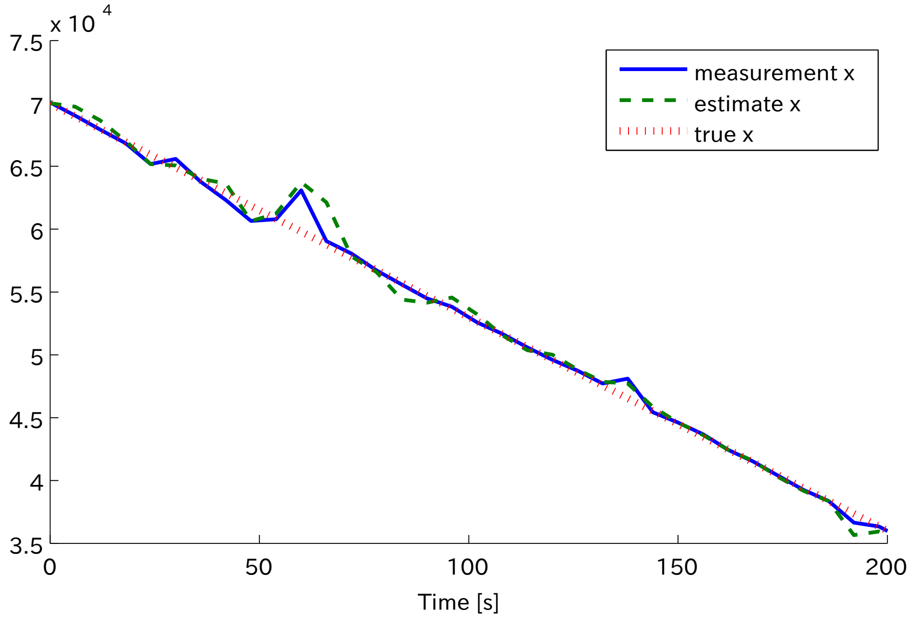


(a) using KF

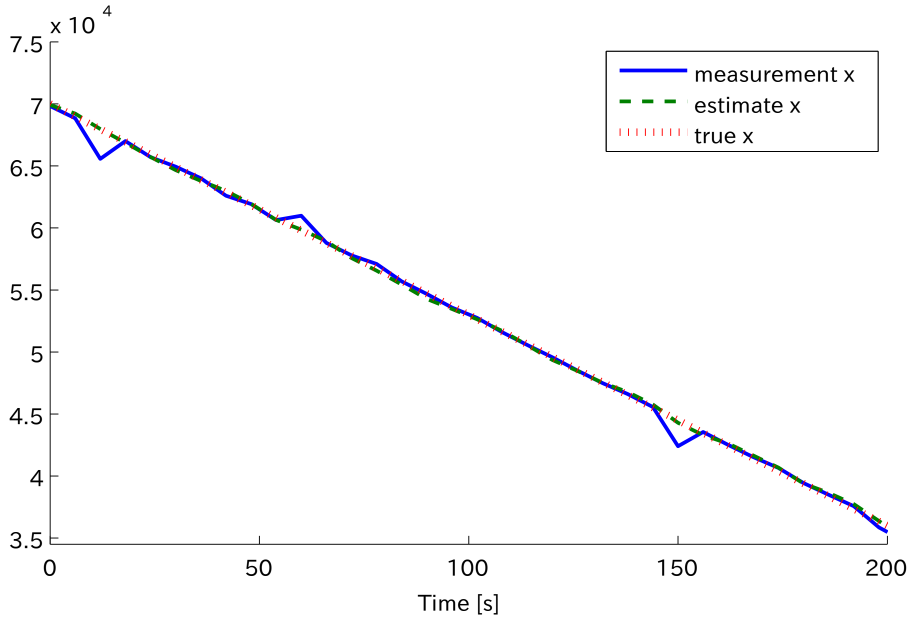


(b) using RKF designed by the proposed method

Figure 3.5: Estimates of z using KF and RKF designed by the proposed method under Cauchy distribution. The solid lines are measurements, dashed lines are estimates, and dotted lines are true signals.

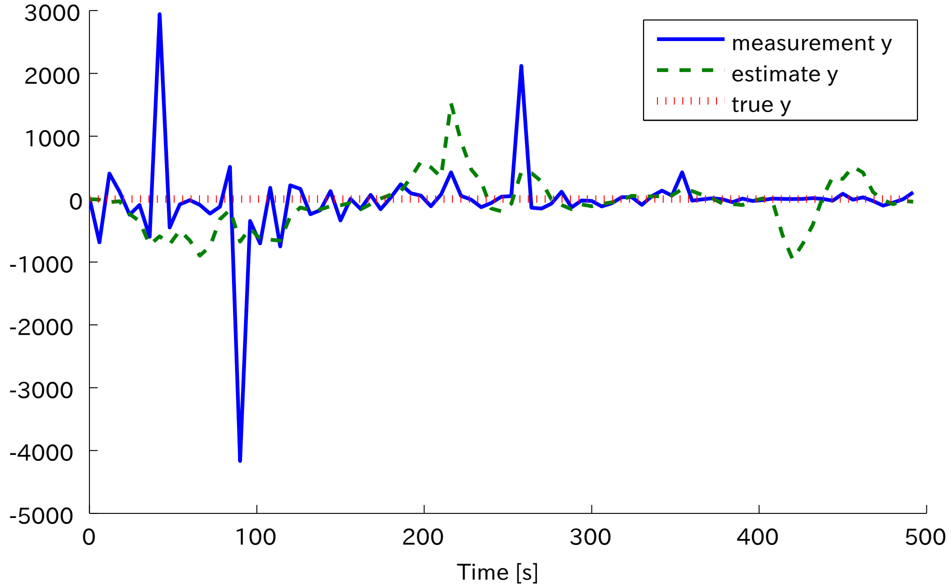


(a) using KF

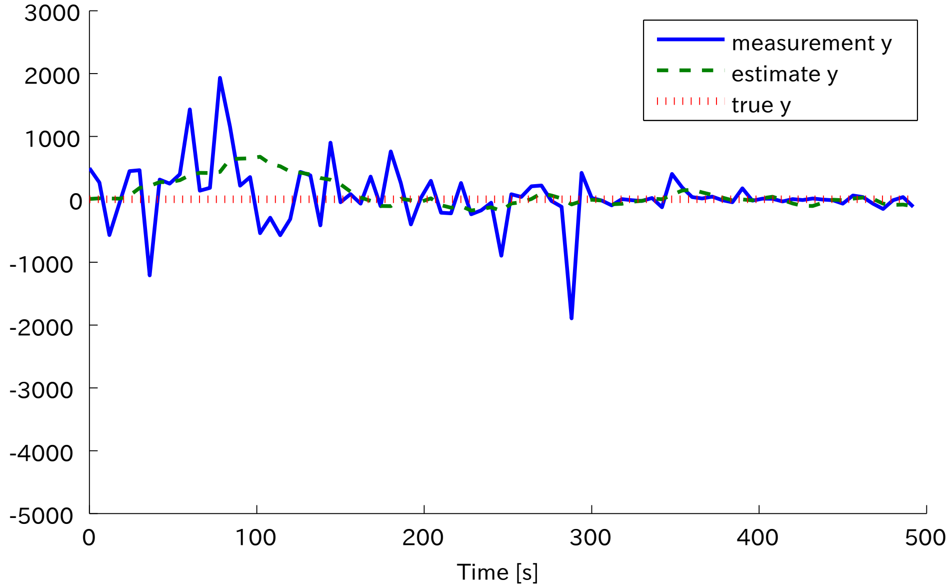


(b) using RKF designed by the proposed method

Figure 3.6: Estimates of x using KF and RKF designed by the proposed method under mixed Gaussian distribution. The solid lines are measurements, dashed lines are estimates, and dotted lines are true signals.

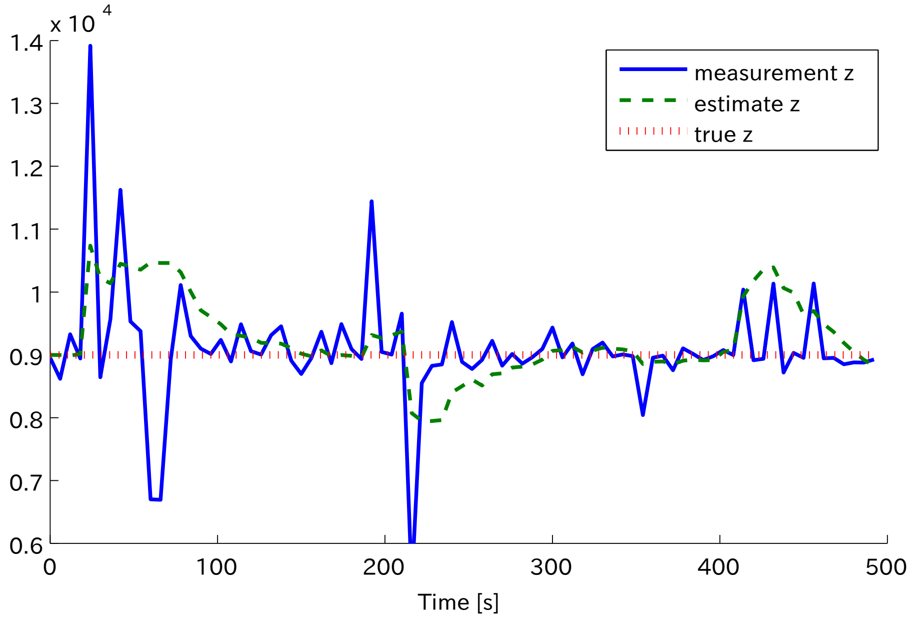


(a) using KF

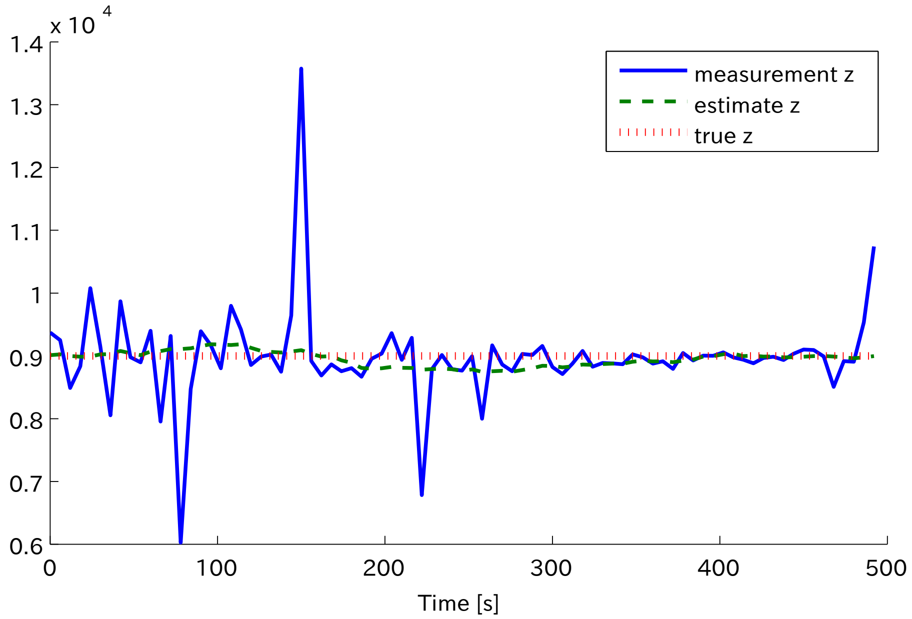


(b) using RKF designed by the proposed method

Figure 3.7: Estimates of y using KF and RKF designed by the proposed method under mixed Gaussian distribution. The solid lines are measurements, dashed lines are estimates, and dotted lines are true signals.

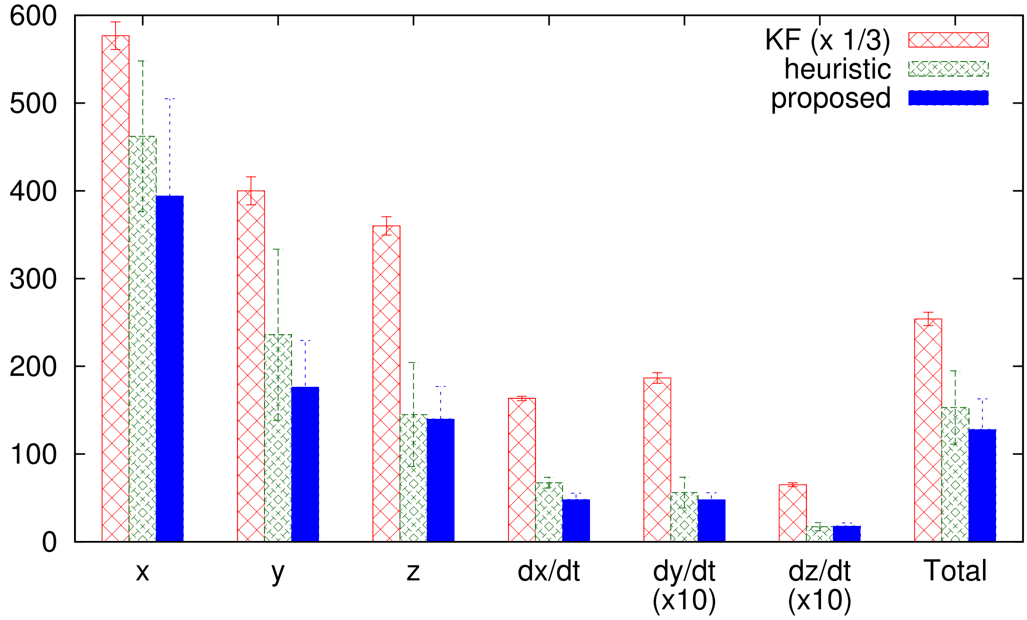


(a) using KF

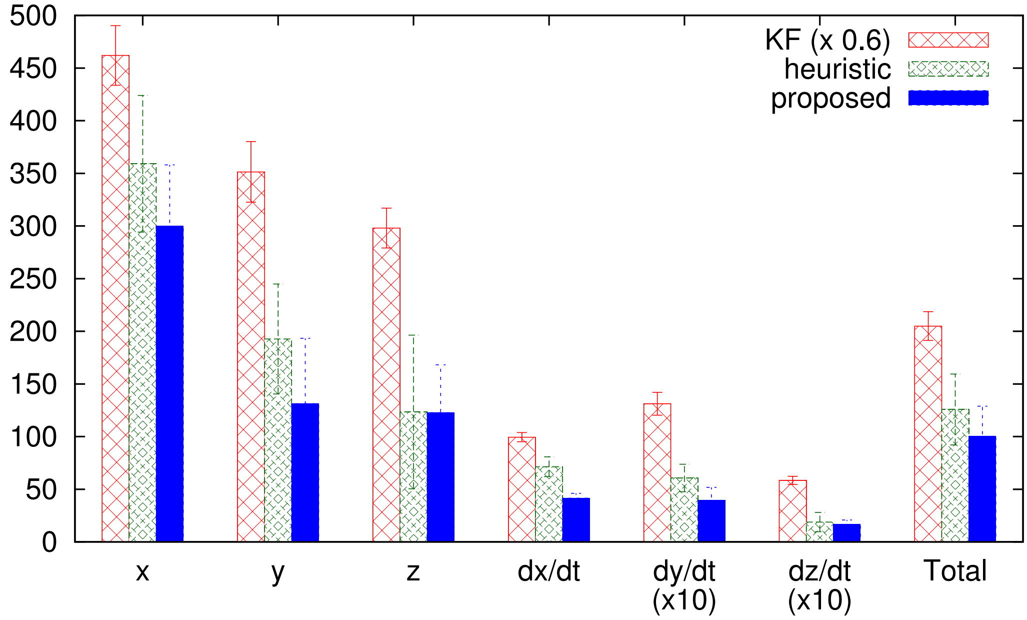


(b) using RKF designed by the proposed method

Figure 3.8: Estimates of z using KF and RKF designed by the proposed method under mixed Gaussian distribution. The solid lines are measurements, dashed lines are estimates, and dotted lines are true signals.



(a) using Cauchy distribution as a clutter model



(b) using mixed Gaussian distribution as a clutter model

Figure 3.9: Root mean square errors of estimates. Values using KF in (a) and (b) are multiplied by 1/3 and 0.6, respectively.

Under the Gaussian mixture noise, the proposed design method results in smaller means and maximums of estimation errors than the heuristic one.

Therefore, the proposed design method gives same or greater performances than the heuristic one. The reason why the proposed design method is superior to the heuristic one is because the regularization parameter with the heuristic design method is scalar, so it is possible that the parameter is determined to be conservative in noise with different variances. It is expected that multi parameters improve performances of RKF with the heuristic design method. However, it is difficult to search many parameters in heuristic senses. In contrast, the proposed design method can determine multi parameters systematically.

3.6 Application to Two-Wheeled Vehicle with Outliers

3.6.1 Problem Settings

In Section 3.5, the target tracking system is treated to evaluate estimation accuracy of the proposed method. In this section, feedback control performances using the estimates are evaluated. As a control problem, we consider a two-wheeled vehicle with a non-holonomic constraint shown in Fig. 3.10. We assume that positions and a angle, i.e., x, y, and θ, are obtained by sensors, and position sensors are contaminated by outliers. Examples of this situation are UGV with GPS and UAV using visual feedback [7, 8].

Let m and J be a mass of the vehicle and moment of inertia about the center of gravity, respectively. Let f and τ_θ denote a driving force in the direction of motion and steering torque, respectively. Assuming $\boldsymbol{q} = \begin{bmatrix} x & \theta & y \end{bmatrix}^T$ and $\boldsymbol{\tau} = \begin{bmatrix} f & \tau_\theta \end{bmatrix}^T$, a dynamic model of

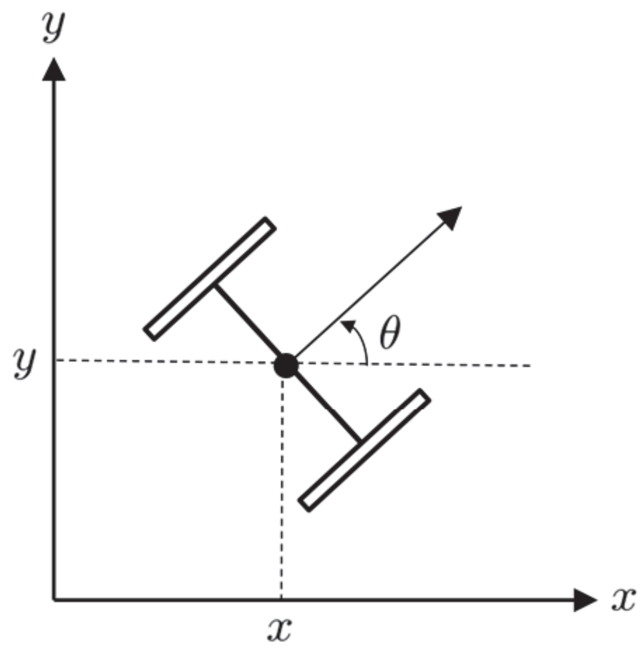


Figure 3.10: Model of two-wheeled vehicle with a non-holonomic constraint.

the vehicle is given by

$$\boldsymbol{M}(\boldsymbol{q})\ddot{\boldsymbol{q}} = \boldsymbol{E}(\boldsymbol{q})\boldsymbol{\tau}, \tag{3.22}$$

where

$$\begin{aligned} \boldsymbol{M}(\boldsymbol{q}) &= \begin{bmatrix} m & 0 & 0 \\ 0 & J & 0 \\ 0 & 0 & m \end{bmatrix}, \\ \boldsymbol{E}(\boldsymbol{q}) &= \begin{bmatrix} \cos\theta & 0 \\ 0 & 1 \\ \sin\theta & 0 \end{bmatrix}. \end{aligned}$$

The vehicle satisfies the following velocity constraint:

$$\dot{x}\sin\theta - \dot{y}\cos\theta = 0.$$

3.6.2 Control Law

We assume that acceleration inputs are given by

$$f = \frac{m\tan\theta\sec^2\theta}{1+\tan\theta}\dot{\theta}\dot{x} - \frac{m\alpha(\dot{x}-u_1)}{\cos\theta}, \tag{3.23}$$

$$\tau_\theta = -J\beta(\dot{\theta}-u_2). \tag{3.24}$$

Application of the acceleration inputs to Eq. (3.22) and increasing of α and β result in the following kinematic model [27]:

$$\dot{\boldsymbol{q}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \tan\theta & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}. \tag{3.25}$$

In this chapter, u_1 and u_2 are designed by the kinematic model (3.25), and the inputs are transformed into the acceleration inputs by Eq. (3.23) and (3.24). Many control laws for the kinematic model have been proposed [28]-[31]. In this chapter, we adopt a time-state control form [28, 29]. Applying the time-state control form to the vehicle gives a linearized system, so we can use linear control theories. In order to verify effectiveness of noise reduction, we selected a simple method.

Assume the following state and input transformations:

$$\begin{aligned} z_1 &= x, \\ z_2 &= \tan\theta, \\ z_3 &= y, \\ u_1 &= v_1, \\ u_2 &= v_2\cos^2\theta, \\ \mu_2 &= \frac{v_2}{v_1}. \end{aligned}$$

The kinematic model can be transformed to the following time-state control form:

$$\begin{aligned}
\dot{z}_1 &= v_1, \\
\frac{d}{dz_1}\begin{bmatrix} z_3 \\ z_2 \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} z_3 \\ z_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}\mu_2.
\end{aligned}$$

One of control laws derived from the time-state control form is as follows:

$$v_1 = \pm k_1 z_1, \tag{3.26}$$

$$v_2 = -k_3 z_3 v_1 - k_2 z_2 |v_1|, \tag{3.27}$$

where k_1, k_2, and k_3 are positive constants. In order to prevent a divergence of z_1, a sign of k_1 needs to be changed appropriately.

Similarly to section 3.5, application of RKF to a constant acceleration model estimates a state of the vehicle in this section.

3.6.3 Conditions

In the demonstration, we compare performances of a controlled vehicle using a heuristic and the proposed design method.

A mean of a nominal Gaussian measurement noise is $\mathbf{0}$, and the following two cases for its covariance matrix are considered:

Case I (same diagonal elements of a covariance matrix):

$$\mathbb{E}[\boldsymbol{v}_k] = \mathbf{0},$$

$$\mathbb{E}[\boldsymbol{v}_k \boldsymbol{v}_k^T] = \begin{bmatrix} 2.5 \times 10^{-3} & 0 & 0 \\ 0 & 2.5 \times 10^{-3} & 0 \\ 0 & 0 & 2.5 \times 10^{-3} \end{bmatrix}.$$

Case II (different diagonal elements of a covariance matrix):

$$\mathbb{E}[\boldsymbol{v}_k] = \mathbf{0},$$

$$\mathbb{E}[\boldsymbol{v}_k \boldsymbol{v}_k^T] = \begin{bmatrix} 2.5 \times 10^{-1} & 0 & 0 \\ 0 & 2.5 \times 10^{-3} & 0 \\ 0 & 0 & 2.5 \times 10^{-3} \end{bmatrix}.$$

Specially, the case II supposes a situation using different measurements like section 3.5, or using sensors with individual differences for noise property. Outliers are assumed to be distributed by a Cauchy distribution used in section 3.5. A center and width of the Cauchy distribution are 0 and $\delta = 5.0 \times 10^{-2}$, respectively.

For parameters of KF and RKF, $\boldsymbol{P}_{0|0} = 10\boldsymbol{I}$, $\sigma_{a_x}^2 = \sigma_{a_y}^2 = \sigma_{a_\theta}^2 = 5.0 \times 10^2$, and $\boldsymbol{R}$ is the same value as the aforementioned covariance matrix of measurement noise. In the heuristic design method, a suitable parameter depends on noise, and it is changed at each simulation. In the demonstration, we search the parameter 10 times, and select its average. We use $\lambda = 43.2$ in the case I, and $\lambda = 26.1$ in the case II.

Parameters of the two-wheeled vehicle are $m = 1.0$ and $J = 0.1$. An initial value of the vehicle is $x_0 = [2.5 \;\; 1.0 \;\; 0.0]^T$, and a target is origin. For parameters of the control law, we set $\alpha = 1$ and $\beta = 1$. Parameters of Eq. (3.26) and (3.27) are $k_1 = 1$, $k_2 = 1.1$, and $k_3 = 0.24$. A sign of Eq. (3.26) is changed every 3 second till $|z_2| + |z_3| < 5 \times 10^{-2}$. After the inequality condition is satisfied, we make z_1 converge exponentially.

3.6.4 Results

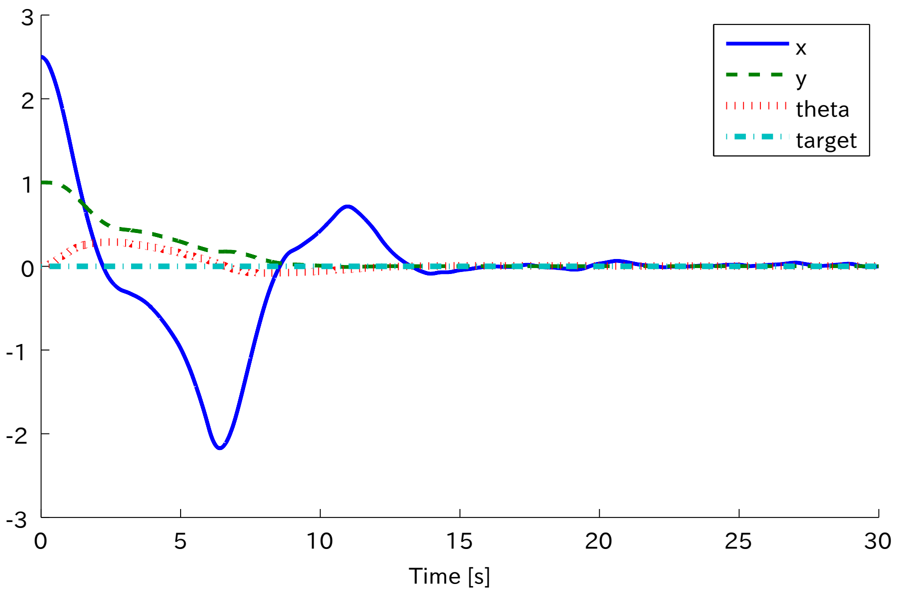
In the case I, Fig. 3.11 shows trajectories of a state of the controlled vehicle using RKF with heuristic and proposed design method. Fig. 3.12 shows absolute values of control errors of the controlled vehicle. It can be seen that the both responses are similar. Fig. 3.13 shows averaged absolute values of the errors before and after 20 second, i.e., transient and steady state responses. The values are averages and STDs of 10 times simulations in similar to section 3.5. These results show that RKF with the proposed design method has almost same performance as RKF with the heuristic one. Actually, a steady state error of y using the proposed design method is only 7.98% larger than that using the heuristic one, and numerical errors of the other values are within 1.69%. Use of the both design methods gives small STDs of control errors.

In the case II, Fig. 3.14 shows trajectories of a state of the controlled vehicle, and Fig. 3.15 shows absolute values of control errors. Fig. 3.16 shows averaged absolute values of the errors before and after 20 second. These results show that the proposed design method makes numerical errors except steady state errors of x and y within 2.08, and the steady state errors of x and y are improved 49.15% and 18.90%, respectively The proposed design method also gives smaller STDs than the heuristic one. The proposed design method gives almost same performances as or smaller errors than the heuristic one. Similarly to the section 3.5, it is possible that the heuristic design method makes the parameter to be conservative in noise with different variances.

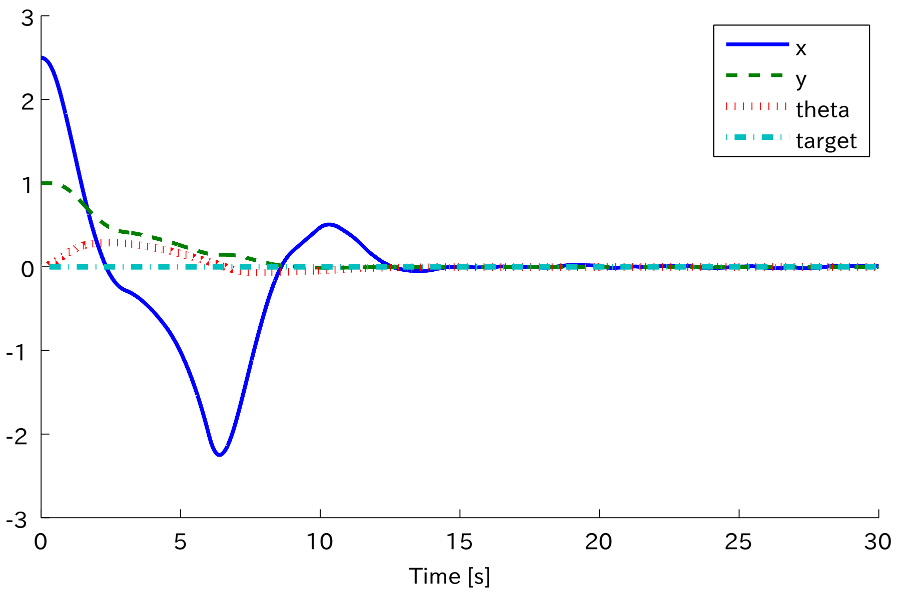
3.7 Conclusion

In this chapter, we proposed a new design method of RKF via l_1 regression. Regularization parameters of RKF are determined by statistics of Gaussian noise. This means that the proposed design method provides the parameters with physical meanings, and we can design the parameters systematically. We applied RKF with the proposed design method to a target tracking system with clutters and a control of a two-wheeled vehicle under outliers. Effectiveness was demonstrated by some numerical simulations.

We will verify the proposed method by real applications as our future works. We will also extend to non-linear systems in chapter 6. Moreover, RKF via l_1 regression needs to compute LMI and l_1 optimization problem. The algorithms need some numerical iterations in general, so a convergence rate and accuracy of the solutions depend on conditions of the iterations. A proposition of an efficient computation algorithm is in chapter 4.

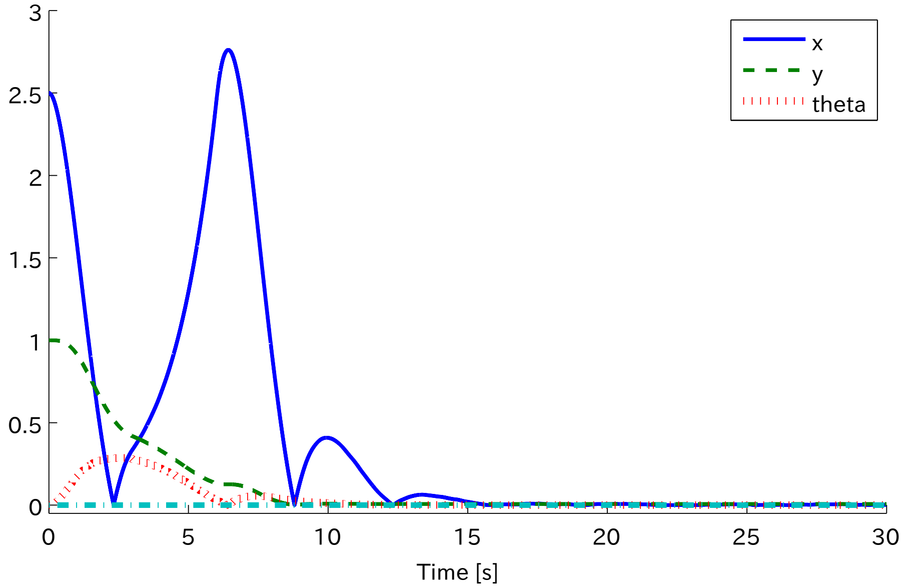


(a) positions using a heuristic design method

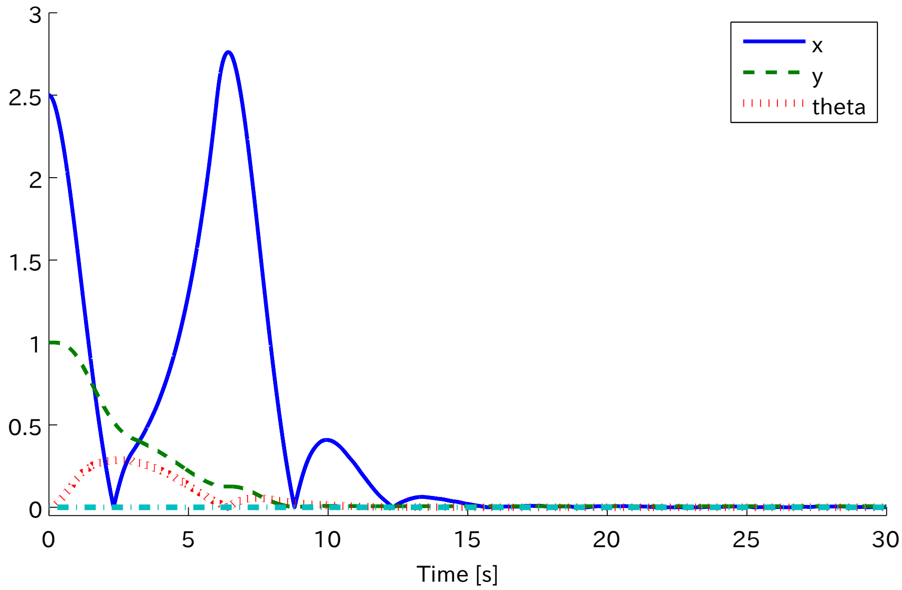


(b) positions using a proposed design method

Figure 3.11: Positions of the vehicle controlled by a time-state control form in the case I. (a) and (b) are results in using a heuristic and proposed design method, respectively. The solid line is x, dashed line is y, dotted line is θ, and dash-dotted line is the target.



(a) using heuristic design method



(b) using proposed design method

Figure 3.12: Absolute values of errors of the controlled vehicle in the case I. (a) and (b) are results in the case using a heuristic and proposed design method, respectively. The solid line is x, dashed line is y, and dotted line is θ.

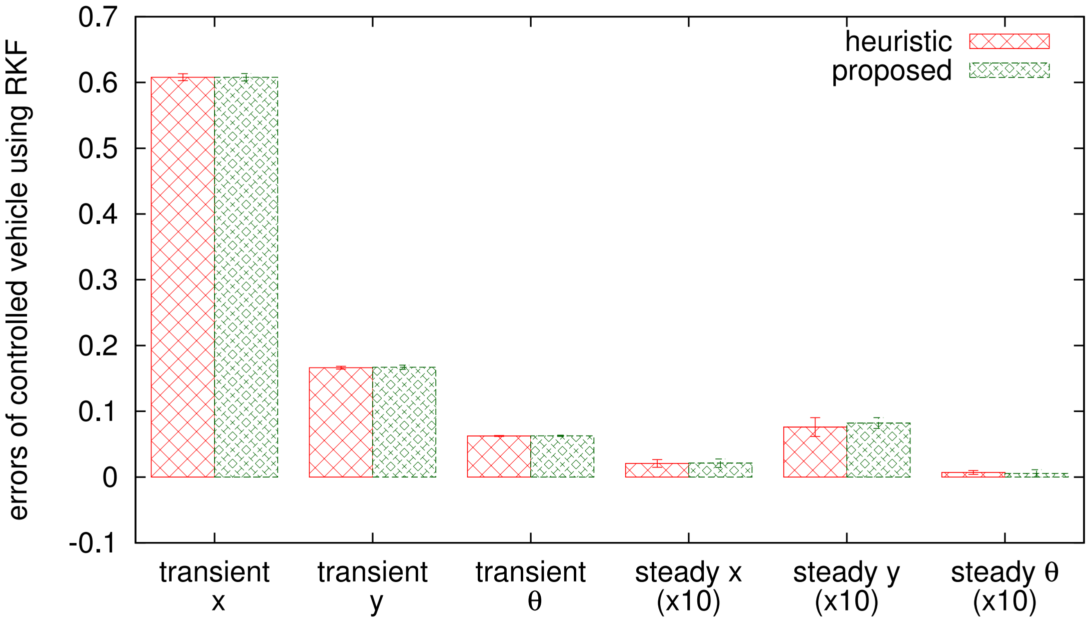
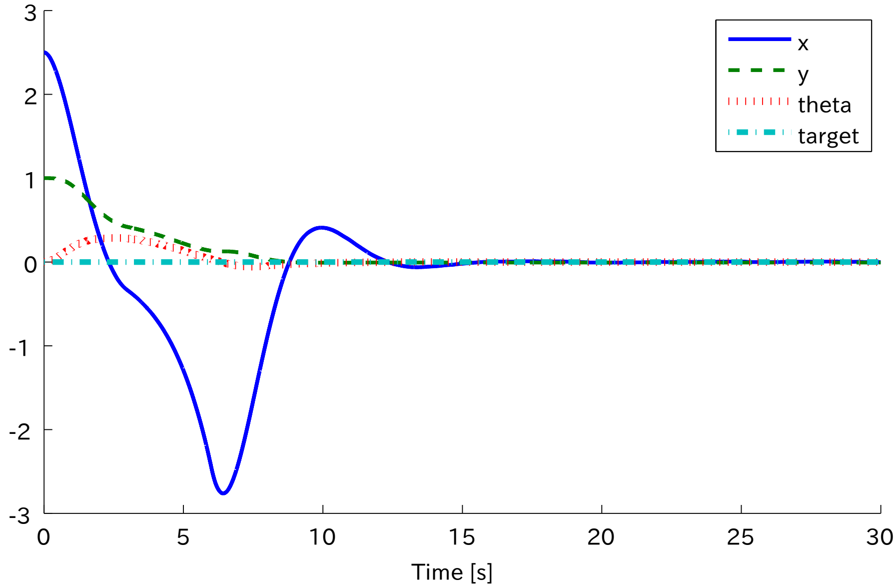
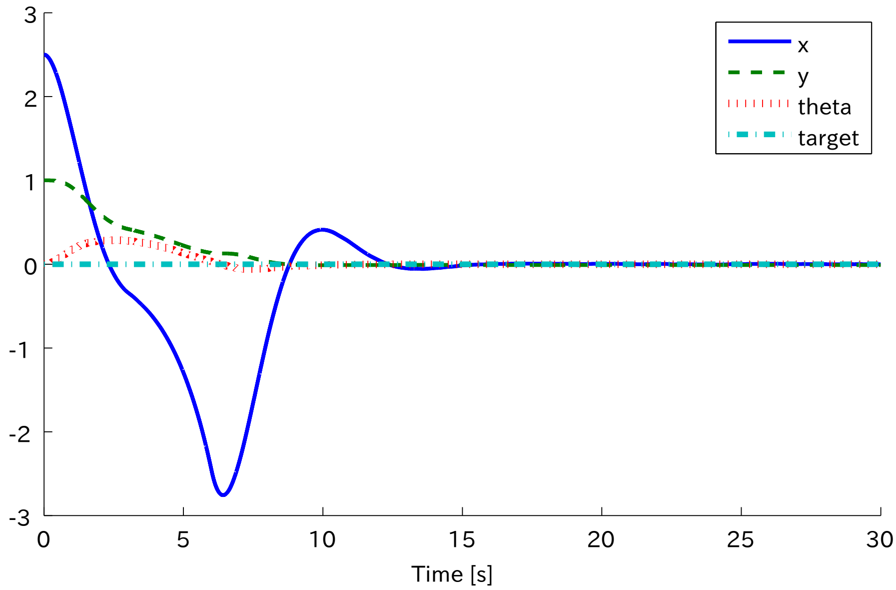


Figure 3.13: Transient errors and steady state errors in the case I.

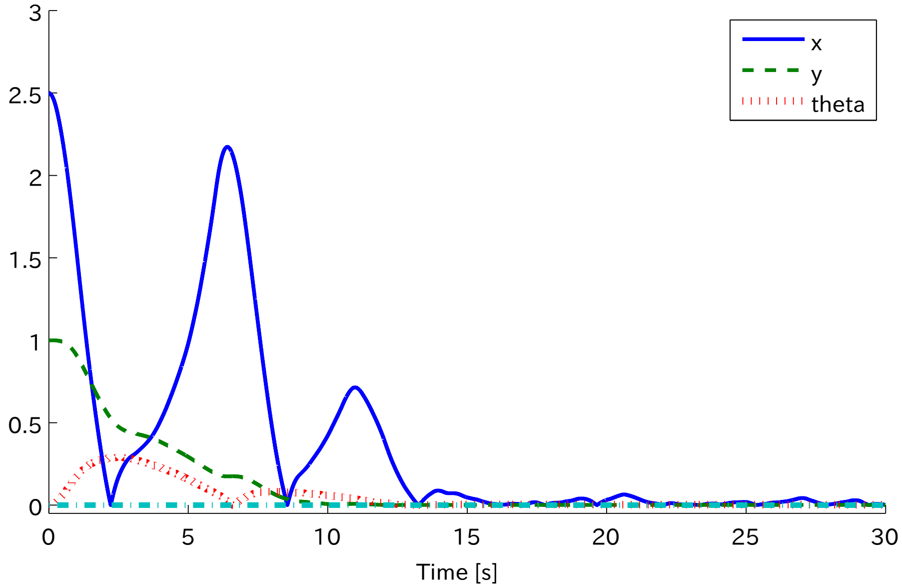


(a) positions using a heuristic design method

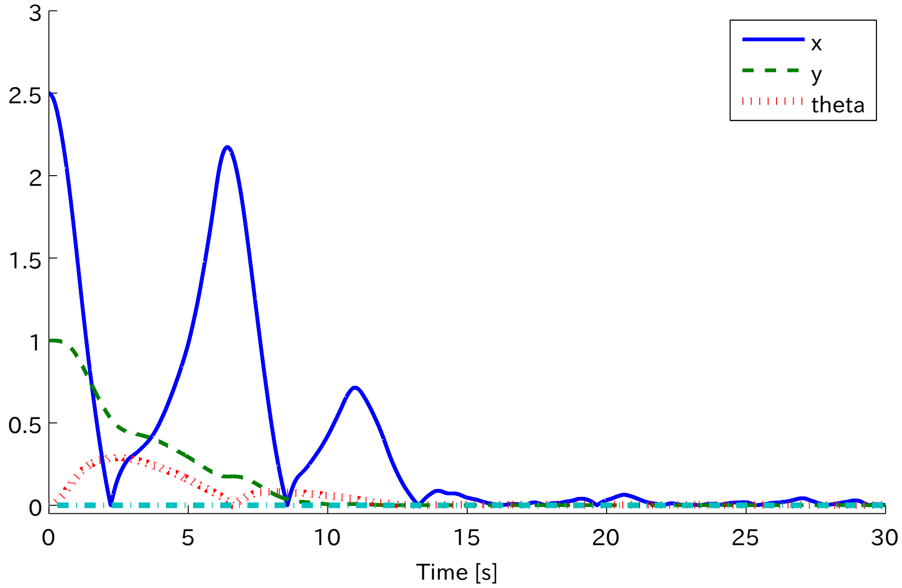


(b) positions using a proposed design method

Figure 3.14: Positions of the vehicle controlled by a time-state control form in the case II. (a) and (b) are results in using a heuristic and proposed design method, respectively. The solid line is x, dashed line is y, dotted line is θ, and dash-dotted line is the target.



(a) using heuristic design method



(b) using proposed design method

Figure 3.15: Absolute values of errors of the controlled vehicle in the case II. (a) and (b) are results in the case using a heuristic and proposed design method, respectively. The solid line is x, dashed line is y, and dotted line is θ.

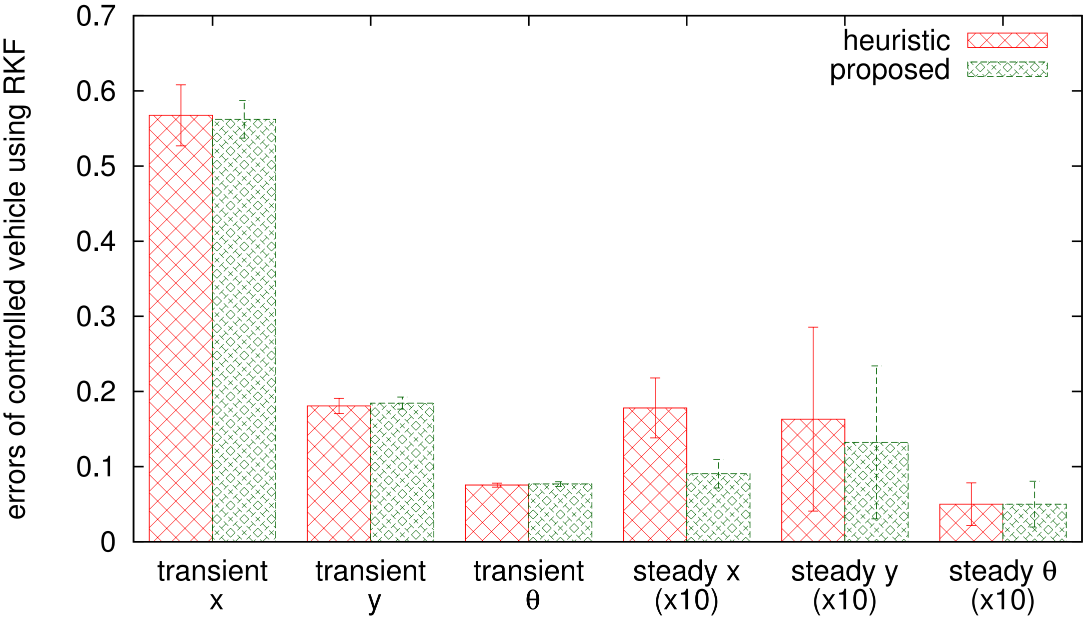


Figure 3.16: Transient errors and steady state errors in the case II.

Chapter 4

Fast Algorithm of Robust Kalman Filter via l_1 Regression by a Closed Form Solution

4.1 Introduction

Kalman filter (KF) is a well known global optimal estimation for linear systems under some conditions. The condition is that an estimation error of an initial state and all noises are distributed by normal distributions. For non-Gaussian measurement noise, KF is a unbiased minimum variance estimator, but not global optimal [32]. Many real applications include non-Gaussian measurement noise. For example, outlier is one of the major non-Gaussian measurement noise. In target tracking systems with a radar measurement, a reflection noise introduces outliers and the outliers are called clutter [6]. In UAV using visual feedback [7], temporary change of image contrast in background causes outliers of position data. Also in UGV using GPS [8], radio disturbances due to some obstacles provide position data with outliers.

Many algorithms of KF for outliers have been proposed. A major KF for non-Gaussian measurement noise including outliers is a Gaussian sum filter [34, 35]. Gaussian sum filter can approximate arbitrary distributions by using a Gaussian mixture distribution, so it can provide global optimal estimates. Particle filter [36] can also approximate arbitrary distributions by using a Monte Carlo method, and no prior information of distributions is needed. However, computational costs of the both methods are very high. In contrast, application of l_1 regression to KF gives a robust estimation under outliers and the method is called robust KF (RKF) via l_1 regression [20]. The method truncates the outliers by some thresholds generated by l_1 regression. Therefore, the method has little time delay to reduce effects of the outliers and it attracts many attentions. In chapter 3, the new systematic design method of the RKF was proposed. However, RKF via l_1 regression needs to compute LMI and l_1 optimization problem, so the performance of the RKF depends on algorithms to compute the problems in practice.

Many useful tools to compute LMI are existing. For example, CVX [20] and YALMIP [37] are description languages of convex optimization problems, and the languages can model LMI easily in MATLAB and the other softwares. SeDuMi [38] and SDPT3 [39] are solvers for optimization problems and implement some algorithms, e.g., an interior point method is the most famous algorithm. The algorithms need some numerical iterations in general, so a convergence rate and accuracy of the solutions depend on conditions of the iterations.

l_1 optimization problems can be formulated as a quadratic programming (QP) optimization problems, and CVX can also model the problems. Moreover, CVXGEN [40] can generate custom C codes of the QP problems for online computations. Fast iterative shrinkage thresholding algorithm (FISTA) are proposed as effective computation methods for l_1 optimization problems [33]. However, similarly to computations of LMI, these methods demand some iterative methods. Homotopy method is also proposed and can compute a solution in a closed form [41], but the method uses all past data.

In this chapter, we derive upper and lower bounds of an optimal estimate of the RKF and compute an approximated estimate using the both bounds. In addition, the approximated estimate is given by a closed form, so no iteration is needed and it gives a fast computation.

This chapter is organized as follows: In section 2, RKF via l_1 regression is explained briefly. In section 3, a closed form solution of the RKF is proposed and estimation errors of the algorithm are analyzed. In section 4, some numerical simulations demonstrate effectiveness of the proposed algorithm. Conclusion is given in section 5.

4.2 Robust Kalman Filter via l_1 Regression

4.2.1 Formula

In this section, RKF via l_1 regression is explained briefly, again.

Consider the LTI system (2.6). Given $\boldsymbol{Q}$, $\boldsymbol{R}$, and an initial value of $\boldsymbol{P}$, i.e., $\boldsymbol{P}_{0|0}$, an update law of RKF via l_1 regression are expressed as

$$\begin{cases} \boldsymbol{L} = \boldsymbol{P}_{k|k-1}\boldsymbol{C}^T(\boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{R})^{-1}, \\ \boldsymbol{e}_k = \boldsymbol{y}_k - \boldsymbol{C}\hat{\boldsymbol{x}}_{k|k-1}, \\ \hat{\boldsymbol{x}}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{L}(\boldsymbol{e}_k - \boldsymbol{z}_k^*), \\ \boldsymbol{P}_{k|k} = (\boldsymbol{I} - \boldsymbol{L}\boldsymbol{C})\boldsymbol{P}_{k|k-1}, \end{cases} \tag{4.1}$$

where $\boldsymbol{z}_k^*$ is given by a solution of the following optimization problem with l_1 regression:

$$\boldsymbol{z}_k^* = \arg\min_{\boldsymbol{z}_k} \ (\boldsymbol{e}_k - \boldsymbol{z}_k)^T \boldsymbol{W} (\boldsymbol{e}_k - \boldsymbol{z}_k) + \sum_{i=1}^{m} \lambda_i |z_{k,i}|, \tag{4.2}$$

where $\boldsymbol{\lambda} = [\lambda_1, \cdots, \lambda_m]^T \in \mathbb{R}^m$ is a regularization parameter, and $\boldsymbol{W}$ is the following positive definite matrix:

$$\boldsymbol{W} = \left(\boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{R}\right)^{-1}. \tag{4.3}$$

Chapter 3 showed that $\boldsymbol{\lambda}$ can be designed by the covariance matrices, $\boldsymbol{Q}$ and $\boldsymbol{R}$, systematically. In chapter 3, an estimate of $\boldsymbol{z}_k$ was expressed as $\hat{\boldsymbol{z}}_k$. However, this chapter represents the estimate as $\boldsymbol{z}_k^*$ instead of $\hat{\boldsymbol{z}}_k$.

Generally, the optimization problem (4.2) requires iterative computation methods like an interior point method. However, a convergence rate and accuracy of solutions depend on conditions of the iterative methods. For example, accuracy of a gradient method depends on its step size. And also, solutions rely on stop conditions in any algorithms. Parameters in iterative algorithms need to be tuned for each application and the tuning is heuristic in general. In this chapter, we propose a new algorithm without iterations by an approximation of the optimal solution, and we analyze performances of the algorithm.

4.2.2 Condition to Design Regularization Parameters of RKF

From section 3.2, a condition to design regularization parameters of RKF is as follows:

$$\begin{bmatrix} \lambda_1^2 & & \\ & \ddots & \\ & & \lambda_m^2 \end{bmatrix} \geq 4\boldsymbol{W}. \tag{4.4}$$

Therefore, a solution of LMI (4.4) determines the regularization parameter $\boldsymbol{\lambda}$.

Since a sparse solution, i.e., $z_{k,i}^* = 0$, can be often obtained more than necessary if λ_i is large, $\boldsymbol{\lambda}$ should be determined in a small residual of both sides of Eq. (4.4). One of the solutions is given by the following semi-definite programming (SDP):

$$\begin{aligned} \min_{\lambda_1^2,\cdots,\lambda_m^2} \ & \lambda_1^2 + \cdots + \lambda_m^2, \\ \text{s.t.} \ & \begin{bmatrix} \lambda_1^2 & & \\ & \ddots & \\ & & \lambda_m^2 \end{bmatrix} \geq 4\boldsymbol{W}. \end{aligned} \tag{4.5}$$

Remark 4.1 In general, Eq. (4.2) cannot be solved analytically, and some iterative methods are needed. Also in Eq. (4.5), iterative methods are required to solve the LMI. □

4.3 Fast Algorithm of Robust Kalman Filter by a Closed Form Solution

4.3.1 Derivation

Let $\hat{\boldsymbol{z}}_k^*$ be an approximated solution of $\boldsymbol{z}_k^*$. This section shows that use of upper and lower bounds of $\boldsymbol{z}_k^*$ provides an approximated solution $\hat{\boldsymbol{z}}_k^*$, and $\hat{\boldsymbol{z}}_k^*$ can be written in a closed form.

The following inequality is sufficient to satisfy Eq. (4.4):

$$\begin{aligned} \boldsymbol{\Lambda} &\geq 2\sqrt{\boldsymbol{W}}, \\ \boldsymbol{\Lambda} &= \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{bmatrix}, \end{aligned}$$

where $\sqrt{\boldsymbol{W}}\left(\sqrt{\boldsymbol{W}}\right)^T = \boldsymbol{W}$, and it can be computed by Cholesky decomposition, so $\sqrt{\boldsymbol{W}}$ becomes a lower triangle matrix. The first order necessary condition of an optimality condition can be rewritten as

$$\boldsymbol{W}\left(\boldsymbol{e}_k - \boldsymbol{z}_k^*\right) = \frac{1}{2}\boldsymbol{\Lambda}\frac{\partial||\boldsymbol{z}_k^*||_1}{\partial \boldsymbol{z}_k^*}.$$

$$\therefore \left(\frac{\partial ||\boldsymbol{z}_k^*||_1}{\partial \boldsymbol{z}_k^*}\right)^T \boldsymbol{W}(\boldsymbol{e}_k - \boldsymbol{z}_k^*) = \frac{1}{2}\left(\frac{\partial ||\boldsymbol{z}_k^*||_1}{\partial \boldsymbol{z}_k^*}\right)^T \boldsymbol{\Lambda}\frac{\partial ||\boldsymbol{z}_k^*||_1}{\partial \boldsymbol{z}_k^*}.$$

Therefore, the following inequality is satisfied:

$$\left(\frac{\partial ||\boldsymbol{z}_k^*||_1}{\partial \boldsymbol{z}_k^*}\right)^T \boldsymbol{W}(\boldsymbol{e}_k - \boldsymbol{z}_k^*) \geq \left(\frac{\partial ||\boldsymbol{z}_k^*||_1}{\partial \boldsymbol{z}_k^*}\right)^T \sqrt{\boldsymbol{W}}\frac{\partial ||\boldsymbol{z}_k^*||_1}{\partial \boldsymbol{z}_k^*},$$

$$\left(\left(\sqrt{\boldsymbol{W}}\right)^T \frac{\partial ||\boldsymbol{z}_k^*||_1}{\partial \boldsymbol{z}_k^*}\right)^T \left(\left(\sqrt{\boldsymbol{W}}\right)^T (\boldsymbol{e}_k - \boldsymbol{z}_k^*) - \frac{\partial ||\boldsymbol{z}_k^*||_1}{\partial \boldsymbol{z}_k^*}\right) \geq 0, \tag{4.6}$$

where $\left(\sqrt{\boldsymbol{W}}\right)^T$ is an upper triangle matrix and represented by the following equation:

$$\left(\sqrt{\boldsymbol{W}}\right)^T = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1m} \\ 0 & w_{22} & \ddots & w_{2m} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & w_{mm} \end{bmatrix}. \tag{4.7}$$

A sufficient condition for Eq. (4.6) is as follows:

$$\begin{cases} \left(\left(\sqrt{\boldsymbol{W}}\right)^T \frac{\partial ||\boldsymbol{z}_k^*||_1}{\partial \boldsymbol{z}_k^*}\right)_i \geq 0, \\ \left(\left(\sqrt{\boldsymbol{W}}\right)^T (\boldsymbol{e}_k - \boldsymbol{z}_k^*) - \frac{\partial ||\boldsymbol{z}_k^*||_1}{\partial \boldsymbol{z}_k^*}\right)_i \geq 0, \end{cases} \tag{4.8}$$

or

$$\begin{cases} \left(\left(\sqrt{\boldsymbol{W}}\right)^T \frac{\partial ||\boldsymbol{z}_k^*||_1}{\partial \boldsymbol{z}_k^*}\right)_i \leq 0, \\ \left(\left(\sqrt{\boldsymbol{W}}\right)^T (\boldsymbol{e}_k - \boldsymbol{z}_k^*) - \frac{\partial ||\boldsymbol{z}_k^*||_1}{\partial \boldsymbol{z}_k^*}\right)_i \leq 0, \end{cases} \tag{4.9}$$

where $(\cdot)_i$ is represented as a i-th element of a vector.

First, consider a case of $i = m$. Assuming that diagonal elements of Eq. (4.7) are selected to be positive, conditions (4.8) and (4.9) result in the following inequalities:

$$z_{k,m}^* \leq e_{k,m} - w_{mm}^{-1}\frac{\partial |z_{k,m}^*|}{z_{k,m}^*}, \quad \frac{\partial |z_{k,m}^*|}{z_{k,m}^*} \geq 0, \tag{4.10}$$

$$z_{k,m}^* \geq e_{k,m} - w_{mm}^{-1}\frac{\partial |z_{k,m}^*|}{z_{k,m}^*}, \quad \frac{\partial |z_{k,m}^*|}{z_{k,m}^*} \leq 0. \tag{4.11}$$

Right hand sides of the inequalities are interpreted as computations of upper and lower bounds of $z_{k,m}^*$. Let $\bar{z}_{k,m}^* \geq 0$ and $\underline{z}_{k,m}^* \leq 0$ be the upper and lower bounds of $z_{k,m}^*$, respectively. Assuming that signs of the upper and lower bounds are equal to one of the optimal solution, the both bounds can be calculated by the following equations, respectively:

$$\bar{z}^*_{k,m} = \begin{cases} e_{k,m} - w^{-1}_{mm} & e_{k,m} > w^{-1}_{mm} \\ 0 & \text{otherwise} \end{cases}, \tag{4.12}$$

$$\underline{z}^*_{k,m} = \begin{cases} e_{k,m} + w^{-1}_{mm} & e_{k,m} < -w^{-1}_{mm} \\ 0 & \text{otherwise} \end{cases}. \tag{4.13}$$

Since the both Eq. (4.12) and (4.13) are 0 in a common domain, an estimate of $z^*_{k,m}$, i.e., $\hat{z}^*_{k,m}$, is defined as

$$\begin{aligned} \hat{z}^*_{k,m} &= \bar{z}^*_{k,m} + \underline{z}^*_{k,m} \\ &= \begin{cases} e_{k,m} - w^{-1}_{mm} & e_{k,m} > w^{-1}_{mm} \\ 0 & \text{otherwise} \\ e_{k,m} + w^{-1}_{mm} & e_{k,m} < -w^{-1}_{mm} \end{cases}. \end{aligned} \tag{4.14}$$

Assume that elements from $i+1$ to m are calculated. The condition (4.8) provides the following condition for a i-th element:

$$\begin{cases} \frac{\partial |z^*_{k,i}|}{z^*_{k,i}} \geq -\frac{1}{w_{ii}} \sum_{j=i+1}^{m} w_{ij} \frac{\partial |\hat{z}^*_{k,j}|}{\hat{z}^*_{k,j}}, \\ z^*_{k,i} \leq e'_{k,i} - w^{-1}_{ii} \frac{\partial |z^*_{k,i}|}{z^*_{k,i}}, \\ e'_{k,i} = e_{k,i} + \frac{1}{w_{ii}} \sum_{j=i+1}^{m} w_{ij}(e_{k,j} - \hat{z}^*_{k,j}). \end{cases} \tag{4.15}$$

In the same way as $i = m$, Eq. (4.15) means a calculation of an upper bound of $z^*_{k,i}$, i.e., $\bar{z}^*_{k,i} \geq 0$. Note that $\bar{z}^*_{k,i} > 0$ gives $\frac{\partial |\bar{z}^*_{k,i}|}{\bar{z}^*_{k,i}} = 1$. Assuming that signs of $z^*_{k,i}$ and $\bar{z}^*_{k,i}$ are same, $\bar{z}^*_{k,i}$ can be computed by

$$\begin{aligned} &\text{if } -\frac{1}{w_{ii}} \sum_{j=i+1}^{m} w_{ij} \frac{\partial |\hat{z}^*_{k,j}|}{\hat{z}^*_{k,j}} \leq 1 \\ &\qquad \text{then } \bar{z}^*_{k,i} = \begin{cases} e'_{k,i} - w^{-1}_{ii} & e'_{k,i} > w^{-1}_{ii} \\ 0 & \text{otherwise} \end{cases} \\ &\qquad \text{else } \bar{z}^*_{k,i} = 0. \end{aligned} \tag{4.16}$$

where, for convenience, $\bar{z}^*_{k,i} = 0$ if the condition (4.8) is not satisfied. Similarly, note that $\frac{\partial |\underline{z}^*_{k,i}|}{\underline{z}^*_{k,i}} = -1$ for $\underline{z}^*_{k,i} < 0$, $\underline{z}^*_{k,i}$ is given by the following equation:

$$\begin{aligned} &\text{if } -\frac{1}{w_{ii}} \sum_{j=i+1}^{m} w_{ij} \frac{\partial |\hat{z}^*_{k,j}|}{\hat{z}^*_{k,j}} \geq -1 \\ &\qquad \text{then } \underline{z}^*_{k,i} = \begin{cases} e'_{k,i} + w^{-1}_{ii} & e'_{k,i} < -w^{-1}_{ii} \\ 0 & \text{otherwise} \end{cases} \\ &\qquad \text{else } \underline{z}^*_{k,i} = 0. \end{aligned} \tag{4.17}$$

Algorithm 1 Fast algorithm of measurement update of robust Kalman filter via l_1 regression at time k

$\boldsymbol{e}_k = \boldsymbol{y}_k - \boldsymbol{C}\hat{\boldsymbol{x}}_{k|k-1}$
$\boldsymbol{W} = (\boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{R})^{-1}$
compute $\left(\sqrt{\boldsymbol{W}}\right)^T$ using Cholesky decomposition, where w_{ij} is a (i,j)-th element of $\left(\sqrt{\boldsymbol{W}}\right)^T$
$\hat{z}^*_{k,m} = \max(|e_{k,m}| - w^{-1}_{mm}, 0)\ \mathrm{sign}(e_{k,m})$
for $i = m-1$ down to 1 **do**
  $e'_{k,i} = e_{k,i} - \frac{1}{w_{ii}}\sum_{j=i+1}^{m} \sigma_{ij}\frac{\partial|\hat{z}^*_{k,j}|}{\partial\hat{z}^*_{k,j}}$.
  if $-\frac{1}{w_{ii}}\sum_{j=i+1}^{m} w_{ij}\frac{\partial|\hat{z}^*_{k,j}|}{\hat{z}^*_{k,j}} > 1$ **then**
    $\overline{z}^*_{k,i} = 0$
  else
    $\overline{z}^*_{k,i} = \max\left(e'_{k,i} - w^{-1}_{ii}, 0\right)$
  end if
  if $-\frac{1}{w_{ii}}\sum_{j=i+1}^{m} w_{ij}\frac{\partial|\hat{z}^*_{k,j}|}{\hat{z}^*_{k,j}} < -1$ **then**
    $\underline{z}^*_{k,i} = 0$
  else
    $\underline{z}^*_{k,i} = \min\left(e'_{k,i} + w^{-1}_{ii}, 0\right)$
  end if
  $\hat{z}^*_{k,i} = \overline{z}^*_{k,i} + \underline{z}^*_{k,i}$.
end for
$\boldsymbol{L} = \boldsymbol{P}_{k|k-1}\boldsymbol{C}^T\boldsymbol{W}$
$\hat{\boldsymbol{x}}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{L}(\boldsymbol{e}_k - \hat{\boldsymbol{z}}^*_k)$
$\boldsymbol{P}_{k|k} = (\boldsymbol{I} - \boldsymbol{L}\boldsymbol{C})\boldsymbol{P}_{k|k-1}$

Also in this case, the both Eq. (4.16) and (4.17) become 0 in a common domain. Therefore, $\hat{z}^*_{k,i}$ is defined as

$$\hat{z}^*_{k,i} = \overline{z}^*_{k,i} + \underline{z}^*_{k,i}. \tag{4.18}$$

Algorithm 1 shows a fast algorithm of RKF by a closed form computation.

4.3.2 Analysis of an Estimation Error of Outliers and Innovation of RKF

In order to show performances of the proposed algorithm, an estimation error of the solution is analyzed. Moreover, an innovation of RKF using the solution is also analyzed.

For the proposed algorithm, the following theorem is satisfied.

Theorem 4.1 Assuming that the following condition is satisfied,

$$-1 \le -\frac{1}{w_{ii}}\sum_{j=i+1}^{m} w_{ij}\frac{\partial|\hat{z}^*_{k,j}|}{\partial\hat{z}^*_{k,j}} \le 1, \quad \forall\hat{z}^*_{k,i}, \tag{4.19}$$

then covariance matrices of an estimation error of outliers and innovation are given by

$$\mathbb{E}\left[(\hat{\boldsymbol{z}}_k^* - \boldsymbol{z}_k)(\hat{\boldsymbol{z}}_k^* - \boldsymbol{z}_k)^T\right] \leq 2\left(\boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{R}\right), \tag{4.20}$$

$$\mathbb{E}\left[(\boldsymbol{e}_k - \hat{\boldsymbol{z}}_k^*)(\boldsymbol{e}_k - \hat{\boldsymbol{z}}_k^*)^T\right] \leq \boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{R}. \tag{4.21}$$

□

Proof: If Eq. (4.19) is satisfied, Eq. (4.18) can be simplified as

$$\hat{z}_{k,i}^* = \begin{cases} e'_{k,i} - w_{ii}^{-1} & e'_{k,i} > w_{ii}^{-1} \\ 0, & \text{otherwise} \\ e'_{k,i} + w_{ii}^{-1} & e'_{k,i} < -w_{ii}^{-1} \end{cases}. \tag{4.22}$$

Eq. (4.22) is equivalent to a solution of the following equation:

$$\left(\sqrt{\boldsymbol{W}}\right)^T (\boldsymbol{e}_k - \hat{\boldsymbol{z}}_k^*) = \frac{\partial||\hat{\boldsymbol{z}}_k^*||}{\partial \hat{\boldsymbol{z}}_k^*}. \tag{4.23}$$

For an estimation error of outliers, Eq. (2.6) and (4.23) yield the following equation:

$$\hat{\boldsymbol{z}}_k^* - \boldsymbol{z}_k = -\left(\sqrt{\boldsymbol{W}}\right)^{-T} \frac{\partial||\hat{\boldsymbol{z}}_k^*||_1}{\partial \hat{\boldsymbol{z}}_k^*} + \boldsymbol{C}(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}) + \boldsymbol{v}_k.$$

Note that $\frac{\partial||\hat{\boldsymbol{z}}_k^*||_1}{\partial \hat{\boldsymbol{z}}_k^*}$ is a vector consisting of sub-gradients. This means that each element of the vector can be interpreted as mutually independent stochastic variable which is in $[-1, 1]$. Since $\mathbb{E}\left[\frac{\partial||\hat{\boldsymbol{z}}_k^*||_1}{\partial \hat{\boldsymbol{z}}_k^*}\left(\frac{\partial||\hat{\boldsymbol{z}}_k^*||_1}{\partial \hat{\boldsymbol{z}}_k^*}\right)^T\right] \leq \boldsymbol{I}$ and Eq. (4.3) is satisfied, a covariance matrix of an estimation error of outliers is given by

$$\begin{aligned}\mathbb{E}\left[(\hat{\boldsymbol{z}}_k^* - \boldsymbol{z}_k)(\hat{\boldsymbol{z}}_k^* - \boldsymbol{z}_k)^T\right] &\leq \left(\sqrt{\boldsymbol{W}}\right)^{-T}\left(\sqrt{\boldsymbol{W}}\right)^{-1} + \boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{R} \\ &= \boldsymbol{W}^{-1} + \boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{R} \\ &= 2\left(\boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{R}\right).\end{aligned}$$

Moreover, for an innovation of RKF, Eq. (4.23) yields

$$\boldsymbol{e}_k - \hat{\boldsymbol{z}}_k^* = \left(\sqrt{\boldsymbol{W}}\right)^{-T} \frac{\partial||\hat{\boldsymbol{z}}_k^*||}{\hat{\boldsymbol{z}}_k^*}.$$

Therefore, its covariance matrix is given by

$$\begin{aligned}\mathbb{E}\left[(\boldsymbol{e}_k - \hat{\boldsymbol{z}}_k^*)(\boldsymbol{e}_k - \hat{\boldsymbol{z}}_k^*)^T\right] &\leq \left(\sqrt{\boldsymbol{W}}\right)^{-T}\left(\sqrt{\boldsymbol{W}}\right)^{-1} \\ &= \boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{R}.\end{aligned}$$

■

Remark 4.2 Note that Eq. (4.22) is an approximated solution, not optimal, again. However, a covariance matrix of an estimation error of $\hat{\boldsymbol{z}}_k^*$ is bounded by Eq. (4.20) under the condition (4.19). □

Remark 4.3 The covariance matrix of an innovation of RKF is bounded by that of standard KF without outliers. The fact shows that performances of RKF computed by the proposed algorithm become ideal ones in some meanings. □

Remark 4.4 The covariance matrices are satisfied only under the condition (4.19). In other words, the condition (4.19) can judge whether the proposed algorithm is good or not. □

Remark 4.5 Theorem 3.1 shows that performances of RKF are given by Eq. (4.20) and (4.21) only if $\boldsymbol{W}$ is a diagonal matrix. The condition of Theorem 4.1 includes such a situation. Conversely, even if $\boldsymbol{W}$ has non-zero cross terms, performances of RKF are given by Eq. (4.20) and (4.21) under the condition (4.19). □

4.3.3 Analysis of a State Estimation Error

Eq. (2.6) and (4.1) yield the following equation:

$$\begin{aligned} \boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k} &= \boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1} - \boldsymbol{L}(\boldsymbol{e}_k - \hat{\boldsymbol{z}}_k^*) \\ &= (\boldsymbol{I} - \boldsymbol{LC})(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}) - \boldsymbol{L}\boldsymbol{v}_k - \boldsymbol{L}\left(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k^*\right), \end{aligned}$$

where $\hat{\boldsymbol{z}}_k^*$ is used in Eq. (4.1) instead of $\boldsymbol{z}_k^*$. Both in KF and RKF, Eq. (4.1) updates a covariance matrix of a state estimation error. However, under outliers, an actual updated covariance matrix of a state estimation error is given by

$$\begin{aligned} \boldsymbol{P}_{k|k} &= \mathbb{E}\left[(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k})(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k})^T\right] \\ &= (\boldsymbol{I} - \boldsymbol{LC})\boldsymbol{P}_{k|k-1}(\boldsymbol{I} - \boldsymbol{LC})^T + \boldsymbol{LRL}^T + \boldsymbol{L}\mathbb{E}\left[(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k^*)(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k^*)^T\right]\boldsymbol{L}^T \\ &= (\boldsymbol{I} - \boldsymbol{LC})\boldsymbol{P}_{k|k-1} + \boldsymbol{L}\mathbb{E}\left[(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k^*)(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k^*)^T\right]\boldsymbol{L}^T. \end{aligned} \tag{4.24}$$

$\hat{\boldsymbol{z}}_k^* = \boldsymbol{0}$ in the standard KF, so an actual $\boldsymbol{P}_{k|k}$ under outliers depends on a second moment of $\boldsymbol{z}_k$. For example, if $\boldsymbol{z}_k$ is distributed by a distribution whose second moment is infinite, like a Cauchy distribution[26], the updated covariance matrix $\boldsymbol{P}_{k|k}$ should be infinite in ideal, and it results in no update of a state. On the other hand, in RKF using the proposed algorithm, Eq. (4.24) satisfies the following inequality and bounded.

$$\begin{aligned} \boldsymbol{P}_{k|k} &\leq (\boldsymbol{I} - \boldsymbol{LC})\boldsymbol{P}_{k|k-1} + 2\boldsymbol{LCP}_{k|k-1} \\ &= (\boldsymbol{I} + \boldsymbol{LC})\boldsymbol{P}_{k|k-1}. \end{aligned} \tag{4.25}$$

In RKF using the proposed algorithm, the updated covariance matrix of a state estimation error can be selected among solutions satisfying Eq. (4.25). The update law (4.1) is one of the solutions.

4.4 Simulation

4.4.1 Problem Settings

In this section, we consider a state estimation problem of two-wheel vehicle under outliers shown in the section 3.6. A model used in the estimation is constant acceleration model (3.20).

Examples of this situation are UGV with GPS and UAV using a visual feedback [7, 8]. In feedback systems using non-contact sensors, outliers are often occurred because of temporary change of image contrast in background and radio disturbances due to some obstacles. Moreover, a coordinate and dynamics are different from the situation, but a clutter which is one of outliers is happened in a target tracking by using a radar measurement [6].

4.4.2 Noise Model

Two cases of distributions are considered as outliers, i.e., Cauchy and Gaussian mixture distributions, p_c and p_g. Cauchy distribution has heavier tails more than Gaussian distribution and is often used to represent impulsive unexpected values of sensors [26]. Gaussian mixture distribution is also used to express unusual outliers, e.g., clutter of target tracking systems [6].

4.4.3 Conditions

An initial value of the vehicle is $\boldsymbol{x}_0 = [2.5 \quad 1.0 \quad 0.0]^T$, and parameters of the vehicle are $m = 1.0$ and $J = 0.1$.

A nominal measurement noise is Gaussian white noise whose mean is $\mathbf{0}$. Consider two casess of covariance matrices of the nominal measurement noise. In the case I, the covariance matrix has small non-zero cross terms. On the other hand, cross terms of the covariance matrix are about as large as its diagonal elements in the case II. Concretely, the covarinace matrices in the case I and II are given by the following equations, respectively:

$$\text{Case I:} \qquad \boldsymbol{R} = \begin{bmatrix} 0.25 & 0.16 & 0.01 \\ 0.16 & 0.25 & 0.09 \\ 0.01 & 0.09 & 0.25 \end{bmatrix}.$$

$$\text{Case II:} \qquad \boldsymbol{R} = \begin{bmatrix} 0.29 & 0.30 & 0.36 \\ 0.30 & 0.53 & 0.30 \\ 0.36 & 0.30 & 0.49 \end{bmatrix}.$$

In the use of the Cauchy distribution, Cauchy noise is added to the nominal measurement noise. Its parameters are $x_0 = 0$ and $\delta = 5 \times 10^{-2}$. In the use of the Gaussian mixture distribution, we set $p = 0.3$. Additionally, $N_x(0, \boldsymbol{\Sigma}_1^2)$ is a distribution of the nominal measurement noise, and $N_x(0, \boldsymbol{\Sigma}_2^2)$ is a distribution whose STD is 20 times more than that of the nominal measurement noise.

Parameters of RKF are $\boldsymbol{P}_{0|0} = \boldsymbol{I}$, $\sigma_{a_x}^2 = \sigma_{a_y}^2 = 1.0 \times 10^4$, and $\sigma_{a_\theta}^2 = 5.0 \times 10$. A covariance matrix of the nominal measurement noise assumes to be known.

MATLAB is used to compute the simulation. A CPU of a computer used in the simulation is Xeon X5550 (2.66GHz) and memory is 3GB. The proposed algorithm is compared with

Table 4.1: Sum of root mean squared errors of each estimate in the case I.

Type of methods	Cauchy noise	Gaussian mixture noise
KF without outliers	1.41	
KF with outliers	82.2	13.3
RKF using CVX	1.46	2.24
RKF using CVXGEN with fixed regularization parameter	1.55	2.42
RKF using FISTA with fixed regularization parameter	1.47	2.77
RKF using only diagonal elements	1.43	1.78
proposed method	1.41	1.81
proposed method (Compiled ver.)	1.42	1.80

Table 4.2: Sum of root mean squared errors of each estimate in the case II.

Type of methods	Cauchy noise	Gaussian mixture noise
KF without outliers	1.31	
KF with outliers	50.1	12.2
RKF using CVX	1.66	4.35
RKF using CVXGEN with fixed regularization parameter	1.54	4.94
RKF using FISTA with fixed regularization parameter	2.18	4.35
RKF using only diagonal elements	2.04	2.17
proposed method	1.52	1.80
proposed method (Compiled ver.)	1.52	1.78

four methods, i.e., CVX, CVXGEN, FISTA and RKF using only diagonal elements of $\boldsymbol{W}$. Since RKF can compute its solution analytically if $\boldsymbol{W}$ is a diagonal matrix, RKF using only the diagonal elements results in a fast computation. CVXGEN generates C code of QP optimization problems in CVX, and the compiled code can be used in MATLAB to accelerate a computation. A solution of Eq. (4.2) requires a computation of LMI (4.5). However, CVXGEN cannot deal with SDP like LMI. FISTA also cannot solve LMI. One regularization parameter calculated by CVX is fixed in CVXGEN and FISTA. Furthermore, an estimation procedure of outlier in the proposed algorithm is implemented in C code, and the compiled code is also compared with them.

4.4.4 Results

Table 4.1 and 4.2 show summations of root mean square errors (RMSEs) of each state in the case I and II, respectively. Table 4.3 shows averaged computation times of each algorithm

Table 4.3: Average of computation time at one time step.

Type of methods	Computation time [ms]
KF	0.06
RKF using CVX	239
RKF using CVXGEN with fixed regularization parameter	0.11
RKF using FISTA with fixed regularization parameter	0.30
RKF using only diagonal elements	0.09
proposed method	0.10
proposed method (Compiled ver.)	0.08

at one time step. These values are averages of 10 times simulations. Results of standard KF with and without outliers are also shown in the table for comparison. In the simulation, results using the proposed algorithm satisfy the condition (4.19).

Table 4.1 shows that, under Cauchy noise, RMSEs of RKF using any algorithms are close to that of KF without outliers. It also shows that RKF can reduce effects of Gaussian mixture noise. However, accuracy of the solutions depends on the algorithms, and iterative algorithms give larger RMSEs than the proposed algorithm and RKF using only diagonal elemnets. In case I, RKF using only diagonal elements has same performances as the proposed algorithm because the cross terms of the covariance matrix $\boldsymbol{R}$ are small.

Table 4.2 also shows that RKF can reduce effects of both Cauchy and Gaussian mixture noises. In addition, it can be seen that the proposed algorithm gives smaller RMSE than the other algorithms including RKF using only diagonal elements. If the covariance matrix $\boldsymbol{R}$ has large cross terms, RKF using only diagonal elements deteriorates its performances. However, the proposed algorithm provides good performances also under such a situation.

Table 4.3 shows that, using CVXGEN, FISTA, RKF using only diagonal elements, and the proposed algorithm, computation times are about 1/1000 times less than one using CVX. Moreover, the compiled version of the proposed algorithm is more accelerated, and a computation time of the compiled version comes close to that of KF.

4.5 Conclusion

In this chapter, we proposed a fast algorithm of RKF via l_1 regression, which consists of a l_1 optimization problem. The proposed algorithm approximates the optimal solution by using its upper and lower bounds, and the approximated solution is given by a closed form. Moreover, it was shown that the proposed algorithm has almost same performances as KF without outliers. Effectiveness was demonstrated by some numerical simulations.

In larger scale optimization problems, or in some conditions of covariance matrix of Gaussian noise, the condition (4.19) was not satisfied at times, then performances of RKF were sometimes deteriorated. A proposition of efficient algorithms for the large scale problems is one of our future works.

Chapter 5

Robust Self-Tuning Controller under Outliers

5.1 Introduction

Self-tuning controller (STC) has been studied as one of control strategies for systems with unknown parameters and varying systems [42]–[47]. And also, STC has been applied to various industrial applications. Recently, an application of STC to an engine control was reported [47], and STC still has many attentions not only in theory, but also in practice.

In a field of studies for STC theory, since Kalman [42] proposed a method combining a least squares (LS) estimation and feedback control, many researchers have proposed various types of STC and discussed about their stabilities. For example, in [43], a method to non-minimum phase systems was proposed. However, stability of the method was proven only for the case of known parameters, and stability using a recursive LS (RLS) to estimate the parameters on-line could not be proven. In [44], global stability of MIMO systems for the noise-free case was proven. Many other methods were proposed, e.g., an estimation method of parameters of PID in on-line [45], but a proof of global stability for general MIMO systems had been a difficult problem for a long time. However, recently, in [46, 47], a global stability for MIMO systems was proven using an idea of sliding mode control.

In this manner, many researchers have studied and improved STC until now. However, these methods consider no measurement noise or only measurement Gaussian noise as measurement noise, and these methods cannot deal with non-Gaussian measurement noise like outliers. Outliers are often happened in many applications. For example, non-contact sensors, e.g., radar measurement, GPS, image measurements, and so on, attract attentions. However, it is known that external environments introduce outliers, which cannot be represented as a normal distribution, into these sensor signals [6, 7]. A use of these sensors in feedback control systems deteriorates control performances, so some methods to reduce the effect of the outliers are required.

In this chapter, we extend a method in [47] and propose a robust STC (RSTC) under outliers. A parameter update law of the conventional STC is given by a RLS estimation. This means that the parameter update law is equivalent to a solution of a minimization problem which consists of a quadratic form of estimated errors. We apply an idea of robust Kalman filter (RKF) via l_1 regression [20] to the minimization problem, and we estimate outliers explicitly by adding l_1 regression term to the minimization problem. We construct

RSTC under outliers by removing the estimated outliers from a controller. Moreover, we analyze control performances of the proposed RSTC.

The organization of this chapter is as follows. In section 2, a conventional STC is explained. In section 3, we propose a RSTC under outliers. And also, we analyze control performances of the proposed method, and it is shown that steady state errors in the proposed method with outliers are nearly equal to ones in the conventional STC without outliers. In section 4, we demonstrate its effectiveness by some numerical simulations. Conclusion is given in section 5.

5.2 Self Tuning Controller

5.2.1 Formula

In this section, STC in [46, 47] is described.

Consider the following p-input p-output linear polynomial systems:

$$\boldsymbol{A}(z^{-1})\boldsymbol{y}_k^* = \boldsymbol{B}(z^{-1})z^{-d}\boldsymbol{u}_k, \tag{5.1}$$

where $\boldsymbol{u}_k \in \mathbb{R}^p$ is an input, $\boldsymbol{y}_k^* \in \mathbb{R}^p$ is an output, and $d \in \mathbb{R}$ is a known delay index. Since an output with outliers are treated in this chapter, an output without outliers is expressed as $\boldsymbol{y}_k^*$ to distinguish the output with outliers. $\boldsymbol{A}(z^{-1})$ and $\boldsymbol{B}(z^{-1})$ are the following $p \times p$ square matrix polynomials:

$$\boldsymbol{A}(z^{-1}) = \boldsymbol{I} + \boldsymbol{A}_1 z^{-1} + \cdots + \boldsymbol{A}_n z^{-n},$$
$$\boldsymbol{B}(z^{-1}) = \boldsymbol{B}_0 + \boldsymbol{B}_1 z^{-1} + \cdots + \boldsymbol{B}_m z^{-m},$$

where $\boldsymbol{B}_0$ is assumed to be non-singular.

Consider the following constraint $s_{k,i}^*$ $(i = 1, \cdots, p)$:

$$\boldsymbol{s}_{k+d}^* := \boldsymbol{C}(z^{-1})(\boldsymbol{y}_{k+d}^* - \boldsymbol{r}_{k+d}) + \boldsymbol{Q}(z^{-1})\boldsymbol{u}_k, \tag{5.2}$$

where $\boldsymbol{r}_k \in \mathbb{R}^p$ is a reference signal, $\boldsymbol{C}(z^{-1})$ and $\boldsymbol{Q}(z^{-1})$ are given by the following diagonal matrices:

$$\boldsymbol{C}(z^{-1}) = \begin{bmatrix} C_{11}(z^{-1}) & & 0 \\ & \ddots & \\ 0 & & C_{pp}(z^{-1}) \end{bmatrix}, \tag{5.3}$$

$$\boldsymbol{Q}(z^{-1}) = \begin{bmatrix} Q_{11}(z^{-1}) & & 0 \\ & \ddots & \\ 0 & & Q_{pp}(z^{-1}) \end{bmatrix}, \tag{5.4}$$

where an element of Eq. (5.3) is $C_{ii}(z^{-1}) = 1 + c_{ii}^1 z^{-1} + \cdots + c_{ii}^{n-1} z^{n-1}$ and designed to be a Schur polynomial. Eq. (5.4) is chosen to be $Q_{ii}(z^{-1}) = q_{ii}(1 - z^{-1})$. $\boldsymbol{G}(z^{-1})$ is defined as the following equation:

$$\boldsymbol{G}(z^{-1}) = \boldsymbol{E}(z^{-1})\boldsymbol{B}(z^{-1}) + \boldsymbol{Q}(z^{-1}),$$

where $\boldsymbol{E}(z^{-1})$ is $p \times p$ symmetric square matrix polynomial and satisfies the following Diophantine equation:

$$\boldsymbol{C}(z^{-1}) = \boldsymbol{E}(z^{-1})\boldsymbol{A}(z^{-1}) + z^{-d}\boldsymbol{F}(z^{-1}),$$

where $\boldsymbol{E}(z^{-1})$ and $\boldsymbol{F}(z^{-1})$ are given by the following equations:

$$\boldsymbol{E}(z^{-1}) = \boldsymbol{I} + \boldsymbol{E}^1 z^{-1} + \cdots + \boldsymbol{E}^{d-1} z^{-(d-1)},$$

$$\boldsymbol{F}(z^{-1}) = \boldsymbol{F}^0 + \boldsymbol{F}^1 z^{-1} + \cdots + \boldsymbol{F}^{n-1} z^{-(n-1)}.$$

At that time, the constraint (5.2) can be rewritten as the following equation:

$$\boldsymbol{s}^*_{k+d} = \boldsymbol{G}(z^{-1})\boldsymbol{u}_k + \boldsymbol{F}(z^{-1})\boldsymbol{y}^*_k - \boldsymbol{C}(z^{-1})\boldsymbol{r}_{k+d}. \tag{5.5}$$

Moreover, the following theorem is satisfied:

Theorem 5.1 Consider p-input p-output linear polynomial systems described as Eq. (5.1). For design parameters of the constraint (5.2), it is assumed that Eq. (5.3) is designed to satisfy a Schur polynomial and Eq. (5.4) is chosen to be $Q_{ii}(z^{-1}) = q_{ii}(1 - z^{-1})$. Then, a control law (5.6) results in $\boldsymbol{s}^*_k \to \boldsymbol{0}$ $(k \to \infty)$.

$$\boldsymbol{u}_k = \hat{\boldsymbol{G}}_k^{-1}(z^{-1}) \left[\boldsymbol{C}(z^{-1})\boldsymbol{r}_{k+d} - \hat{\boldsymbol{F}}_k(z^{-1})\boldsymbol{y}^*_k\right], \tag{5.6}$$

where $\hat{\boldsymbol{F}}_k(z^{-1})$ and $\hat{\boldsymbol{G}}_k(z^{-1})$ are estimated polynomials of $\boldsymbol{F}(z^{-1})$ and $\boldsymbol{G}(z^{-1})$ at time k, respectively. (i,j)-th elements of the polynomials are given by the following polynomials:

$$\hat{F}_{k,ij}(z^{-1}) = \hat{f}^0_{k,ij} + \hat{f}^1_{k,ij} z^{-1} + \cdots + \hat{f}^{n-1}_{k,ij} z^{-(n-1)},$$

$$\hat{G}_{k,ij}(z^{-1}) = \hat{g}^0_{k,ij} + \hat{g}^1_{k,ij} z^{-1} + \cdots + \hat{g}^{m+d-1}_{k,ij} z^{-(m+d-1)}.$$

A coefficient vector consisting of i-th row of the estimated polynomial matrix is given by

$$\begin{aligned}\hat{\boldsymbol{\theta}}^i_k = \big[\ & \hat{g}^0_{k,i1} \ \cdots \ \hat{g}^{m+d-1}_{k,i1} \ \cdots \ \hat{g}^0_{k,ip} \ \cdots \ g^{m+d-1}_{k,ip} \\ & \hat{f}^0_{k,i1} \ \cdots \ \hat{f}^{n-1}_{k,i1} \ \cdots \ \hat{f}^0_{k,ip} \ \cdots \ f^{n-1}_{k,ip}\ \big]^T,\end{aligned}$$

which is estimated as follows:

$$\hat{\boldsymbol{\theta}}^i_k = \hat{\boldsymbol{\theta}}^i_{k-1} + \boldsymbol{K}^i_k e^*_{k,i}, \tag{5.7}$$

$$e^*_{k,i} = s^*_{k,i} + C_{ii}(z^{-1})r_{k,i} - (\boldsymbol{\phi}^*_{k-d})^T \hat{\boldsymbol{\theta}}^i_{k-1}, \tag{5.8}$$

$$\boldsymbol{K}^i_k = \frac{\boldsymbol{P}^i_{k-1}\boldsymbol{\phi}^*_{k-d}}{(\boldsymbol{\phi}^*_{k-d})^T \boldsymbol{P}^i_{k-1}\boldsymbol{\phi}^*_{k-d} + R^s_i}, \tag{5.9}$$

$$\boldsymbol{P}^i_k = \left(\boldsymbol{I} - \boldsymbol{K}^i_k (\boldsymbol{\phi}^*_{k-d})^T\right) \boldsymbol{P}^i_{k-1}, \tag{5.10}$$

$$\begin{aligned}\boldsymbol{\phi}^*_k = \big[\ & u_{k,1} \ \cdots \ u_{k-(m+d-1),1} \ \cdots \ u_{k,p} \ \cdots \ u_{k-(m+d-1),p} \\ & y^*_{k,1} \ \cdots \ y^*_{k-(n-1),1} \ \cdots \ y^*_{k,p} \ \cdots \ y^*_{k-(n-1),p}\ \big]^T,\end{aligned} \tag{5.11}$$

where $R^s_i \in \mathbb{R},\ i = 1, \cdots, p$, is a weight of the error $e^*_{k,i}$. □

Proof: See [47]. ■

In [47], Eq. (5.9) is given by a RLS estimation without R_i^s. However, use of the following Lyapunov function candidate can prove Eq. (5.9):

$$V_{k,i} = \frac{1}{2}(e_{k,i}^*)^T (R_i^s)^{-1} e_{k,i}^* + \frac{1}{2}(\boldsymbol{\theta}_k^i - \hat{\boldsymbol{\theta}}_k^i)^T (\boldsymbol{P}_k^i)^{-1} (\boldsymbol{\theta}_k^i - \hat{\boldsymbol{\theta}}_k^i).$$

Remark 5.1 Since design parameters (5.3) and (5.4) in the constraint (5.2) are diagonal matrices, Eq. (5.7)–(5.11) are satisfied independently for each element of $\boldsymbol{s}_k^*$. This means that the problem in the theorem is equivalent to solving a single output system whose output is $s_{k,i}^* + C_{ii}(z^{-1})r_{k,i}$. □

5.2.2 Relationship Between a Parameter Update Law of Self Tuning Controller and Optimization Problem

Consider the following system which has measurement Gaussian noise $\boldsymbol{v}_k \in \mathbb{R}^p$ explicitly (see Fig. 5.1):

$$\boldsymbol{A}(z^{-1})\boldsymbol{y}_k^* = \boldsymbol{B}(z^{-1})z^{-d}\boldsymbol{u}_k, \tag{5.12}$$

$$\boldsymbol{y}_k^* = \bar{\boldsymbol{y}}_k + \boldsymbol{v}_k, \tag{5.13}$$

where $\bar{\boldsymbol{y}}_k$ is a nominal output without measurement noise.

Consider the following optimization problem for Eq. (5.12) and (5.13):

$$\begin{aligned} &\hat{\boldsymbol{\theta}}_k^i = \arg\min_{\boldsymbol{\theta}_k^i} {v_{k,i}^s}^T (R_i^s)^{-1} v_{k,i}^s + (\boldsymbol{\theta}_k^i - \hat{\boldsymbol{\theta}}_{k-1}^i)^T (\boldsymbol{P}_{k-1}^i)^{-1} (\boldsymbol{\theta}_k^i - \hat{\boldsymbol{\theta}}_{k-1}^i), \\ &\text{subject to } \; s_{k,i}^* + C_{ii}(z^{-1})r_{k,i} = (\boldsymbol{\phi}_{k-d}^*)^T \boldsymbol{\theta}_k^i + v_{k,i}^s, \end{aligned} \tag{5.14}$$

where $v_{k,i}^s$ is a virtual measurement Gaussian noise in the constraint associated with $v_{k,i}$, and R_i^s is a variance of $v_{k,i}^s$. A solution of the optimization problem (5.14) becomes the parameter update law (5.7). This means that the parameter update law (5.7) is equivalent to the solution of the optimization problem (5.14).

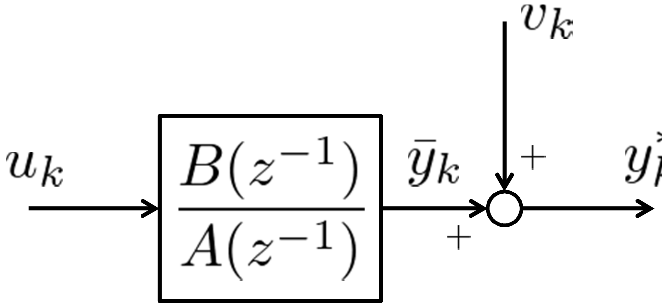


Figure 5.1: Controlled plant with Gaussian measurement noise in the case of SISO

If the regressor is not contaminated by measurement noise, i.e., the regressor is given by

$$\bar{\boldsymbol{\phi}}_k = \begin{bmatrix} u_{k,1} & \cdots & u_{k-(m+d-1),1} & \cdots & u_{k,p} & \cdots & u_{k-(m+d-1),p} \\ \bar{y}_{k,1} & \cdots & \bar{y}_{k-(n-1),1} & \cdots & \bar{y}_{k,p} & \cdots & \bar{y}_{k-(n-1),p} \end{bmatrix}^T, \quad (5.15)$$

then, from Eq. (5.2), noise $v^s_{k,i}$ in the constraint $s^*_{k,i}$ is expressed as

$$v^s_{k,i} = C_{ii}(z^{-1})v_{k,i}. \quad (5.16)$$

If measurement Gaussian noise $\boldsymbol{v}_k$ is white noise, $\boldsymbol{v}^s_k$ is also distributed by a normal distribution due to a reproductive property of the normal distribution.

5.3 Robust Self Tuning Controller

5.3.1 Parameter Update Law of Robust Self-Tuning Controller Using l_1 Regression

In this chapter, the following controlled plant is considered (see Fig. 5.2):

$$\boldsymbol{A}(z^{-1})\boldsymbol{y}_k = \boldsymbol{B}(z^{-1})z^{-d}\boldsymbol{u}_k, \quad (5.17)$$

$$\boldsymbol{y}_k = \bar{\boldsymbol{y}}_k + \boldsymbol{v}_k + \boldsymbol{z}_k = \boldsymbol{y}^*_k + \boldsymbol{z}_k, \quad (5.18)$$

where $\boldsymbol{z}_k \in \mathbb{R}^p$ is an outlier and assumed to be independent of measurement Gaussian noise $\boldsymbol{v}_k$. A control law and parameter update law of a RSTC under $z_{k,i}$ are defined as follows.

Definition 5.1 RSTC is defined as the following equation:

$$\boldsymbol{u}_k = \hat{\boldsymbol{G}}_k^{-1}(z^{-1})\left[\boldsymbol{C}(z^{-1})\boldsymbol{r}_{k+d} - \hat{\boldsymbol{F}}_k(z^{-1})(\boldsymbol{y}_k - \hat{\boldsymbol{z}}_k)\right], \quad (5.19)$$

where $\hat{\boldsymbol{z}}_k$ is an estimate of $\boldsymbol{z}_k$ and can be computed by a parameter update law. □

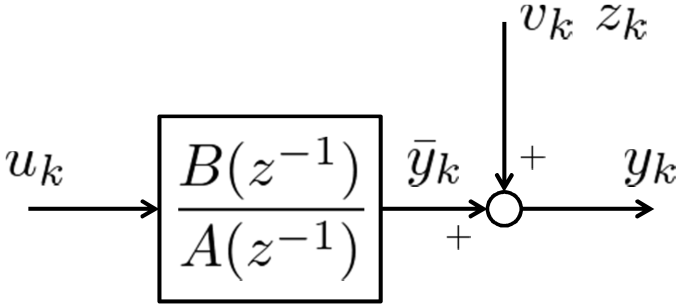


Figure 5.2: Controlled plant with Gaussian measurement noise and outlier in the case of SISO

Definition 5.2 A constraint and regressor are defined as follows:

$$\begin{aligned} \boldsymbol{s}_{k+d} &:= \boldsymbol{C}(z^{-1})(\boldsymbol{y}_{k+d} - \boldsymbol{r}_{k+d}) + \boldsymbol{Q}(z^{-1})\boldsymbol{u}_k \\ &= \boldsymbol{G}(z^{-1})\boldsymbol{u}_k + \boldsymbol{F}(z^{-1})\boldsymbol{y}_k - \boldsymbol{C}(z^{-1})\boldsymbol{r}_{k+d}, \end{aligned} \tag{5.20}$$

$$\begin{aligned} \boldsymbol{\phi}_k := [\ & u_{k,1} \ \cdots \ u_{k-(m+d-1),1} \ \cdots \ u_{k,p} \ \cdots \ u_{k-(m+d-1),p} \\ & y_{k,1} \ \cdots \ y_{k-(n-1),1} \ \cdots \ y_{k,p} \ \cdots \ y_{k-(n-1),p} \ ]^T. \end{aligned} \tag{5.21}$$

Parameters of the control law and an estimate of $z_{k,i}$ are given by

$$\{\hat{\boldsymbol{\theta}}_k^i, \hat{z}_{k,i}\} = \arg\min_{\boldsymbol{\theta}_k^i, z_{k,i}} {v_{k,i}^s}^T (R_i^s)^{-1} v_{k,i}^s (\boldsymbol{\theta}_k^i - \hat{\boldsymbol{\theta}}_{k-1}^i)^T (\boldsymbol{P}_{k-1}^i)^{-1} (\boldsymbol{\theta}_k^i - \hat{\boldsymbol{\theta}}_{k-1}^i) + \lambda_i ||z_{k,i}||_1,$$

$$\begin{aligned} \text{subject to} \quad & s_{k,i} + C_{ii}(z^{-1}) r_{k,i} = \boldsymbol{\phi}_{k-d}^T \boldsymbol{\theta}_k^i + v_{k,i}^s + z_{k,i}^s, \\ & \boldsymbol{\theta}_k^i = \hat{\boldsymbol{\theta}}_{k-1}^i + K_k^i (e_{k,i} - z_{k,i}^s), \\ & e_{k,i} = s_{k,i} + C_{ii}(z^{-1}) r_{k,i} - \boldsymbol{\phi}_{k-d}^T \hat{\boldsymbol{\theta}}_{k-1}^i, \\ & z_{k,i}^s = C_{ii}(z^{-1}) z_{k,i}, \end{aligned} \tag{5.22}$$

where λ_i is a regularization parameter. □

We give some remarks about Definition 5.1 and 5.2.

Remark 5.2 Use of the control law (5.19) means considering the following constraint:

$$\hat{\boldsymbol{s}}_{k+d} = \hat{\boldsymbol{G}}_k(z^{-1})\boldsymbol{u}_k + \hat{\boldsymbol{F}}_k(z^{-1})(\boldsymbol{y}_k - \hat{\boldsymbol{z}}_k) - \boldsymbol{C}(z^{-1})\boldsymbol{r}_{k+d}. \tag{5.23}$$

In order to reduce the effects of outliers, outliers are estimated explicitly, and the estimated outliers are subtracted from outputs. However, a constraint contaminated by outliers, i.e., Eq. (5.20), is used to estimate the outliers. □

Remark 5.3 In order to estimate outliers whose values may often become zero, Eq. (5.22) is constructed by adding a sparse regularization term, by which estimates may contain zero values, to Eq. (5.14). The idea is same as an idea of a derivation of RKF via l_1 regression [20]. Among all sparse regularization techniques, use of l_1 regression results in a convex optimization and the solution can be easily calculated. □

Remark 5.4 In Definition 5.2, $\boldsymbol{\theta}_k^i$, $\boldsymbol{v}_k$, and $\boldsymbol{z}_k$ are regarded as stochastic variables, and an expectation $\hat{\boldsymbol{\theta}}_k^i = \mathbb{E}[\boldsymbol{\theta}_k^i]$ and $\hat{z}_{k,i}$ which is an estimate of $z_{k,i}$ at a time k are computed. Note that $\hat{z}_{k,i}$ is not an expectation of $z_{k,i}$. □

Remark 5.5 An outlier in the constraint, i.e., $z_{k,i}^s$, is written as

$$z_{k,i}^s = C_{ii}(z^{-1}) z_{k,i}, \tag{5.24}$$

only if a regressor is given by Eq. (5.15). However, for simplicity, $z_{k,i}^s$ is assumed to be given by Eq. (5.24) in Eq. (5.22).

Moreover, the relationship between outliers in the constraint and measurement outputs, i.e., the relationship between $\boldsymbol{z}_k^s$ and $\boldsymbol{z}_k$, is one-to-one at each element. Each element of noise in the constraint is decoupled like Remark 5.1, so Eq. (5.22) deals with the elements of $\boldsymbol{z}_k$ independently. □

Lemma 5.1 Eq. (5.22) can be rewritten as

$$\hat{z}_{k,i} = \arg\min_{z_{k,i}} \xi_i^T \ W_i \ \xi_i + \lambda_i ||z_{k,i}||_1, \tag{5.25}$$

$$\hat{\boldsymbol{\theta}}_k^i = \hat{\boldsymbol{\theta}}_{k-1}^i + \boldsymbol{K}_k^i \left(e_{k,i} - \hat{z}_{k,i}^s\right), \tag{5.26}$$

where $\xi_i := e_{k,i} - C_{ii}(z^{-1})z_{k,i}$,
and $W_i := (1 - \boldsymbol{\phi}_{k-d}^T \boldsymbol{K}_k^i)^T (R_i^s)^{-1} (1 - \boldsymbol{\phi}_{k-d}^T \boldsymbol{K}_k^i) + (\boldsymbol{K}_k^i)^T (\boldsymbol{P}_k^i)^{-1} \boldsymbol{K}_k^i \in \mathbb{R}$. □

Eq. (5.25) is solved in a closed form due to a l_1 optimization problem with a single variable (see Appendix C).

In order to derive a theorem for RSTC in Definition 5.1 and 5.2, the following assumption and lemma are given.

Assumption 5.1 Assume that Eq. (5.3) is designed to be a sufficient condition of Schur stable [48]. Namely, either

$$1 > c_{ii}^1 > \cdots > c_{ii}^{n-1} > 0,$$

or

$$1 > |c_{ii}^1| + \cdots + |c_{ii}^{n-1}|,$$

are satisfied. □

Lemma 5.2 Consider p-input p-output linear polynomial systems described as Eq. (5.17) and (5.18). For design parameters of the constraint (5.23), it is assumed that Eq. (5.3) is designed to satisfy a Schur polynomial and Eq. (5.4) is chosen to be $Q_{ii}(z^{-1}) = q_{ii}(1 - z^{-1})$. Assume that $\boldsymbol{z}_k$ is white noise, and stochastic variables $\boldsymbol{\theta}_k^i$ and $\boldsymbol{v}_k$ are mutually independent. In addition, it is assumed that regularization parameters are given by

$$\sigma_i^2 = \boldsymbol{\phi}_{k-d}^T \boldsymbol{P}_{k-1}^i \boldsymbol{\phi}_{k-d} + R_i^s, \tag{5.27}$$

$$\lambda_i = 2W_i \ \sigma_i. \tag{5.28}$$

If Assumption 5.1 is satisfied, a variance of estimated outliers, i.e., $\mathbb{E}[(z_{k,i} - \hat{z}_{k,i})^2]$, is bounded. Moreover, a limit superior of the variance, i.e., $\limsup_{k\to\infty} \mathbb{E}[(z_{k,i} - \hat{z}_{k,i})^2]$, is bounded by the following inequality:

$$\limsup_{k\to\infty} \mathbb{E}\left[(z_{k,i} - \hat{z}_{k,i})^2\right] \leq \alpha_i \left(\boldsymbol{\phi}_\infty^T \boldsymbol{P}_\infty^i \boldsymbol{\phi}_\infty + R_i^s\right), \tag{5.29}$$

where $\alpha_i = \frac{2}{\sum_{j=0}^{n-1} (c_{\infty,ii}^j)^2}$ and $c_{\infty,ii}^0 = 1$. □

Proof: Define $\eta_i :\in [-1, \ 1]$. The necessary condition of an optimality of Eq. (5.25) can be written as

$$\forall z_{k,i}, \ \ -2W_i(e_{k,i} - \hat{z}_{k,i}^s) + \lambda_i \eta_i = 0. \tag{5.30}$$

Since η_i can be chosen randomly in $[-1,\ 1]$, η_i can be regarded as a stochastic variable without loss of generality. Moreover, η_i is independent of the other stochastic variables because η_i can be chosen independently. From Eq. (5.28), the following equation is satisfied:

$$e_{k,i} - \hat{z}^s_{k,i} = \sigma_i\ \eta_i. \tag{5.31}$$

Therefore, from Eq. (5.22),

$$\begin{aligned} z^s_{k,i} - \hat{z}^s_{k,i} &= e_{k,i} - \boldsymbol{\phi}^T_{k-d}\left(\boldsymbol{\theta}^i_k - \hat{\boldsymbol{\theta}}^i_{k-1}\right) - v^s_{k,i} - \hat{z}^s_{k,i}, \\ &= \sigma_i\ \eta_i - \boldsymbol{\phi}^T_{k-d}\left(\boldsymbol{\theta}^i_k - \hat{\boldsymbol{\theta}}^i_{k-1}\right) - v^s_{k,i}. \end{aligned} \tag{5.32}$$

Note that each stochastic variable is mutually independent. From Eq. (5.27) and $\eta_i^2 \le 1$, a variance of estimates of $z^s_{k,i}$ is bounded as

$$\begin{aligned} \mathbb{E}\left[(z^s_{k,i} - \hat{z}^s_{k,i})^2\right] &\le \boldsymbol{\phi}^T_{k-d}\boldsymbol{P}^i_{k-1}\boldsymbol{\phi}_{k-d} + R^s_i + \sigma_i^2 \\ &= 2\left(\boldsymbol{\phi}^T_{k-d}\boldsymbol{P}^i_{k-1}\boldsymbol{\phi}_{k-d} + R^s_i\right). \end{aligned} \tag{5.33}$$

Therefore, $\mathbb{E}[(z^s_{k,i} - \hat{z}^s_{k,i})^2]$ is bounded. From the assumption, $z_{k,i} - \hat{z}_{k,i}$ has a whiteness, then Eq. (5.24) gives $\mathbb{E}[(z^s_{k,i} - \hat{z}^s_{k,i})^2] = \sum_{j=0}^{n-1}(c^j_{ii})^2\ \mathbb{E}[(z_{k-j,i} - \hat{z}_{k-j,i})^2]$. From Assumption 5.1, $\mathbb{E}[(z_{k,i} - \hat{z}_{k,i})^2]$ is also bounded because $(c^j_{ii})^2$ is also a coefficient of Schur polynomials. At $k \to \infty$,

$$\limsup_{k\to\infty} \mathbb{E}[(z^s_{k,i} - \hat{z}^s_{k,i})^2] = \limsup_{k\to\infty} \mathbb{E}[(z_{k,i} - \hat{z}_{k,i})^2]\ \sum_{j=0}^{n-1}(c^j_{\infty,ii})^2.$$

Therefore, from Eq. (5.33), Eq. (5.29) is satisfied. ■

Theorem 5.2 Consider p-input p-output linear polynomial systems described as Eq. (5.17) and (5.18). For design parameters of the constraint (5.23), it is assumed that Eq. (5.3) is designed to satisfy a Schur polynomial and Eq. (5.4) is chosen to be $Q_{ii}(z^{-1}) = q_{ii}(1 - z^{-1})$. Assuming that $\boldsymbol{z}_k$ is white noise and its each element is mutually independent. Stochastic variables, i.e., $\boldsymbol{\theta}^i_k$, $\boldsymbol{v}_k$, and $\boldsymbol{z}_k$, are assumed to be mutually independent. Moreover, it is assumed that regularization parameters are denoted as Eq. (5.28). If a covariance matrix $\boldsymbol{P}^i_k$ $(k \to \infty)$ is sufficient small and Assumption 5.1 is satisfied, steady state errors of RSTC with outliers and normal STC without outliers, i.e., $s_{k,i}$ and $s^*_{k,i}$, satisfy

$$\limsup_{k\to\infty} \mathbb{E}\left[(s^*_{k,i} - \hat{s}_{k,i})^2\right] \le \mathcal{O}\left(\max_i R^s_i\right). \tag{5.34}$$

□

Proof: The constraint of the normal STC without outliers (5.5) can be rewritten as

$$\begin{aligned} s^*_{k+d,i} &= \boldsymbol{g}^i(z^{-1})\boldsymbol{u}_k + \boldsymbol{f}^i(z^{-1})\boldsymbol{y}^*_k - C_{ii}(z^{-1})r_{k+d,i} \\ &= (\boldsymbol{\phi}^*_k)^T\boldsymbol{\theta}^i_k - C_{ii}(z^{-1})r_{k+d,i}, \end{aligned}$$

where $\boldsymbol{f}^i(z^{-1}) = [F_{i1}(z^{-1}), \cdots, F_{ip}(z^{-1})]$ and $\boldsymbol{g}^i(z^{-1}) = [G_{i1}(z^{-1}), \cdots, G_{ip}(z^{-1})]$.

Note that $\boldsymbol{y}_k = \boldsymbol{y}_k^* + \boldsymbol{z}_k$. Similarly, from Eq. (5.23), the constraint of RSTC with outliers can be rewritten as

$$\hat{s}_{k+d,i} = (\boldsymbol{\phi}_k^*)^T \hat{\boldsymbol{\theta}}_k^i - C_{ii}(z^{-1}) r_{k+d,i} + \hat{\boldsymbol{f}}_k^i(z^{-1}) \tilde{\boldsymbol{z}}_k,$$

where $\tilde{\boldsymbol{z}}_k = \boldsymbol{z}_k - \hat{\boldsymbol{z}}_k$. Therefore, the following equation is satisfied:

$$s_{k+d,i}^* - \hat{s}_{k+d,i} = (\boldsymbol{\phi}_k^*)^T \tilde{\boldsymbol{\theta}}_k^i - \hat{\boldsymbol{f}}_k^i(z^{-1}) \tilde{\boldsymbol{z}}_k,$$

where $\tilde{\boldsymbol{\theta}}_k^i = \boldsymbol{\theta}_k^i - \hat{\boldsymbol{\theta}}_k^i$.

From the assumption, $\tilde{\boldsymbol{z}}_{k-d-i}$ is independent of other elements at each time. A limit superior of a variance of $\hat{\boldsymbol{f}}_k^i(z^{-1}) \tilde{\boldsymbol{z}}_{k-d}$ is given by the following inequality:

$$\begin{aligned} \limsup_{k\to\infty} \mathbb{E}\left[\left(\hat{\boldsymbol{f}}_k^i(z^{-1}) \tilde{\boldsymbol{z}}_k\right)^2\right] &= \sum_{r=1}^{p} \sum_{s=0}^{n-1} \left(f_{\infty,ir}^s\right)^2 \limsup_{k\to\infty} \mathbb{E}[\tilde{z}_{k,r}^2] \\ &\le \sum_{r=1}^{p} \sum_{s=0}^{n-1} \alpha_i \left(f_{\infty,ir}^s\right)^2 \left(\boldsymbol{\phi}_\infty^T \boldsymbol{P}_\infty^r \boldsymbol{\phi}_\infty + R_r^s\right). \end{aligned}$$

Therefore, the following inequality is satisfied:

$$\limsup_{k\to\infty} \mathbb{E}\left[(s_{k,i}^* - \hat{s}_{k,i})^2\right] \le (\boldsymbol{\phi}_\infty^*)^T \boldsymbol{P}_\infty^i \boldsymbol{\phi}_\infty^* \sum_{r=1}^{p} \sum_{s=0}^{n-1} \alpha_i \left(f_{\infty,ir}^s\right)^2 \left(\boldsymbol{\phi}_\infty^T \boldsymbol{P}_\infty^r \boldsymbol{\phi}_\infty + R_r^s\right). \tag{5.35}$$

If the covariance matrix $\boldsymbol{P}_\infty^i$ is sufficient small, that is $\boldsymbol{\phi}_\infty^T \boldsymbol{P}_\infty^r \boldsymbol{\phi}_\infty \ll R_r^s$, Eq. (5.35) results in Eq. (5.34). ■

We give some remarks about the steady state errors and variances of noise in constraint.

Remark 5.6 If the variance of Gaussian noise in the constraint, i.e., R_i^s, is either 0 or sufficient small, the steady state errors in the RSTC are equal to ones in the normal STC without outliers. □

Remark 5.7 Theorem 5.2 means that performances of the proposed method depend on the variance R_i^s. Therefore, R_i^s is needed to be determined accurately. Since the proposed method can estimate outlier $z_{k,i}$, we can estimate the variance using the estimated outlier numerically. Namely, R_i^s can be estimated by

$$R_{k,i}^s = \frac{1}{N-1} \sum_{j=0}^{N-1} \varepsilon_{k-j,i}^2 - \boldsymbol{\phi}_{k-d}^T \boldsymbol{P}_k^i \boldsymbol{\phi}_{k-d}, \tag{5.36}$$

where $\varepsilon_{k-j,i} = e_{k-j,i} - C_{ii}(z^{-1}) \hat{z}_{k-j,i}$, and N is a length of previous data used in the estimation. □

5.3.2 Covariance Update Law of Robust Self-Tuning Controller

In the previous subsection, we discussed the parameter update law of the proposed method under outliers. In this subsection, we discuss an update law of covariance matrix under outliers.

Let $\boldsymbol{\theta}_k^i = \boldsymbol{\theta}_{k-1}^i$, then an estimation error is given by

$$\begin{aligned}\boldsymbol{\theta}_k^i - \hat{\boldsymbol{\theta}}_k^i &= \boldsymbol{\theta}_{k-1}^i - \hat{\boldsymbol{\theta}}_{k-1}^i - \boldsymbol{K}_k^i(e_{k,i} - \hat{z}_{k,i}^s) \\ &= \left(\boldsymbol{I} - \boldsymbol{K}_k^i \boldsymbol{\phi}_{k-d}^T\right)\left(\boldsymbol{\theta}_{k-1}^i - \hat{\boldsymbol{\theta}}_{k-1}^i\right) - \boldsymbol{K}_k^i v_{k,i}^s - \boldsymbol{K}_k^i\left(z_{k,i}^s - \hat{z}_{k,i}^s\right). \end{aligned} \tag{5.37}$$

Therefore, a covariance matrix of the estimation error, $\boldsymbol{P}_k^i$, is given by

$$\begin{aligned}\boldsymbol{P}_k^i &= \mathbb{E}\left[(\boldsymbol{\theta}_k^i - \hat{\boldsymbol{\theta}}_k^i)(\boldsymbol{\theta}_k^i - \hat{\boldsymbol{\theta}}_k^i)^T\right] \\ &= (\boldsymbol{I} - \boldsymbol{K}_k^i \boldsymbol{\phi}_{k-d}^T)\boldsymbol{P}_{k-1}^i + \boldsymbol{K}_k^i\ \mathbb{E}\left[(z_{k,i}^s - \hat{z}_{k,i}^s)^2\right]\ (\boldsymbol{K}_k^i)^T. \end{aligned} \tag{5.38}$$

The normal STC ($\hat{z}_{k,i}^s = 0$) computes the covariance matrix $\boldsymbol{P}_k^i$ by Eq. (5.10). However, the actual covariance matrix $\boldsymbol{P}_k^i$ with outliers contains a second order moment of $z_{k,i}^s$. If $z_{k,i}^s$ is distributed by some probability distribution whose second moment is infinite like Cauchy distribution, $\boldsymbol{P}_k^i$ should be infinite. However, the fact is inconsistent with a parameter update. On the other hand, from Eq. (5.38) and (5.33), the covariance matrix of the RSTC satisfies the following equations:

$$\boldsymbol{P}_k^i \le \left(\boldsymbol{I} + \boldsymbol{K}_k^i \boldsymbol{\phi}_{k-d}^T\right)\boldsymbol{P}_{k-1}^i. \tag{5.39}$$

This means that the covariance matrix of the RSTC under outliers should be a solution satisfying Eq. (5.39) Moreover, in the RSTC, the update law of the covariance matrix (5.10) is one of the solutions. Therefore, in the RSTC, Eq. (5.10) can estimate a state even under outliers, and the covariance matrix may converge.

5.4 Simulation

5.4.1 Conditions

Consider the following non-minimum phase system with the unstable zeros at −5.00 and −1.11:

$$\boldsymbol{y}_k = -\boldsymbol{A}_1\boldsymbol{y}_{k-1} + \boldsymbol{B}_0\boldsymbol{u}_{k-1} + \boldsymbol{B}_1\boldsymbol{u}_{k-2},$$

where

$$\begin{aligned}\boldsymbol{A}_1 &= \begin{bmatrix} 0.4 & 0.2 \\ 0.1 & 0.3 \end{bmatrix}, \\ \boldsymbol{B}_0 &= \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.3 \end{bmatrix}, \\ \boldsymbol{B}_1 &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned}$$

Parameters of the constraint is designed as follows:

$$\begin{aligned}
\boldsymbol{C}(z^{-1}) &= \mathrm{diag}(1+0.2z^{-1},\ 1-0.1z^{-1}),\\
\boldsymbol{Q}(z^{-1}) &= \mathrm{diag}(1.3(1-z^{-1}),\ 1.7(1-z^{-1})).
\end{aligned}$$

Then, all roots of $\boldsymbol{A}(z^{-1})\boldsymbol{Q}(z^{-1}) + \boldsymbol{B}(z^{-1})\boldsymbol{C}(z^{-1})$ are located in a unit circle. And also, the Diophantine equation determines ideal polynomials of $\boldsymbol{F}(z^{-1})$ and $\boldsymbol{G}(z^{-1})$, and they are given by

$$\begin{aligned}
\boldsymbol{F}(z^{-1}) &= \boldsymbol{F}_0,\\
\boldsymbol{G}(z^{-1}) &= \boldsymbol{G}_0 + \boldsymbol{G}_1 z^{-1}.
\end{aligned}$$

where

$$\begin{aligned}
\boldsymbol{F}^0 &= \begin{bmatrix} -0.6 & -0.2 \\ -0.1 & -0.4 \end{bmatrix},\\
\boldsymbol{G}^0 &= \begin{bmatrix} 2.1 & 0.3 \\ 0.2 & 2.0 \end{bmatrix},\\
\boldsymbol{G}^1 &= \begin{bmatrix} -0.3 & 0 \\ 0 & -0.7 \end{bmatrix}.
\end{aligned}$$

Consider the two cases of distributions as outliers, i.e., Cauchy distribution and Gaussian mixture distribution, $p_c(x)$ and $p_g(x)$. Consider a Gaussian white noise whose mean is $\mathbf{0}$ and covariance matrix is $\mathrm{diag}(1\times10^{-2}, 1\times10^{-2})$ as a nominal measurement noise. In the case of Cauchy distribution as a model of outliers, outliers distributed by Cauchy distribution are added to the nominal noise. Parameters of Cauchy distribution are $x_0 = 0$ and $\delta = 1\times10^{-3}$. In the case of Gaussian mixture distribution as a model of outliers, $p = 0.1$ and a standard deviation of Gaussian mixture distribution is 5 times as large as one of the nominal noise.

Initial parameters of RSTC are $R^s_{0,1} = R^s_{0,2} = 1$ and $\boldsymbol{P}^1_0 = \boldsymbol{P}^2_0 = I$. A length of past data to estimate the variance in the constraint is $N = 20$. Initial parameters of $\hat{\boldsymbol{F}}^0_k$, $\hat{\boldsymbol{G}}^0_k$, and $\hat{\boldsymbol{G}}^1_k$ are set as

$$\begin{aligned}
\hat{\boldsymbol{F}}^0_0 &= \begin{bmatrix} -0.5 & -0.1 \\ 0 & -0.8 \end{bmatrix},\\
\hat{\boldsymbol{G}}^0_0 &= \begin{bmatrix} 2.0 & -0.1 \\ 0.2 & 3.0 \end{bmatrix},\\
\hat{\boldsymbol{G}}^1_0 &= \begin{bmatrix} -0.1 & -0.2 \\ 0.1 & -1.0 \end{bmatrix}.
\end{aligned}$$

As the variance in the constraint of the normal STC, two cases are considered. One case is that the normal STC uses a fixed variance estimated in RSTC as the variance in the constraint. The other case is that the normal STC updates the variance by using Eq. (5.36). Other parameters of the normal STC are same as ones of RSTC.

Table 5.1: Root mean squared errors of tracking errors the controlled plant under Cauchy noise in the case using sinusoidal wave as a target.

	output 1	output 2
self tuning controller with a fixed variance	9.0×10^{-3}	1.4×10^{-2}
self tuning controller with an updated variance	6.2×10^{-3}	5.1×10^{-3}
robust self tuning controller (proposed method)	5.9×10^{-3}	5.1×10^{-3}

Table 5.2: Root mean squared errors of tracking errors of the controlled plant under Gaussian mixture noise in the case using a sinusoidal wave as a target.

	output 1	output 2
self tuning controller with a fixed variance	2.0×10^{-2}	2.4×10^{-2}
self tuning controller with an updated variance	7.3×10^{-3}	9.5×10^{-3}
robust self tuning controller (proposed method)	3.6×10^{-3}	6.2×10^{-3}

5.4.2 Results

In the Case Using a Sinusoidal Wave as a Target

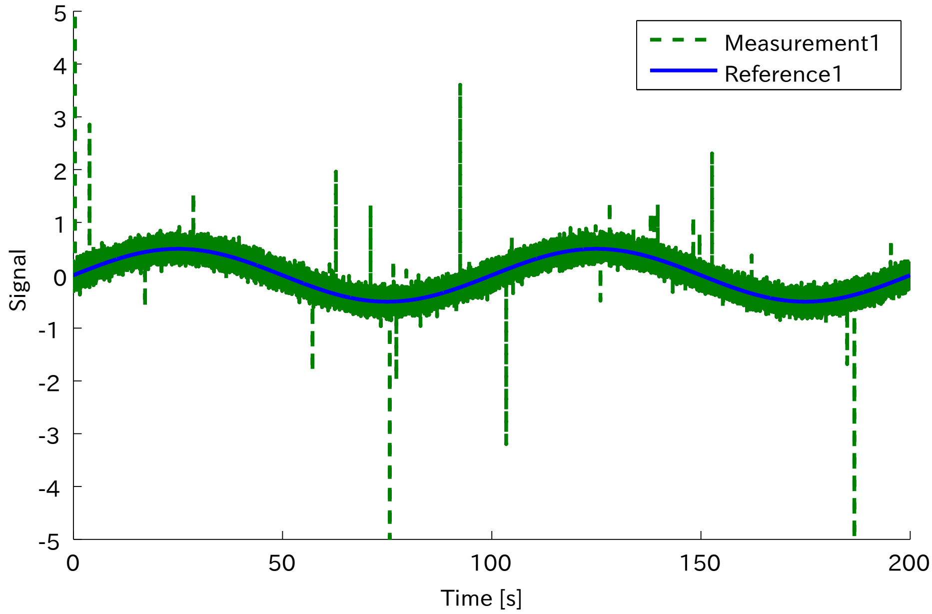
Fig. 5.3 – 5.16 show simulation results in the case using a sinusoidal wave as a target.

Fig. 5.3 – 5.9 show results in the case of Cauchy distribution. Fig. 5.3 shows measurement outputs with outliers. Fig. 5.4 – 5.6 show outputs of STC with a fixed variance, STC with an updated variance, and RSTC, respectively. Fig. 5.7 – 5.9 show estimates of parameters of these methods.

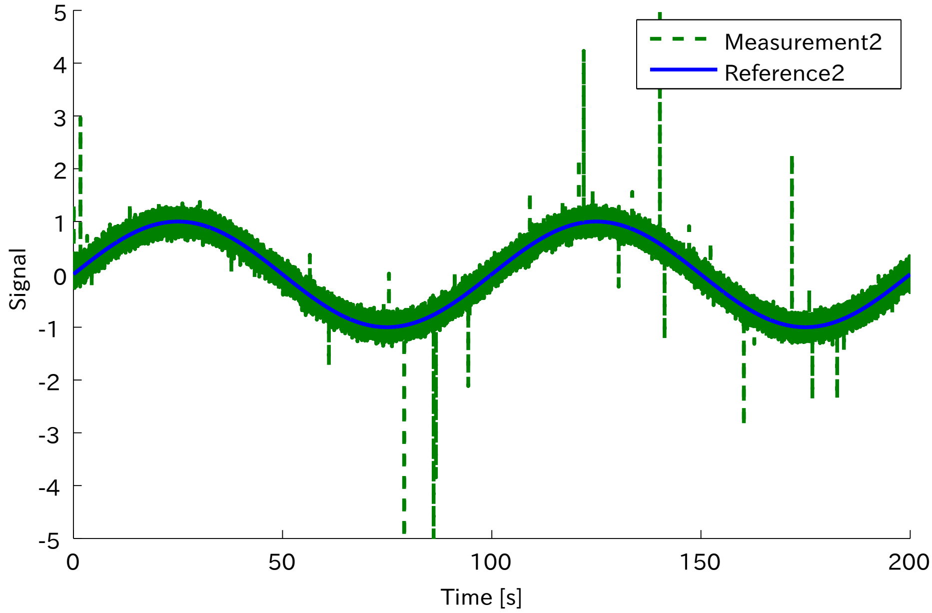
Fig. 5.10 – 5.16 show results in the case of Gaussian mixture distribution. Fig. 5.10 shows measurement outputs with outliers. Fig. 5.12 – 5.13 show outputs of STC with a fixed variance, STC with an updated variance, and RSTC, respectively. Fig. 5.15 – 5.16 show estimates of parameters of these methods.

Table 5.1 and 5.2 show root mean squared errors (RMSEs) of the control errors.

These results show that the proposed RSTC can remove the outliers and reduce control errors more than STC.

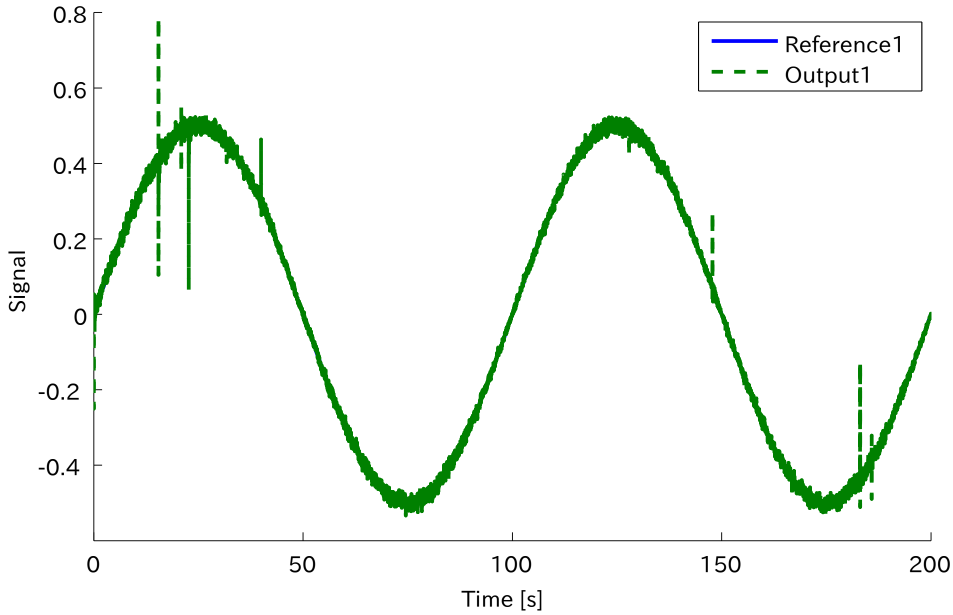


(a) output 1

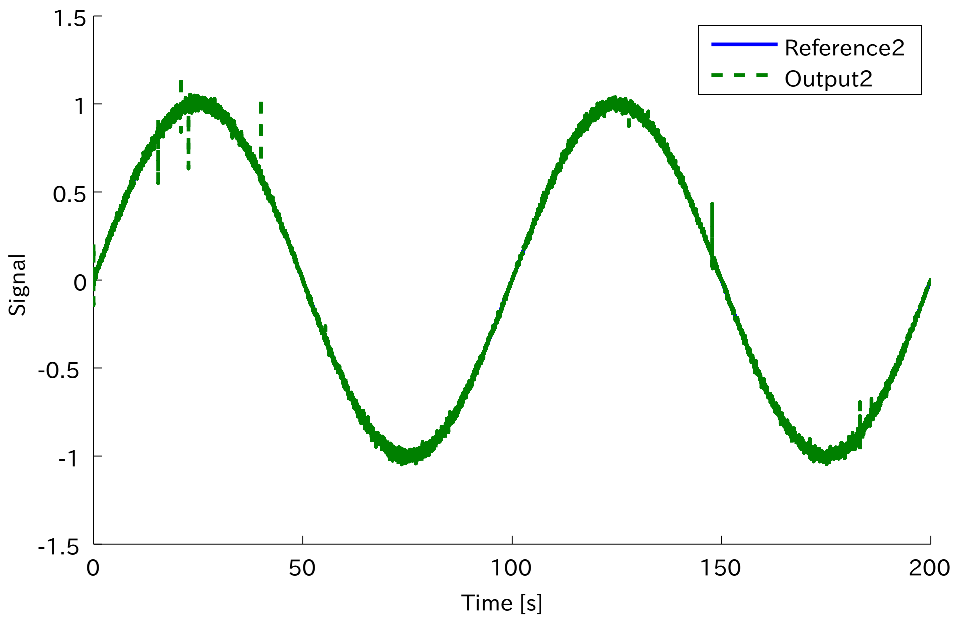


(b) output 2

Figure 5.3: Measurement with outliers using Cauchy distribution in the case using sinusoidal wave as a target. (a) and (b) are output 1 and 2, respectively.

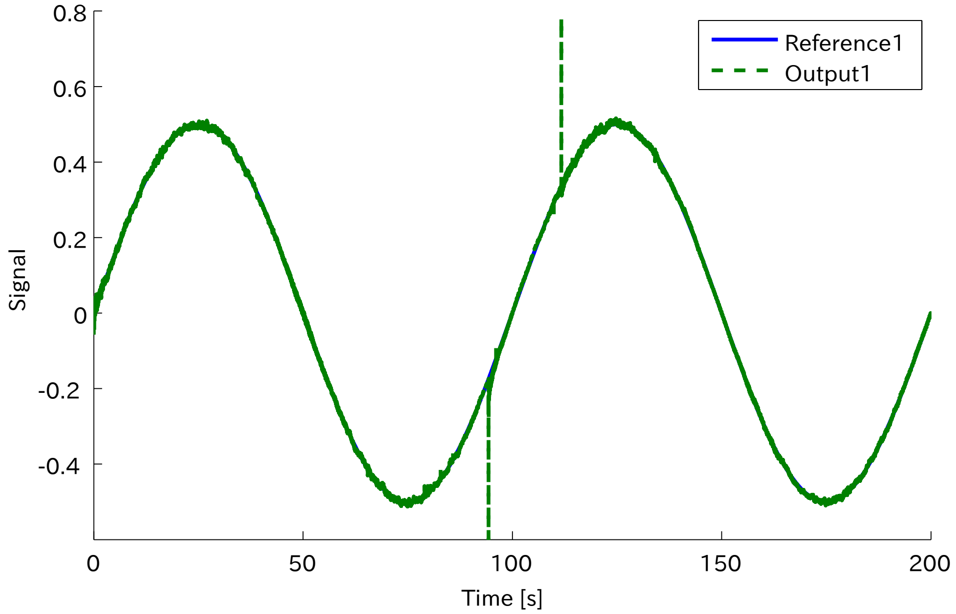


(a) output 1

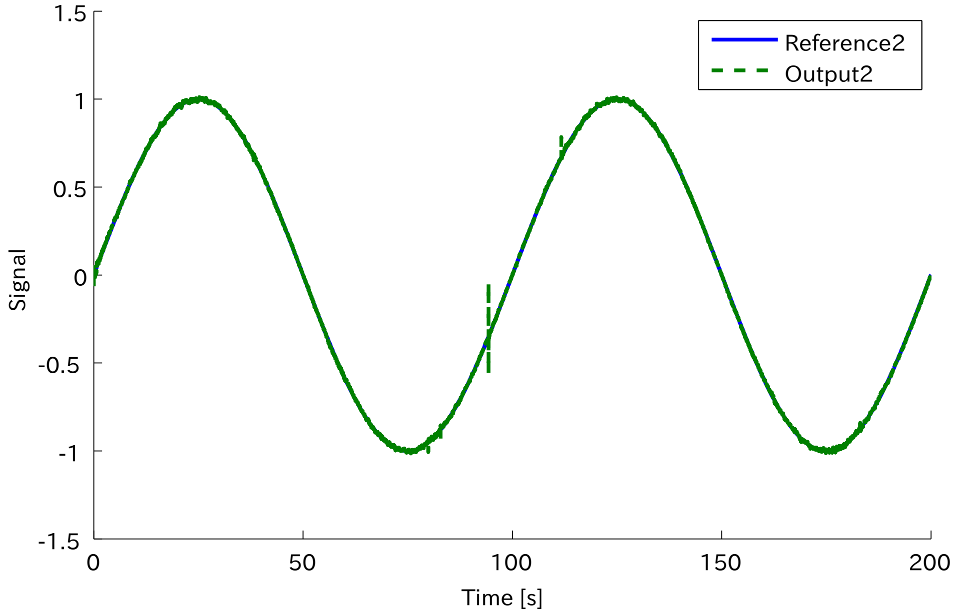


(b) output 2

Figure 5.4: Control performances using normal self tuning controller with a fixed variance under Cauchy noise in the case using sinusoidal wave as a target: The solid line is a reference and dashed line is an output of the controlled object. (a) and (b) are results of output 1 and 2, respectively.

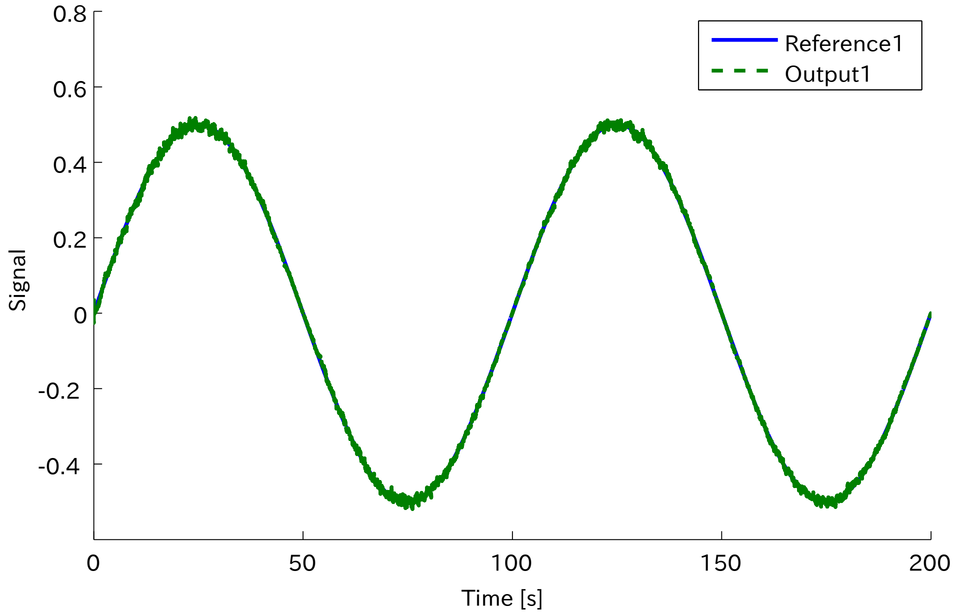


(a) output 1

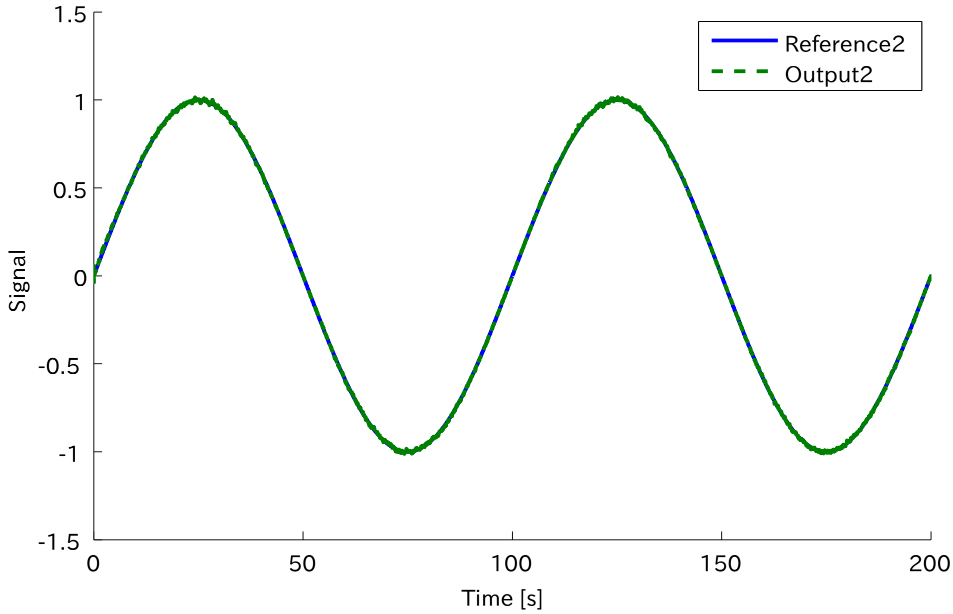


(b) output 2

Figure 5.5: Control performances using normal self tuning controller with an updated variance under Cauchy noise in the case using sinusoidal wave as a target: The solid line is a reference and dashed line is an output of the controlled object. (a) and (b) are results of output 1 and 2, respectively.

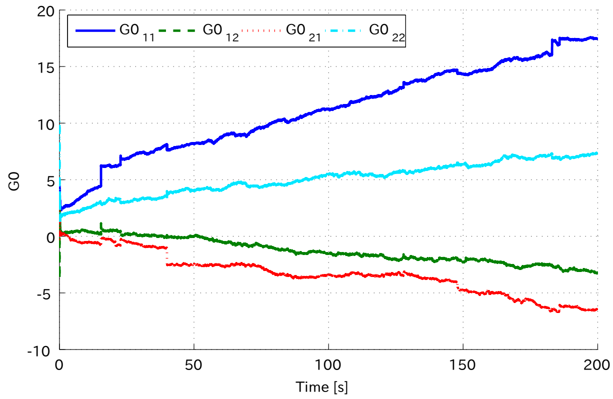


(a) output 1

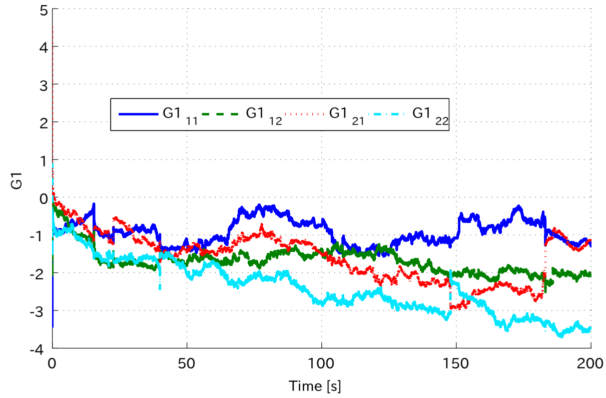


(b) output 2

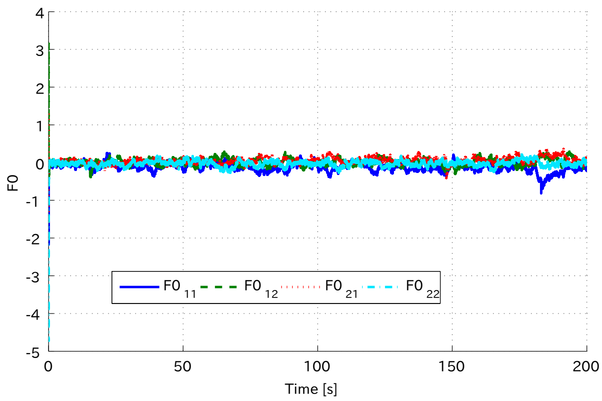
Figure 5.6: Control performances using robust self tuning controller under Cauchy noise in the case using sinusoidal wave as a target: The solid line is a reference and dashed line is an output of the controlled object. (a) and (b) are results of output 1 and 2, respectively.



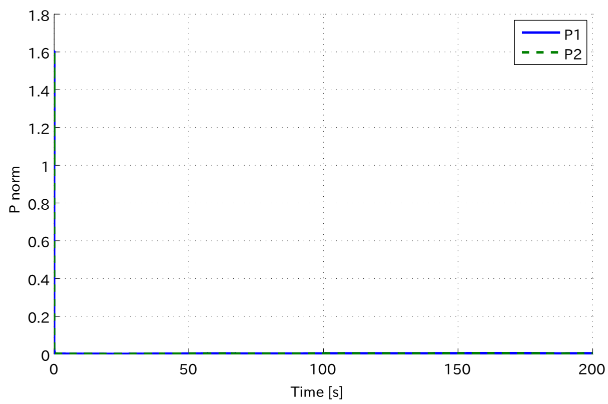
(a) estimates of G_0



(b) estimates of G_1

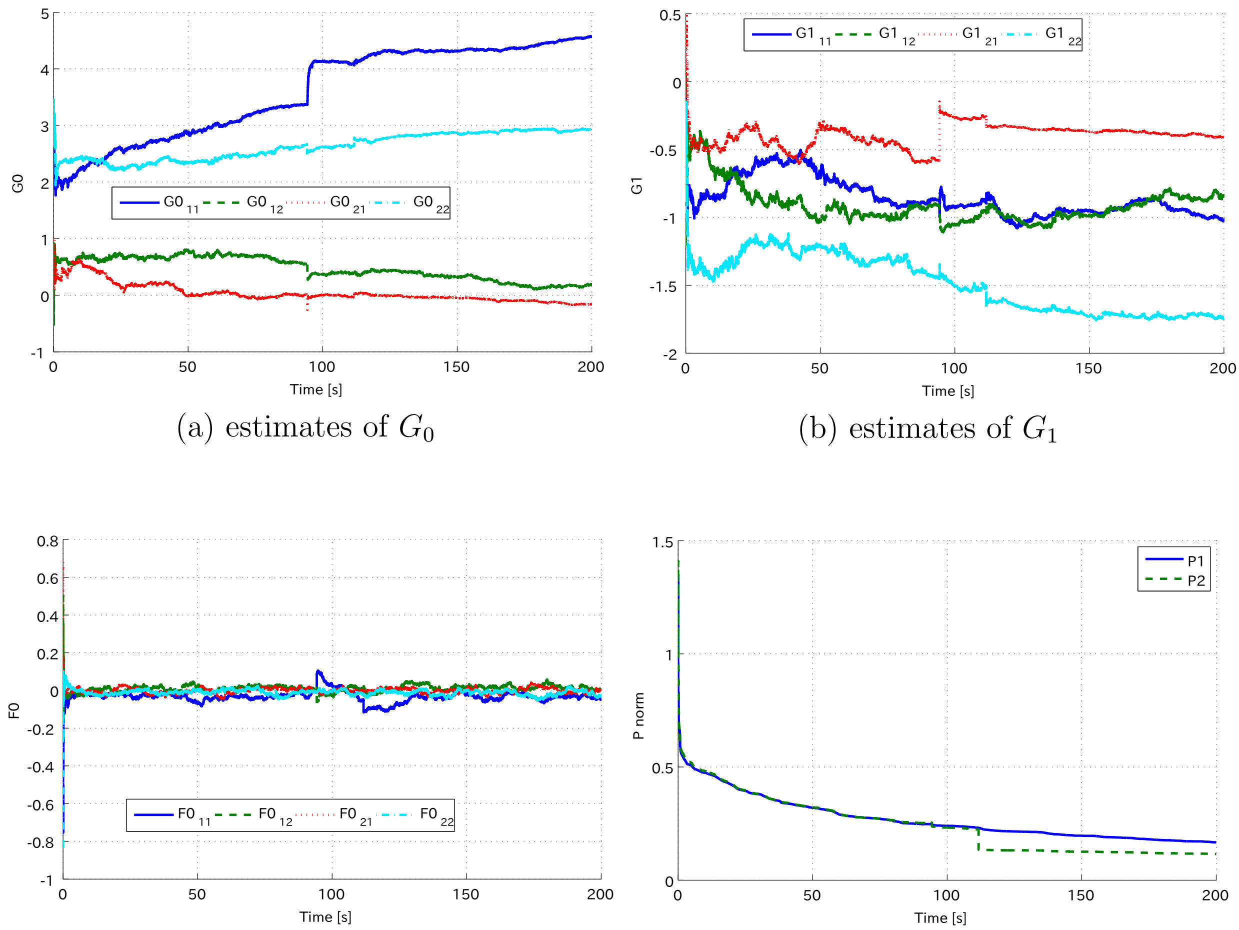


(c) estimates of F_0



(d) ∞-norms of covariance matrices of parameter estimation errors

Figure 5.7: Estimates of parameters and ∞-norms of covariance matrices of parameter estimation errors using normal self tuning controller with a fixed variance under Cauchy noise in the case using sinusoidal wave as a target



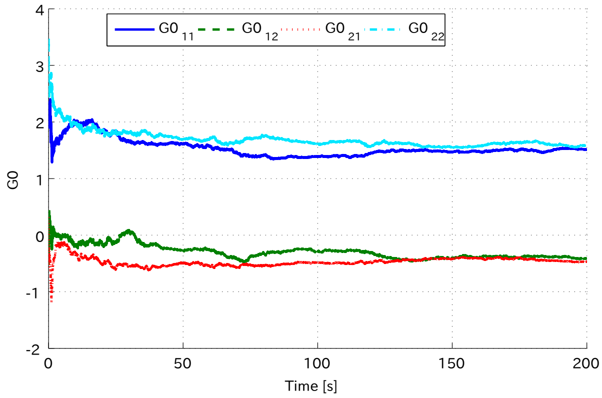
(a) estimates of G_0

(b) estimates of G_1

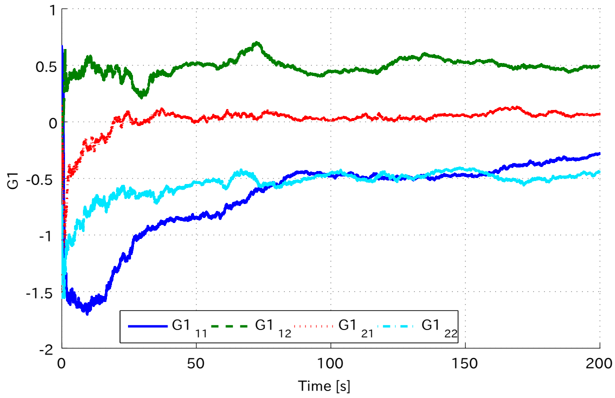
(c) estimates of F_0

(d) ∞-norms of covariance matrices of parameter estimation errors

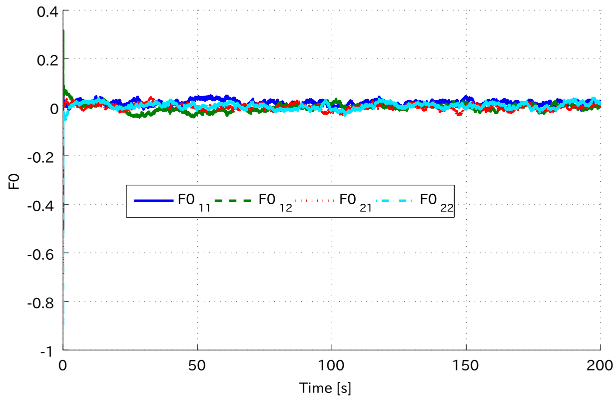
Figure 5.8: Estimates of parameters and ∞-norms of covariance matrices of parameter estimation errors using normal self tuning controller with an updated variance under Cauchy noise in the case using sinusoidal wave as a target



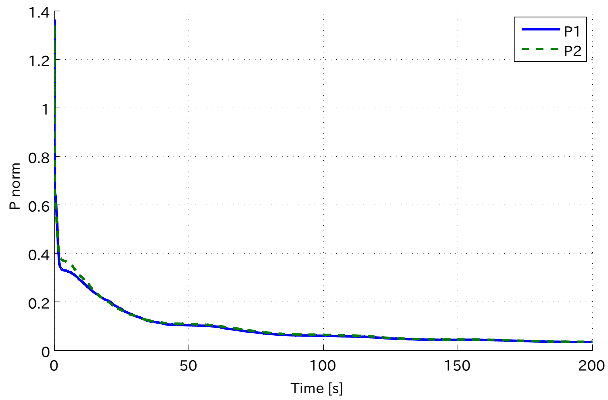
(a) estimates of G_0



(b) estimates of G_1

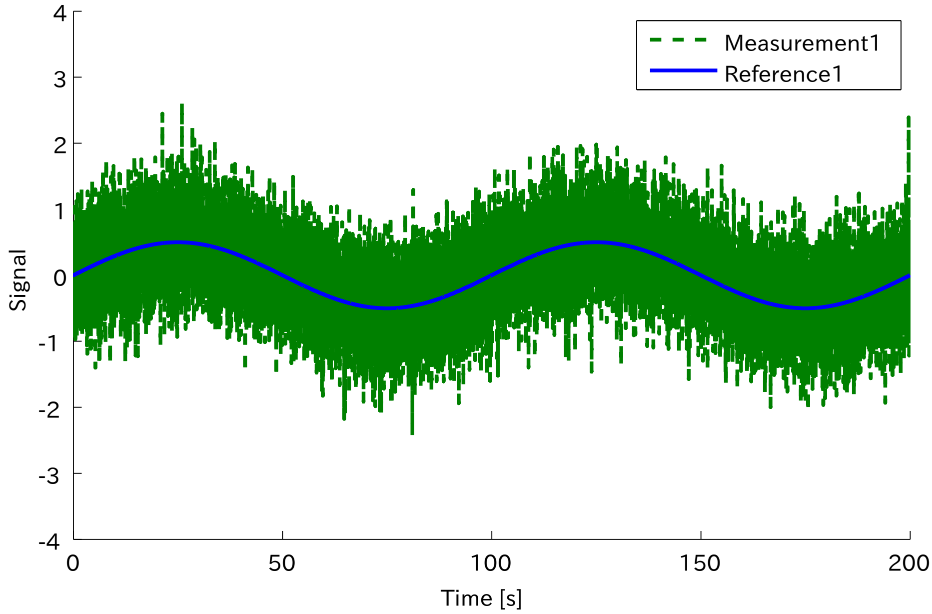


(c) estimates of F_0

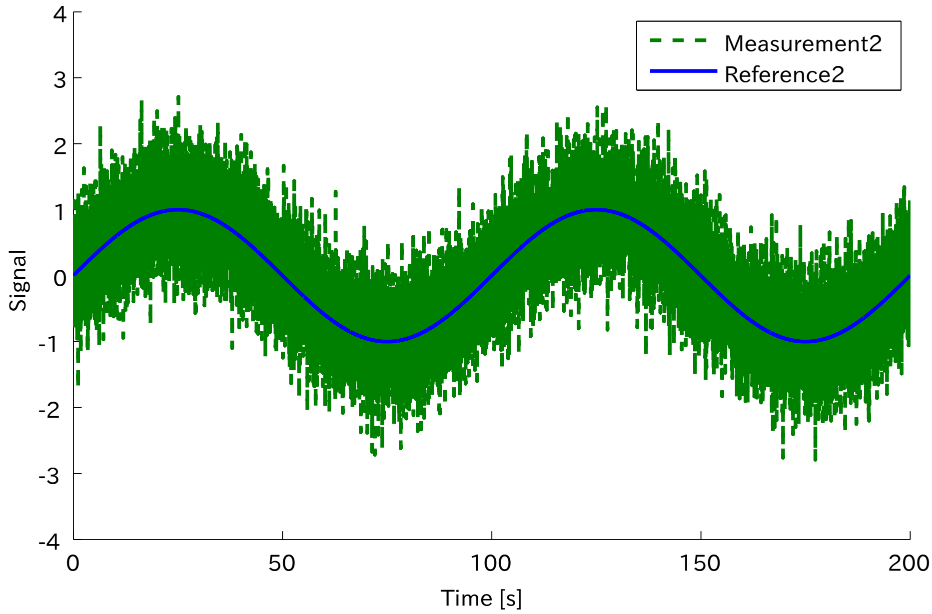


(d) ∞-norms of covariance matrices of parameter estimation errors

Figure 5.9: Estimates of parameters and ∞-norms of covariance matrices of parameter estimation errors using robust self tuning controller under Cauchy noise in the case using sinusoidal wave as a target

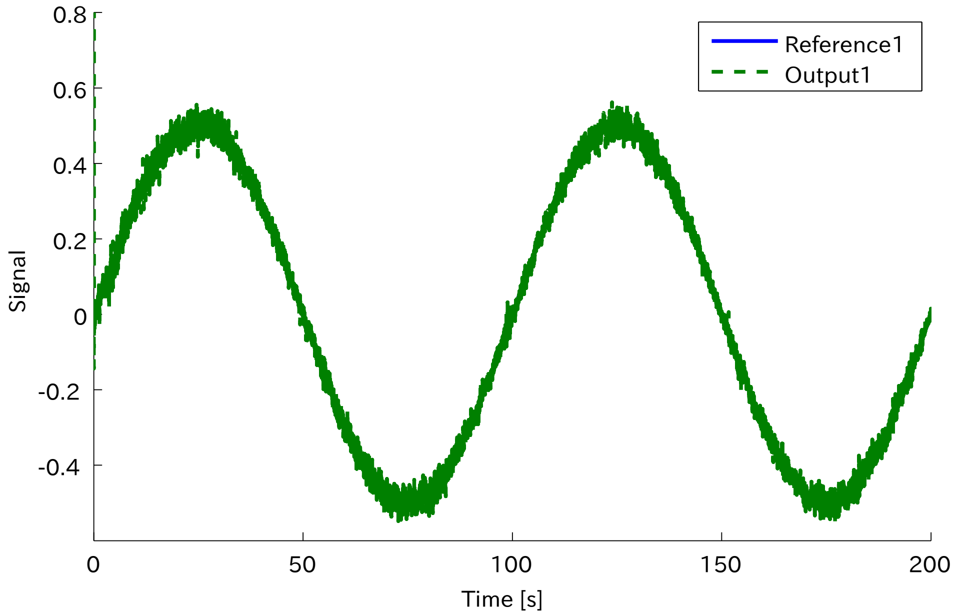


(a) output 1

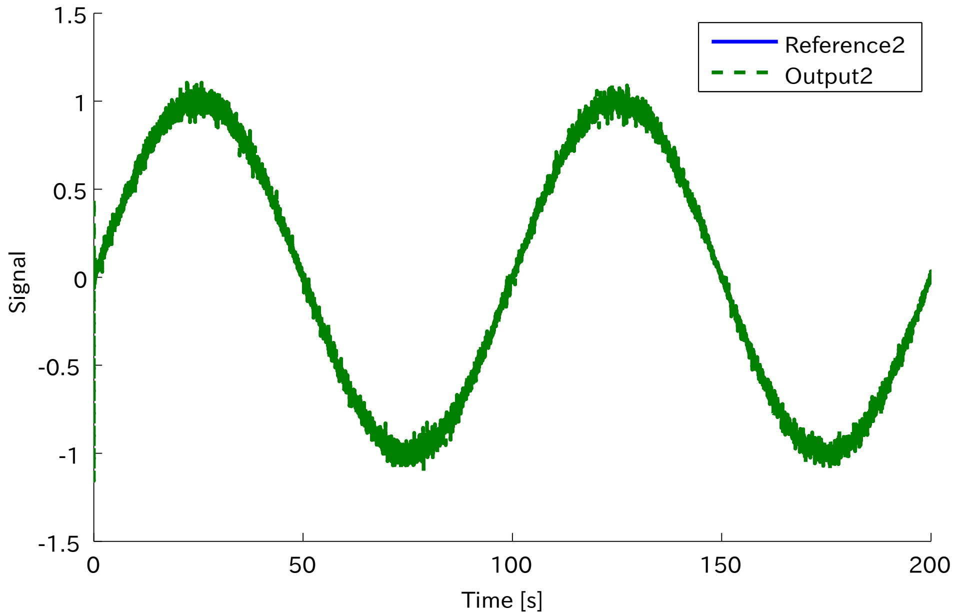


(b) output 2

Figure 5.10: Measurement with outliers using mixed Gaussian distribution in the case using sinusoidal wave as a target. (a) and (b) are output 1 and 2, respectively.

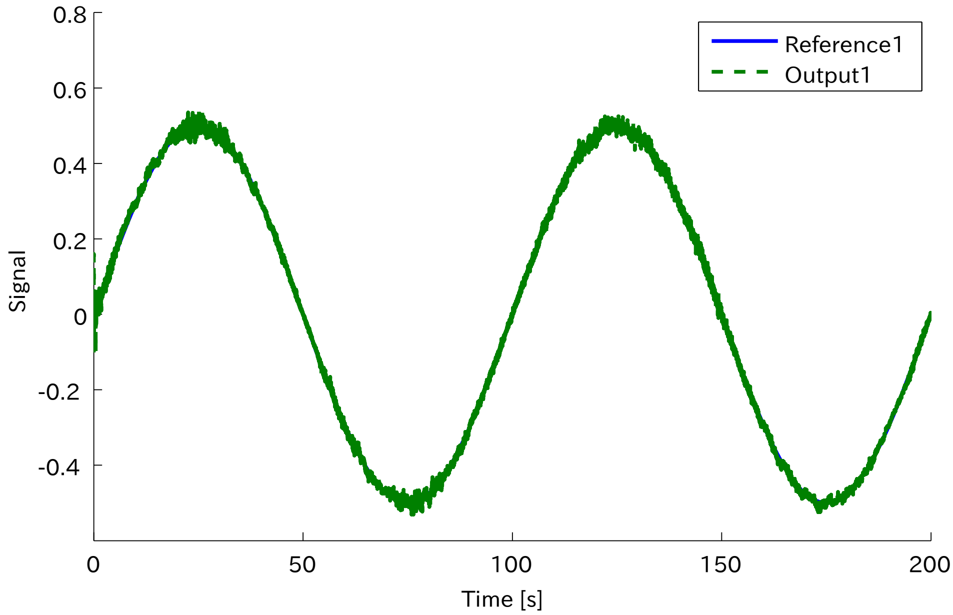


(a) output 1

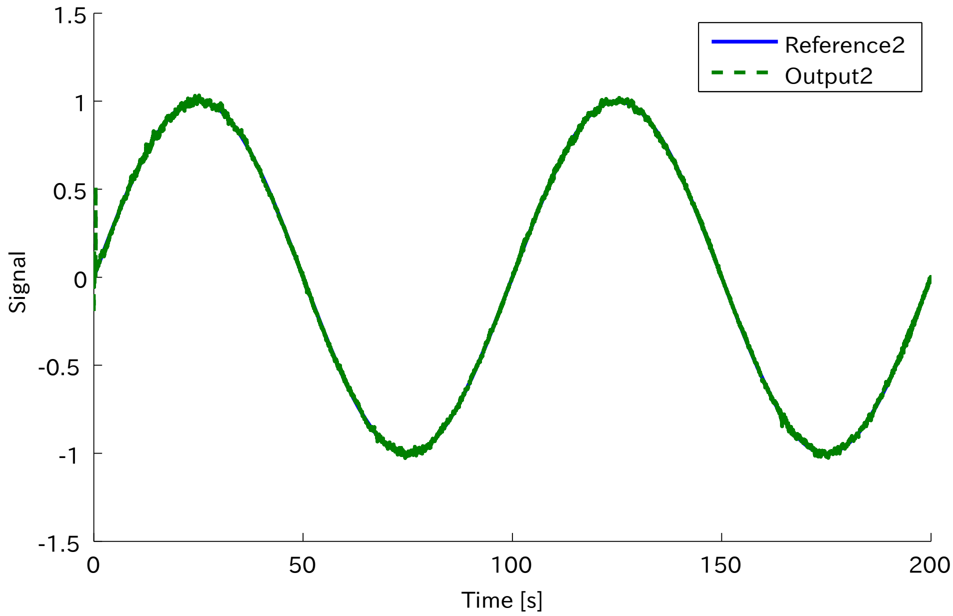


(b) output 2

Figure 5.11: Control performances using normal self tuning controller using a fixed variance under mixed Gaussian noise in the case using sinusoidal wave as a target: The solid line is a reference and dashed line is an output of the controlled object. (a) and (b) are results of output 1 and 2, respectively.

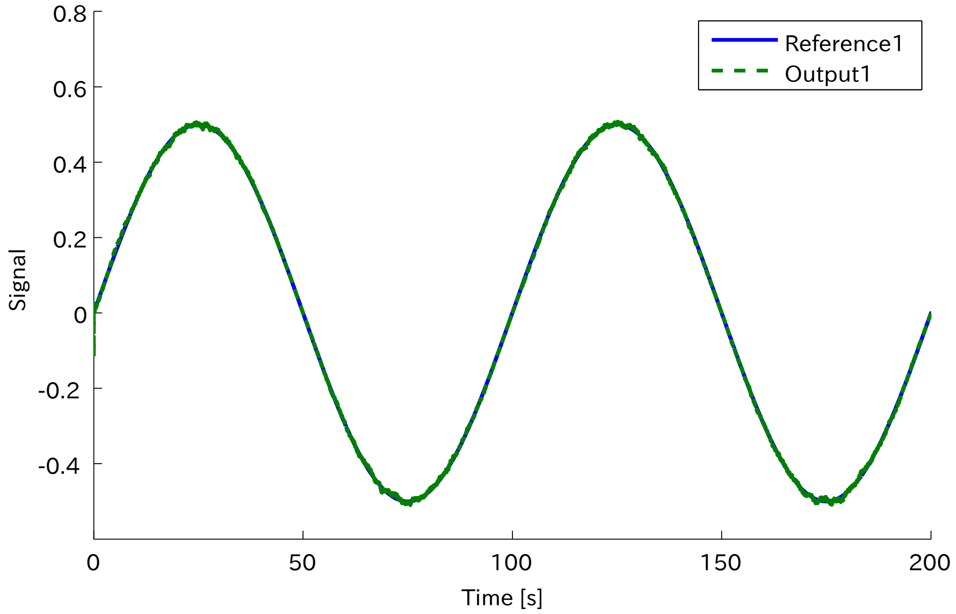


(a) output 1

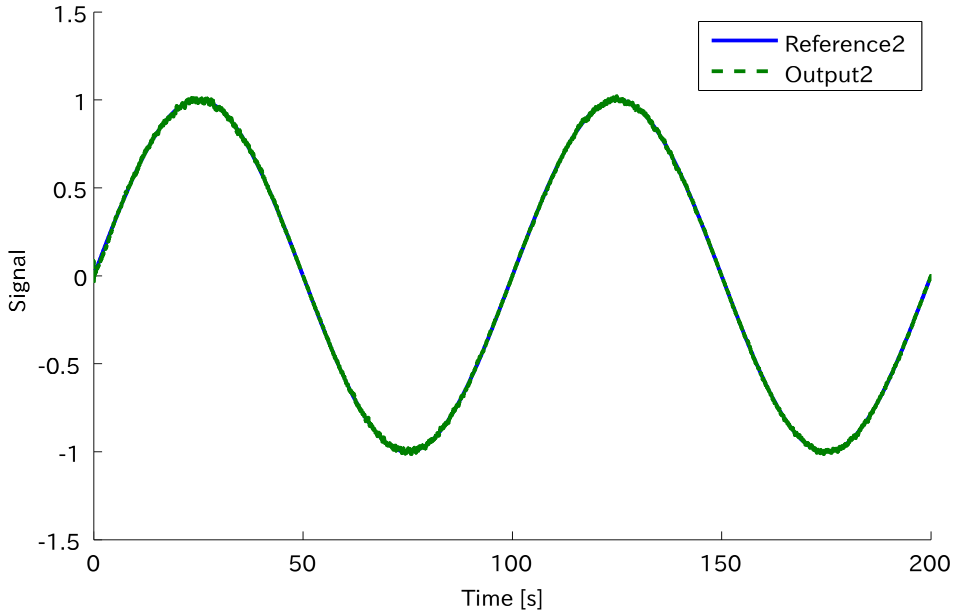


(b) output 2

Figure 5.12: Control performances using normal self tuning controller using an updated variance under mixed Gaussian noise in the case using sinusoidal wave as a target: The solid line is a reference and dashed line is an output of the controlled object. (a) and (b) are results of output 1 and 2, respectively.

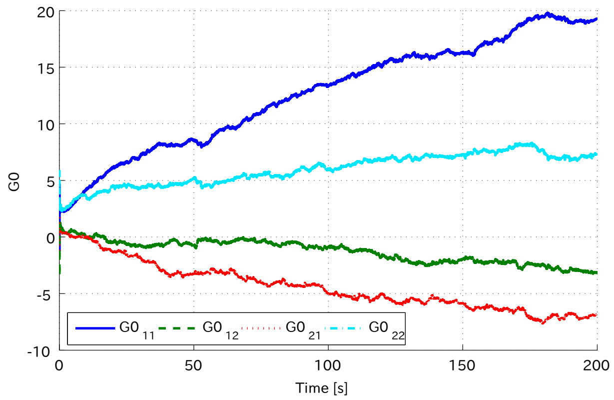


(a) output 1

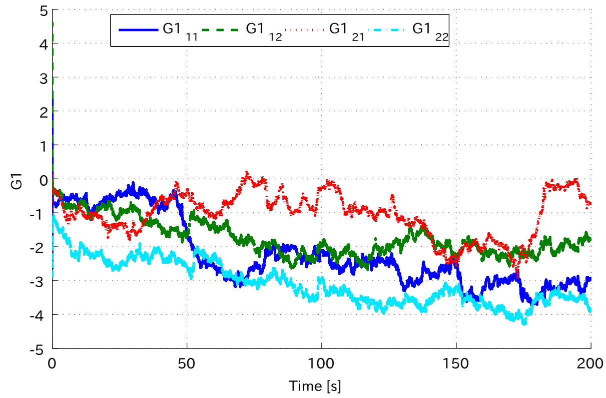


(b) output 2

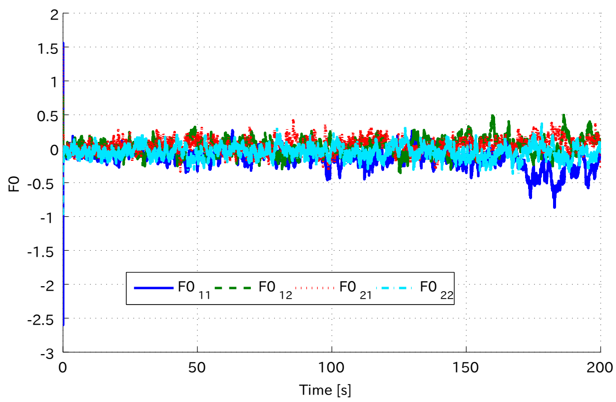
Figure 5.13: Control performances using robust self tuning controller under mixed Gaussian noise in the case using sinusoidal wave as a target: The solid line is a reference and dashed line is an output of the controlled object. (a) and (b) are results of output 1 and 2, respectively.



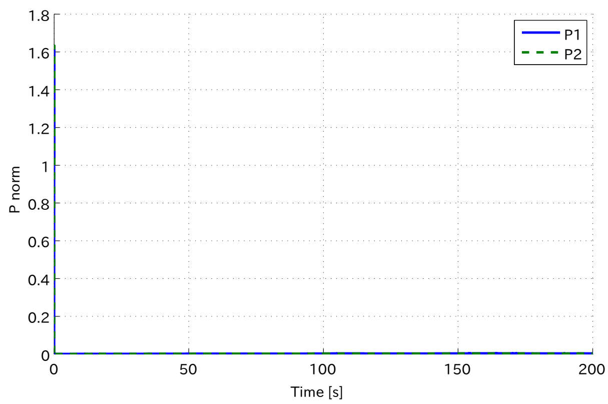
(a) estimates of G_0



(b) estimates of G_1

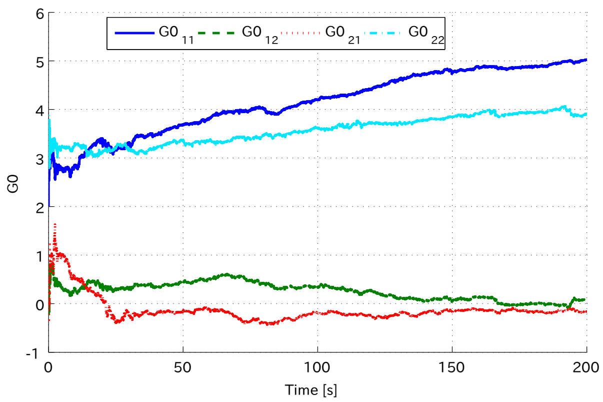


(c) estimates of F_0

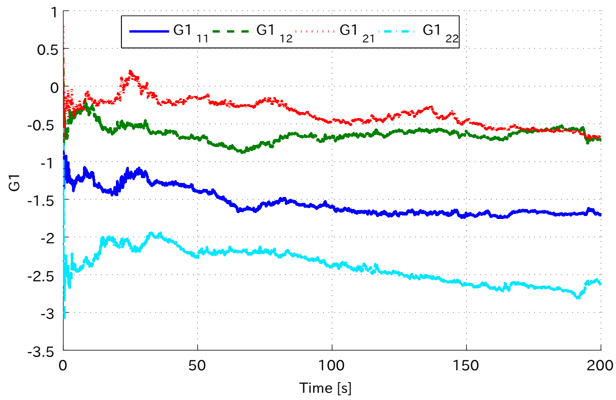


(d) ∞-norms of covariance matrices of parameter estimation errors

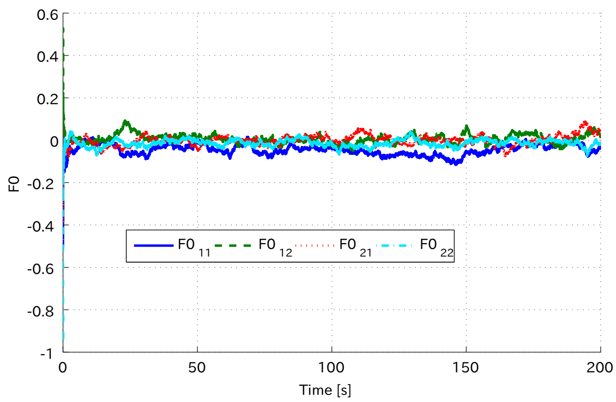
Figure 5.14: Estimates of parameters and ∞-norms of covariance matrices of parameter estimation errors using normal self tuning controller with a fixed variance under Gaussian mixture noise in the case using sinusoidal wave as a target



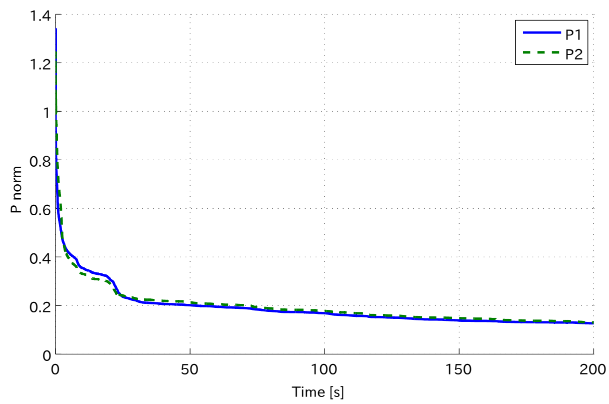
(a) estimates of G_0



(b) estimates of G_1

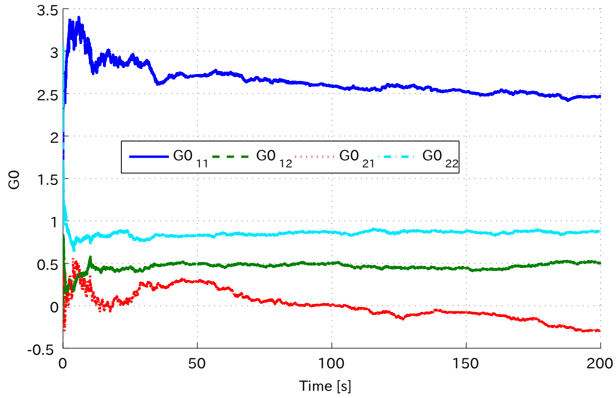


(c) estimates of F_0

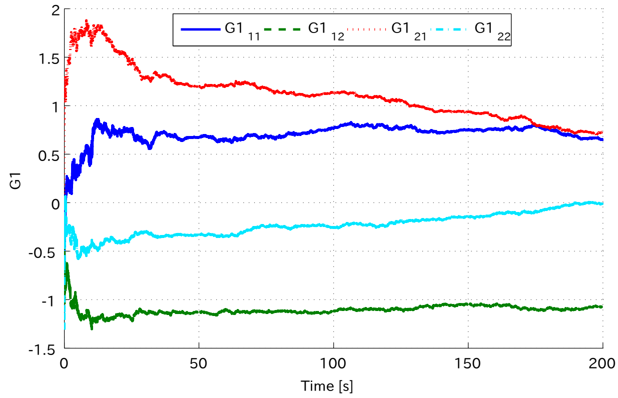


(d) ∞-norms of covariance matrices of parameter estimation errors

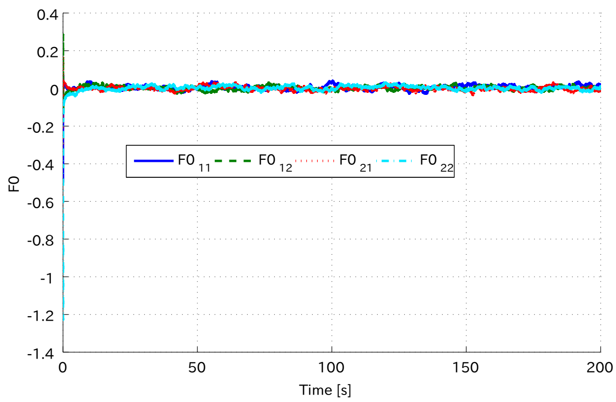
Figure 5.15: Estimates of parameters and ∞-norms of covariance matrices of parameter estimation errors using normal self tuning controller with an updated variance under Gaussian mixture noise in the case using sinusoidal wave as a target



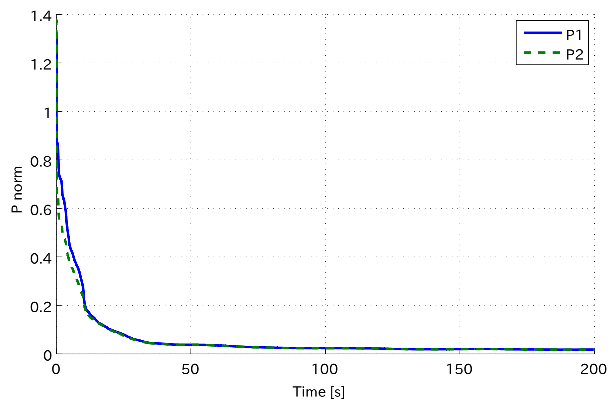
(a) estimates of G_0



(b) estimates of G_1



(c) estimates of F_0



(d) ∞-norms of covariance matrices of parameter estimation errors

Figure 5.16: Estimates of parameters and ∞-norms of covariance matrices of parameter estimation errors using robust self tuning controller under Gaussian mixture noise in the case using sinusoidal wave as a target

In the Case Using a Rectangular Wave as a Target

Fig. 5.17 – 5.30 show simulation results in the case using a rectangular wave as a target.

Fig. 5.17 shows measurement outputs with outliers in the case of Cauchy distribution, Fig. 5.18 – 5.20 show outputs of STC with a fixed variance, STC with an updated variance, and RSTC, respectively. Fig. 5.21 – 5.23 show estimates of parameters of these methods.

Fig. 5.24 shows measurement outputs with outliers in the case of Gaussian mixture distribution, Fig. 5.26 – 5.27 show outputs of STC with a fixed variance, STC with an updated variance, and RSTC, respectively. Fig. 5.29 – 5.30 show estimates of parameters of these methods.

Table 5.3 and 5.4 show RMSEs of the control errors.

These results show that the proposed RSTC can remove the outliers in steady state and reduce control errors more than STC. However, the proposed method deteriorates a transient response. For example, in Fig. 5.20 at 100 second, output 1 has an overshoot more than STC. The proposed method cannot estimate outliers at the change point well, then it results in a degradation of a control performance. The proposed method may estimate the change point as outliers.

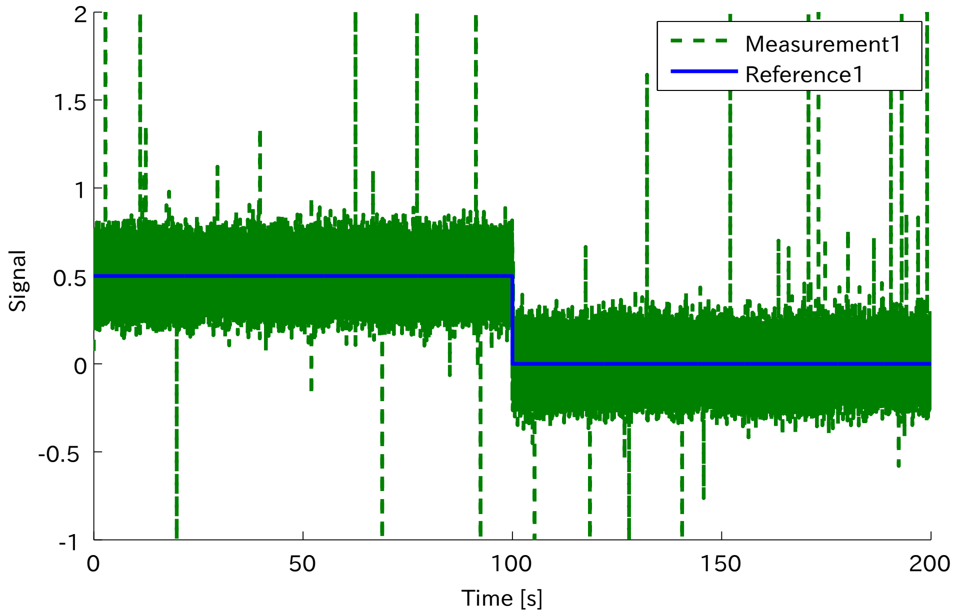
From Fig. 5.21, 5.23, 5.28, and 5.30, covariance matrices of STC with an updated variance and RSTC are changed at 100 second, discontinuously. Kalman gain depends on $R^s_{k,i}$, and $R^s_{k,i}$ is estimated by Eq. (5.36) in which estimated outliers and parameters are used. These estimates have smaller variances after a changing point of a target than before because the target has no offset after the changing point. Therefore, $R^s_{k,i}$ becomes smaller after the changing point than before, and the covariance matrices also become small, discontinuously.

Table 5.3: Root mean squared errors of tracking errors of the controlled plant under Cauchy noise in the case using a rectangle wave as a target.

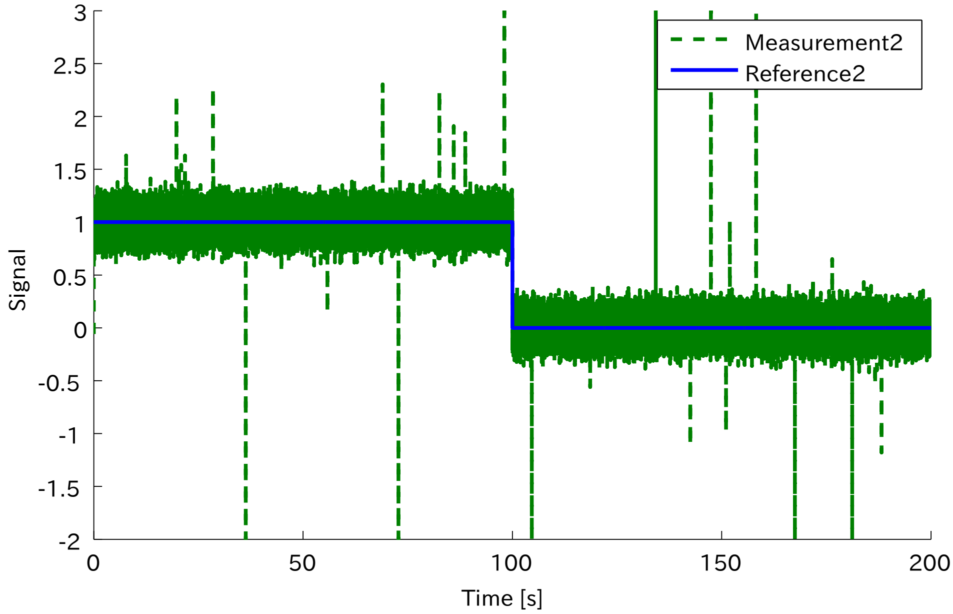
	output 1	output 2
self tuning controller with a fixed variance	1.5×10^{-2}	2.0×10^{-2}
self tuning controller with an updated variance	3.8×10^{-3}	6.1×10^{-3}
robust self tuning controller (proposed method)	3.3×10^{-3}	5.7×10^{-3}

Table 5.4: Root mean squared errors of tracking errors of the controlled plant under Gaussian mixture noise in the case using a rectangle wave as a target.

	output 1	output 2
self tuning controller with a fixed variance	1.3×10^{-2}	2.0×10^{-2}
self tuning controller with an updated variance	9.5×10^{-3}	8.9×10^{-3}
robust self tuning controller (proposed method)	4.6×10^{-3}	5.9×10^{-3}

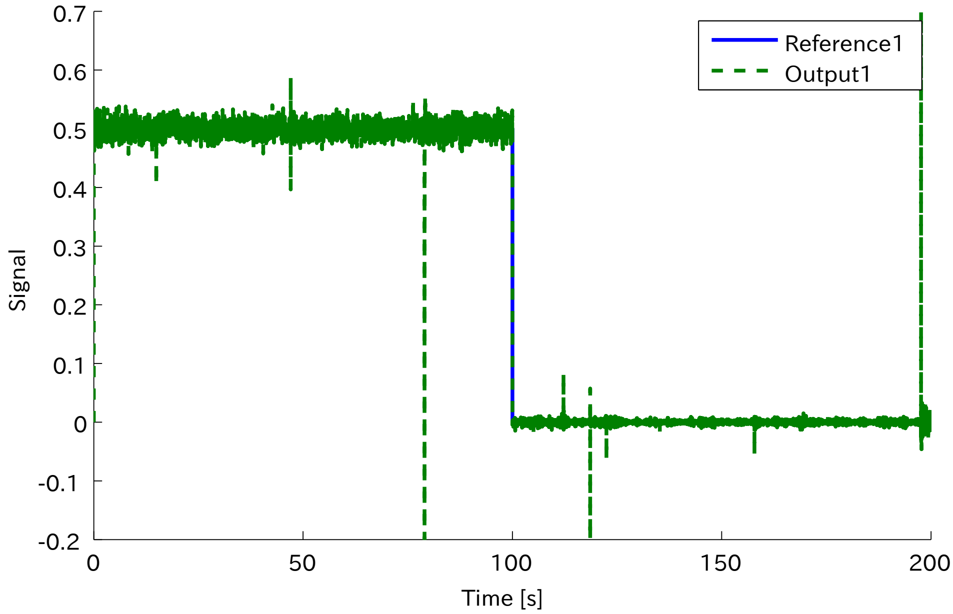


(a) output 1

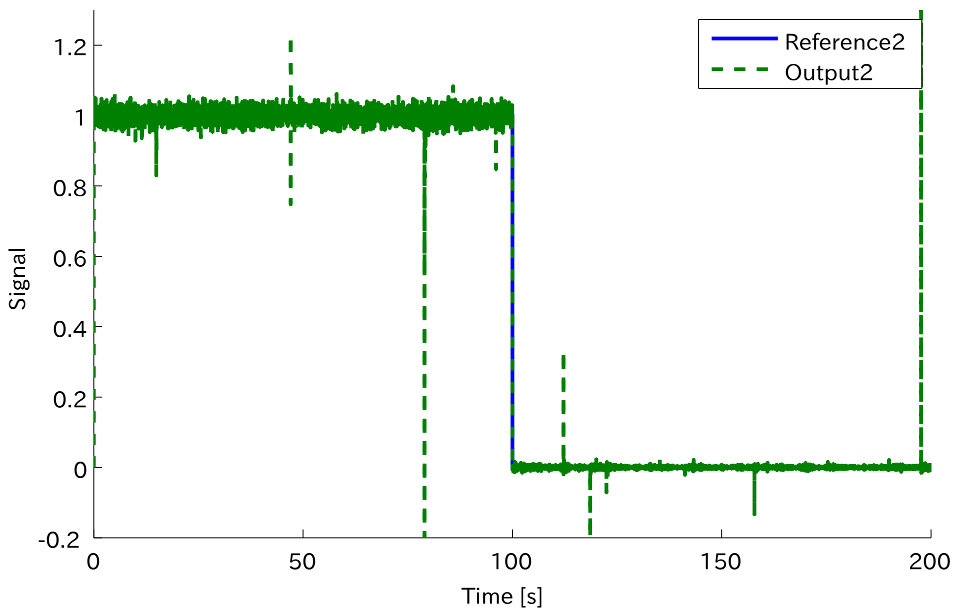


(b) output 2

Figure 5.17: Measurement with outliers using Cauchy distribution. (a) and (b) are output 1 and 2, respectively.

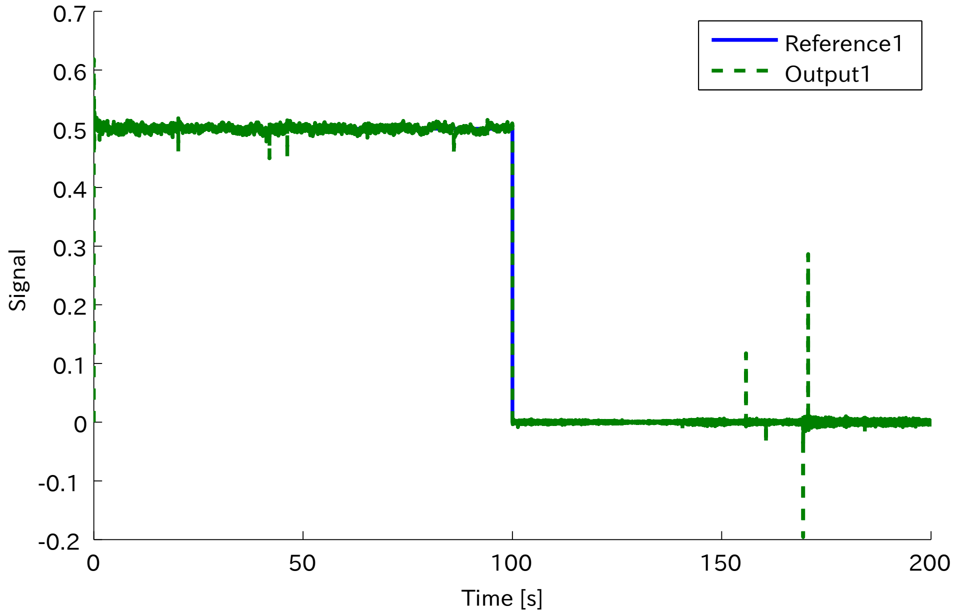


(a) output 1

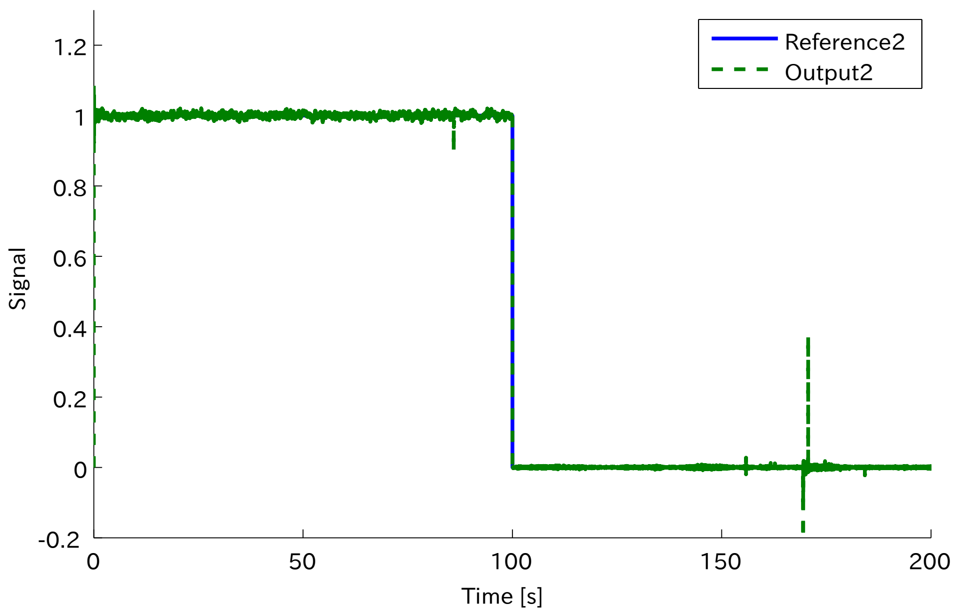


(b) output 2

Figure 5.18: Control performances using normal self tuning controller with a fixed variance under Cauchy noise: The solid line is a reference and dashed line is an output of the controlled object. (a) and (b) are results of output 1 and 2, respectively.

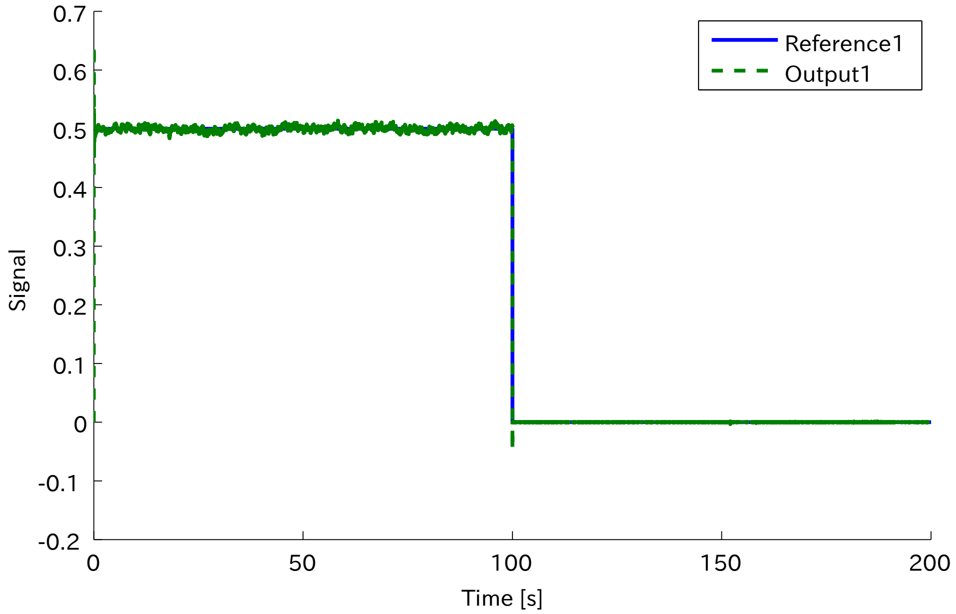


(a) output 1

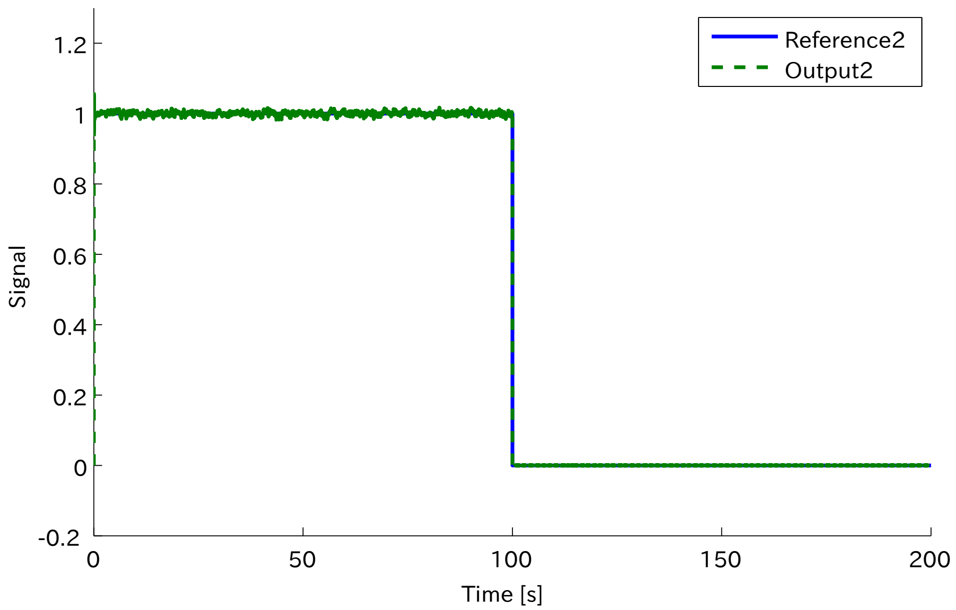


(b) output 2

Figure 5.19: Control performances using normal self tuning controller with an updated variance under Cauchy noise: The solid line is a reference and dashed line is an output of the controlled object. (a) and (b) are results of output 1 and 2, respectively.

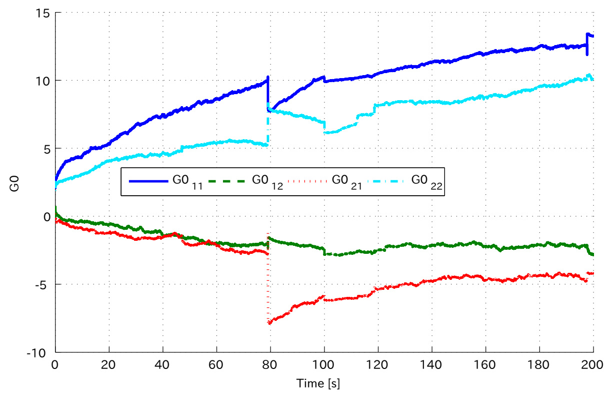


(a) output 1

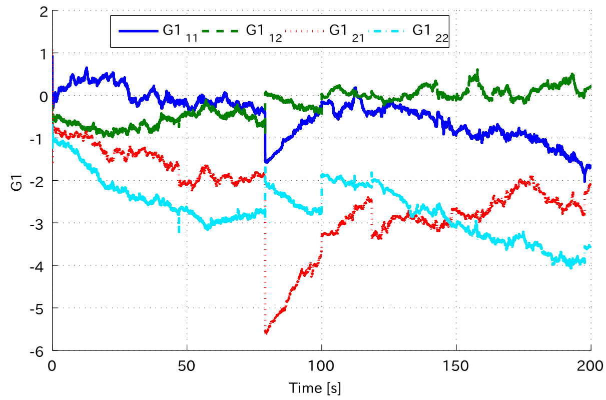


(b) output 2

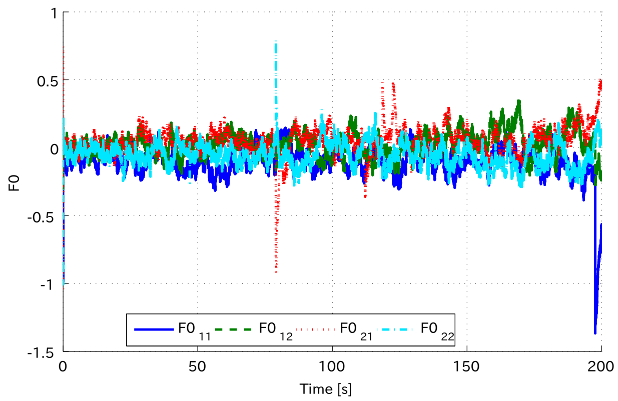
Figure 5.20: Control performances using robust self tuning controller under Cauchy noise: The solid line is a reference and dashed line is an output of the controlled object. (a) and (b) are results of output 1 and 2, respectively.



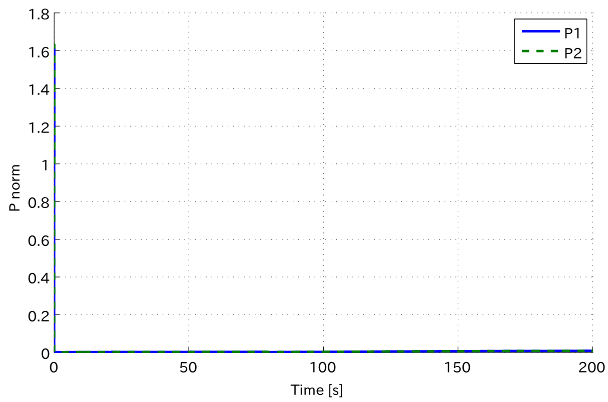
(a) estimates of G_0



(b) estimates of G_1

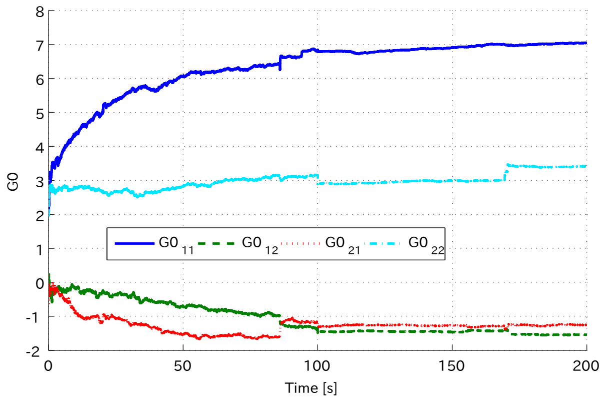


(c) estimates of F_0

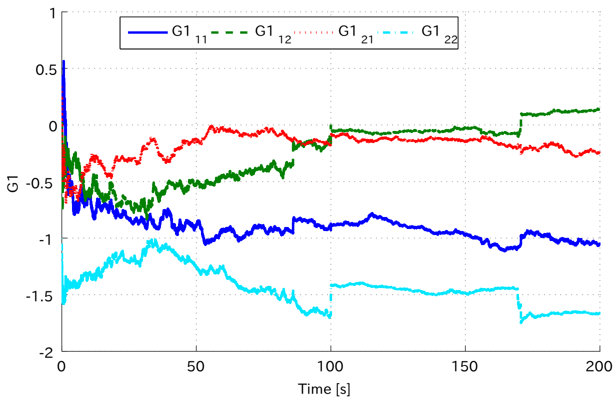


(d) ∞-norms of covariance matrices of parameter estimation errors

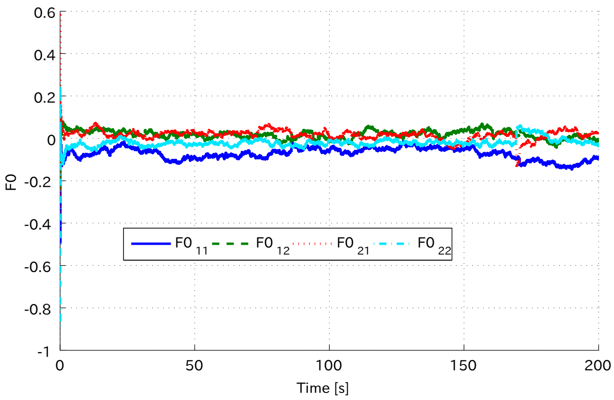
Figure 5.21: Estimates of parameters and ∞-norms of covariance matrices of parameter estimation errors using normal self tuning controller with a fixed variance under Cauchy noise



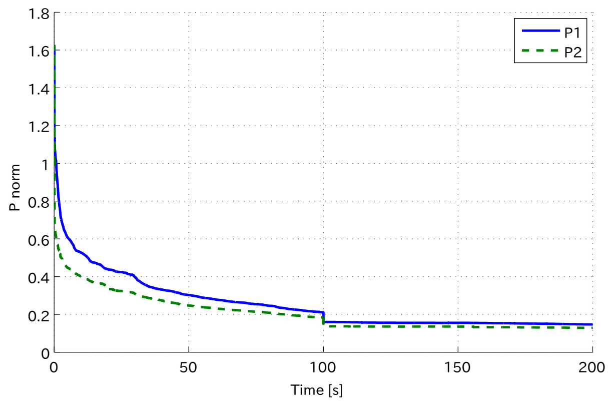
(a) estimates of G_0



(b) estimates of G_1

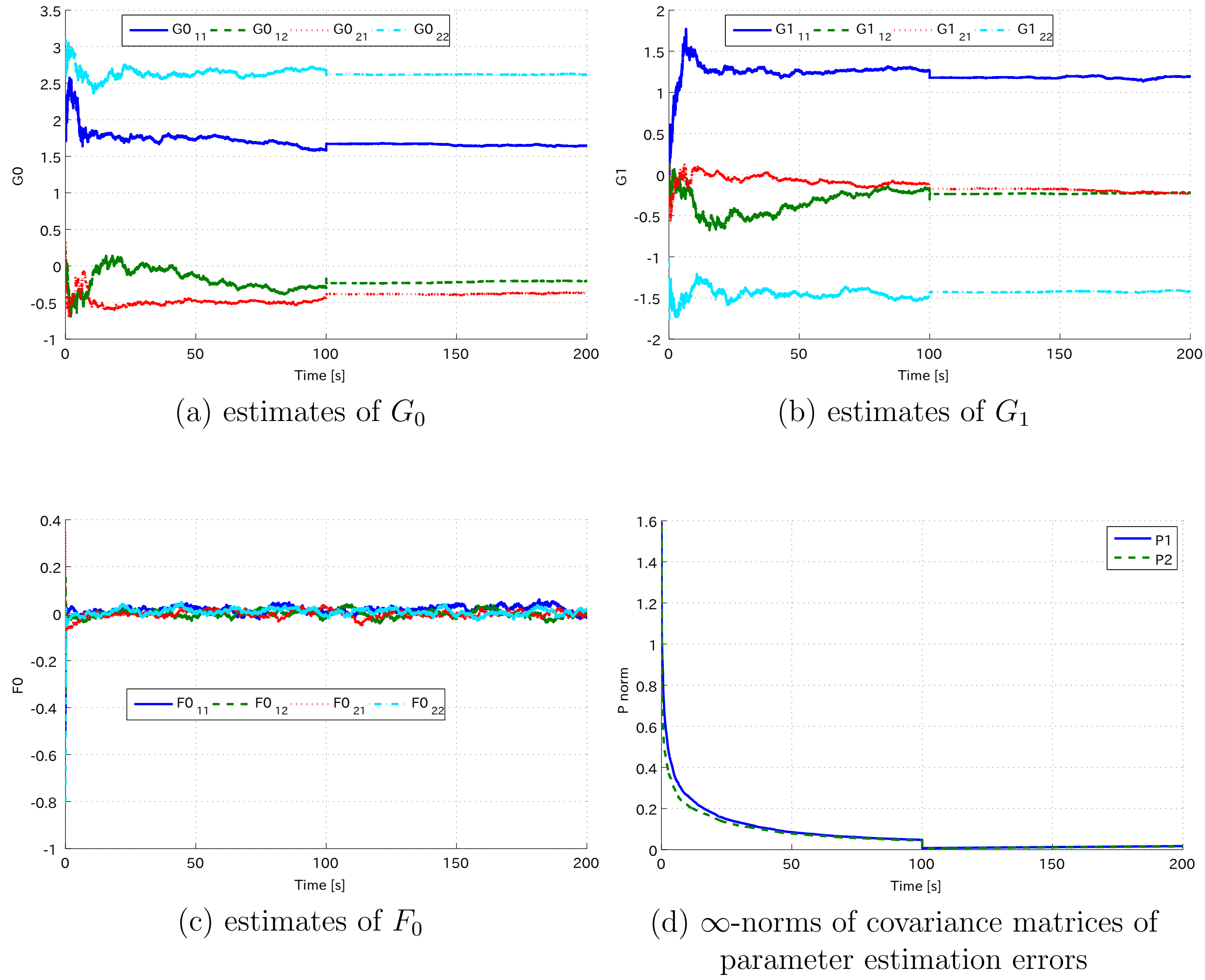


(c) estimates of F_0



(d) ∞-norms of covariance matrices of parameter estimation errors

Figure 5.22: Estimates of parameters and ∞-norms of covariance matrices of parameter estimation errors using normal self tuning controller with an updated variance under Cauchy noise



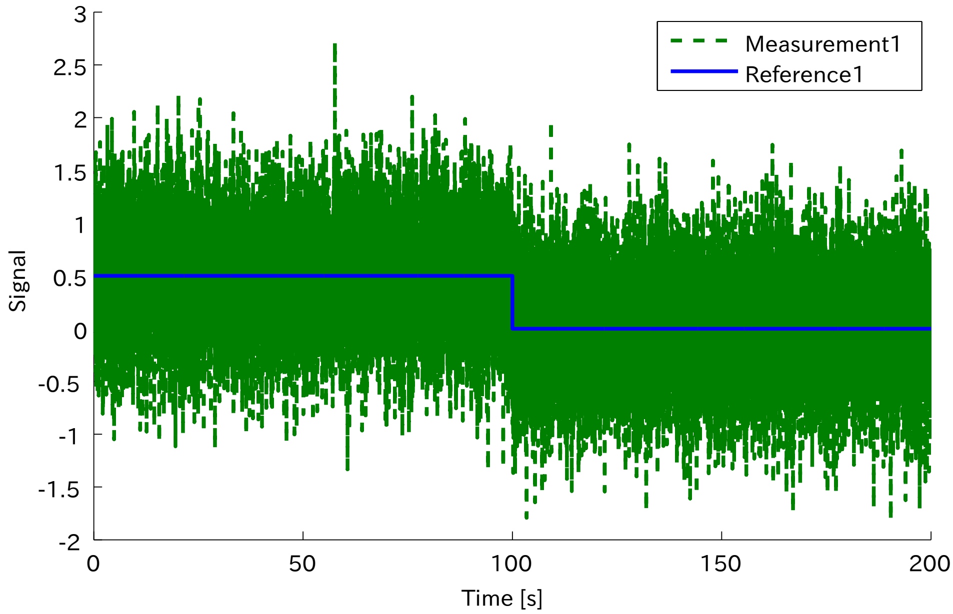
(a) estimates of G_0

(b) estimates of G_1

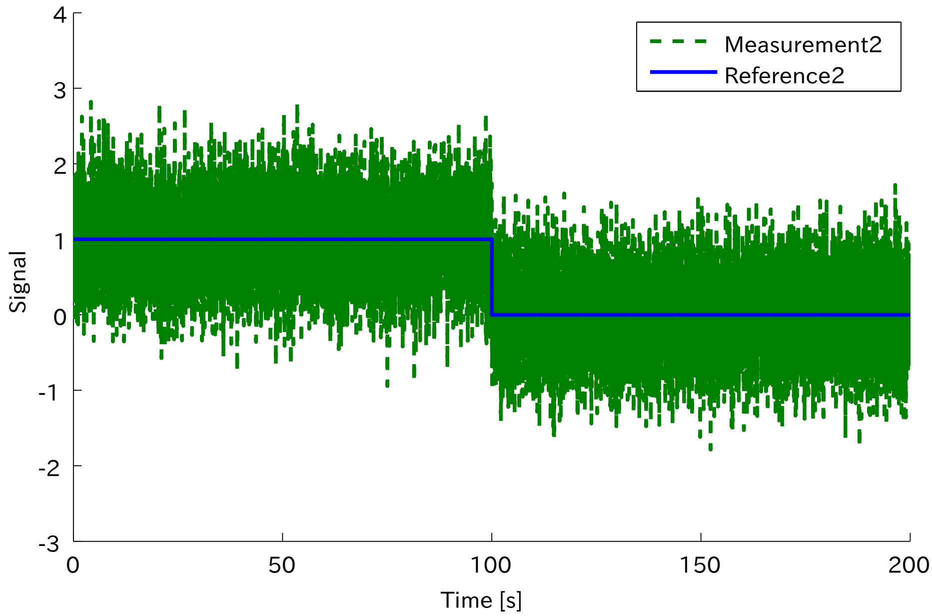
(c) estimates of F_0

(d) ∞-norms of covariance matrices of parameter estimation errors

Figure 5.23: Estimates of parameters and ∞-norms of covariance matrices of parameter estimation errors using robust self tuning controller under Cauchy noise

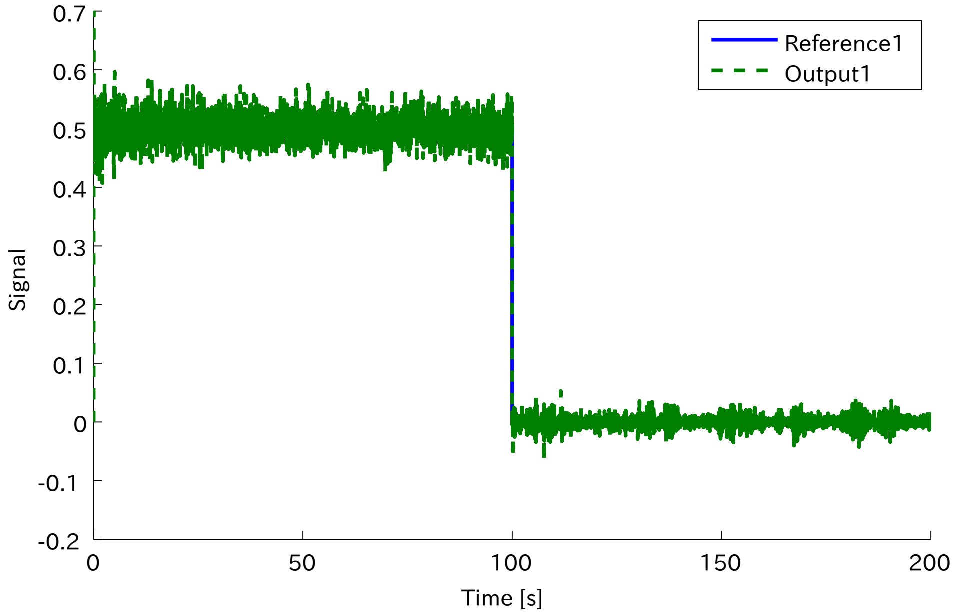


(a) output 1

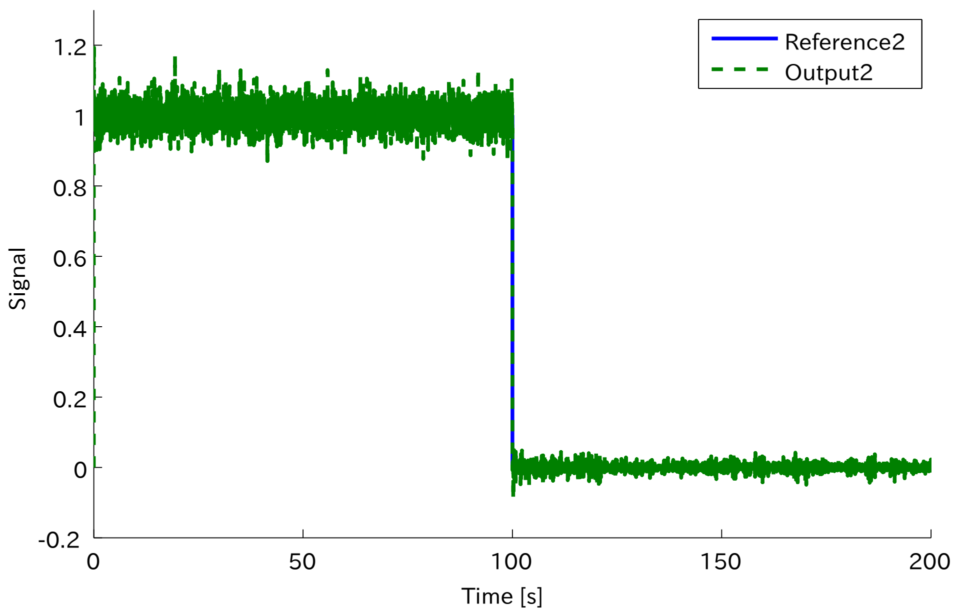


(b) output 2

Figure 5.24: Measurement with outliers using mixed Gaussian distribution. (a) and (b) are output 1 and 2, respectively.

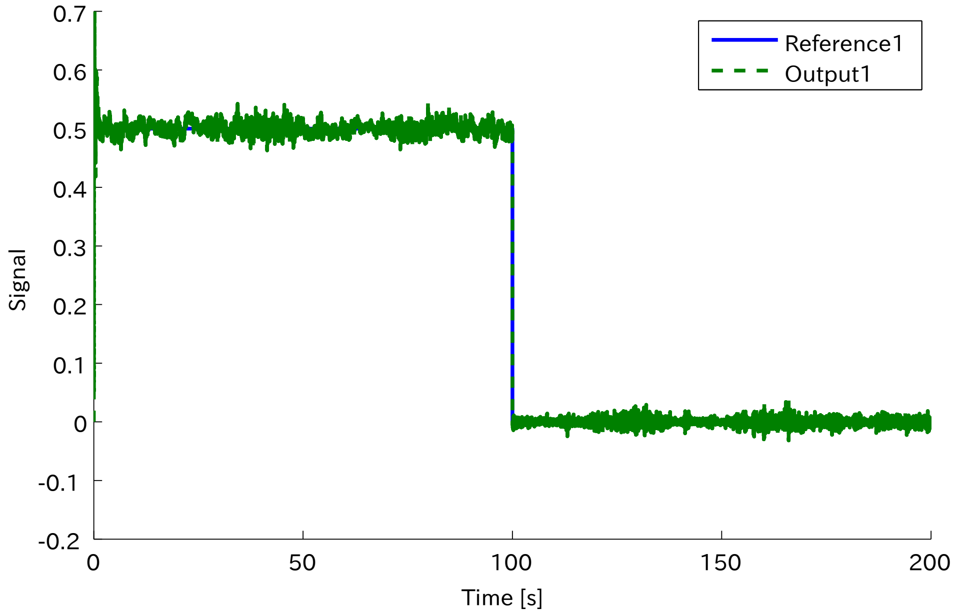


(a) output 1

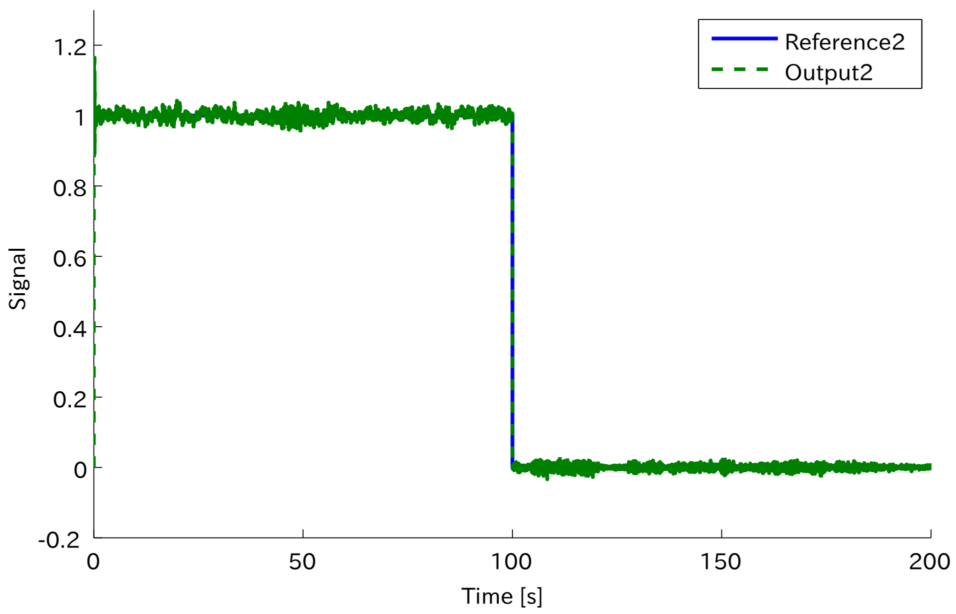


(b) output 2

Figure 5.25: Control performances using normal self tuning controller using a fixed variance under mixed Gaussian noise: The solid line is a reference and dashed line is an output of the controlled object. (a) and (b) are results of output 1 and 2, respectively.

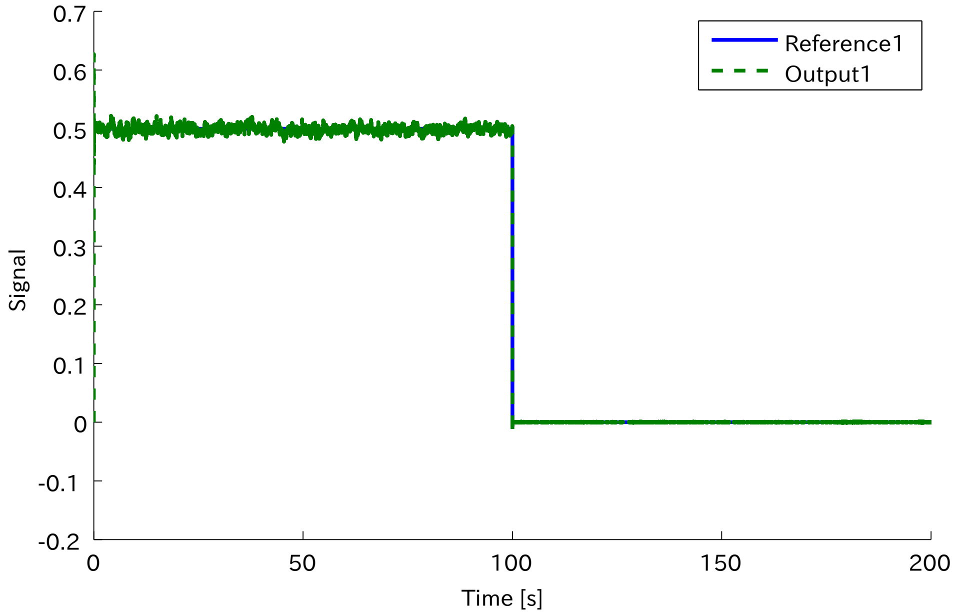


(a) output 1

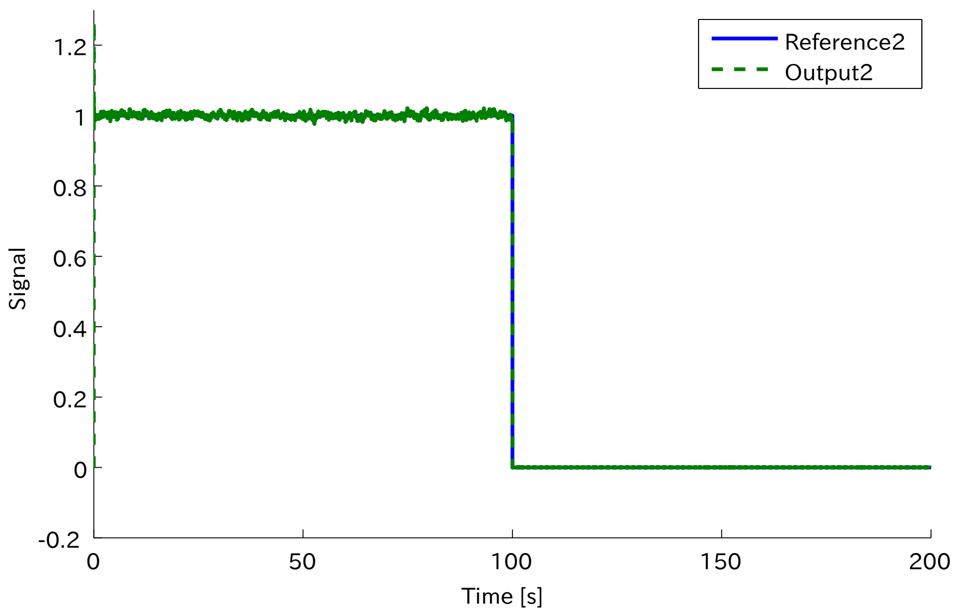


(b) output 2

Figure 5.26: Control performances using normal self tuning controller using an updated variance under mixed Gaussian noise: The solid line is a reference and dashed line is an output of the controlled object. (a) and (b) are results of output 1 and 2, respectively.

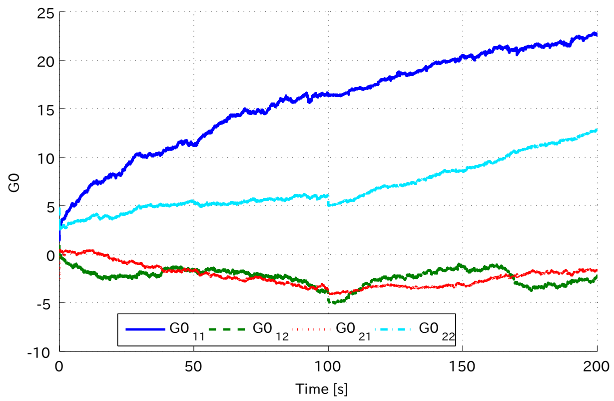


(a) output 1

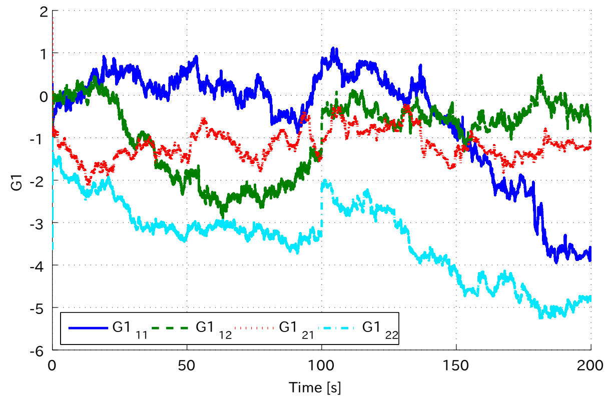


(b) output 2

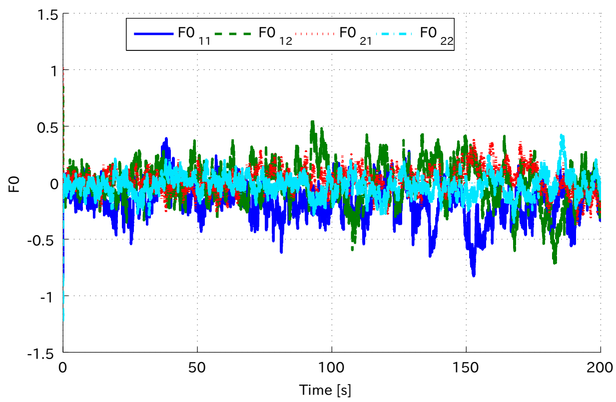
Figure 5.27: Control performances using robust self tuning controller under mixed Gaussian noise: The solid line is a reference and dashed line is an output of the controlled object. (a) and (b) are results of output 1 and 2, respectively.



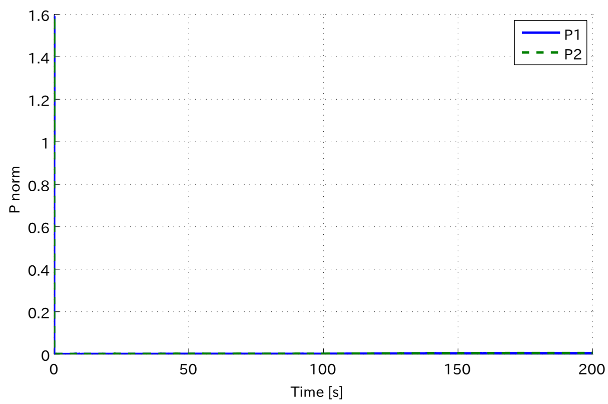
(a) estimates of G_0



(b) estimates of G_1

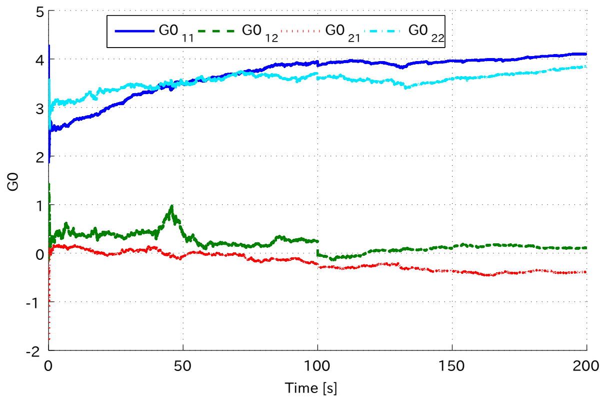


(c) estimates of F_0

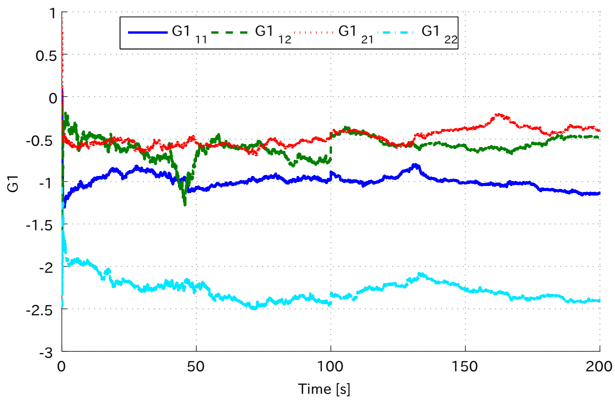


(d) ∞-norms of covariance matrices of parameter estimation errors

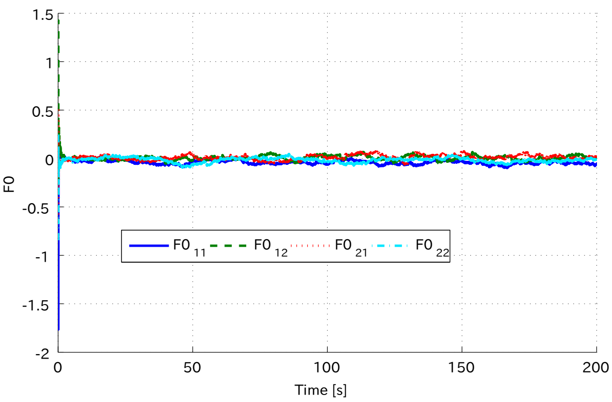
Figure 5.28: Estimates of parameters and ∞-norms of covariance matrices of parameter estimation errors using normal self tuning controller with a fixed variance under Gaussian mixture noise



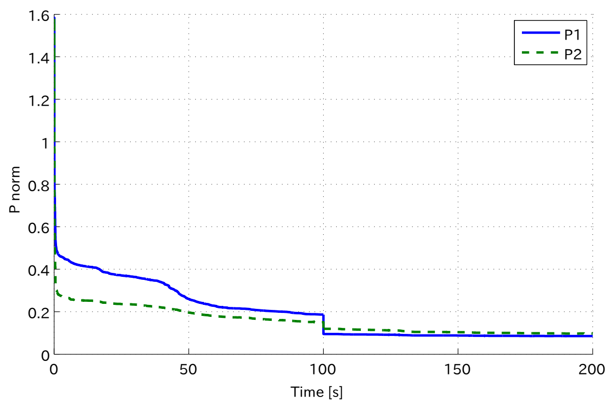
(a) estimates of G_0



(b) estimates of G_1

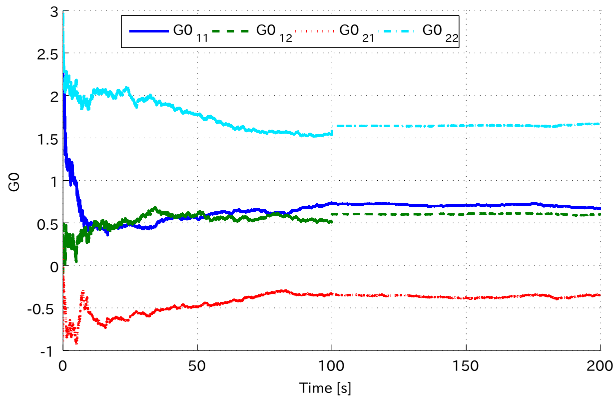


(c) estimates of F_0

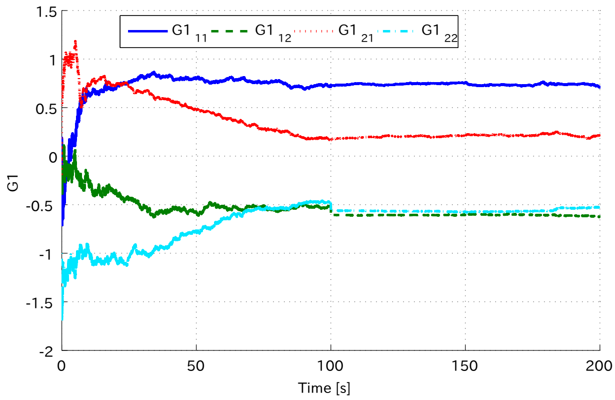


(d) ∞-norms of covariance matrices of parameter estimation errors

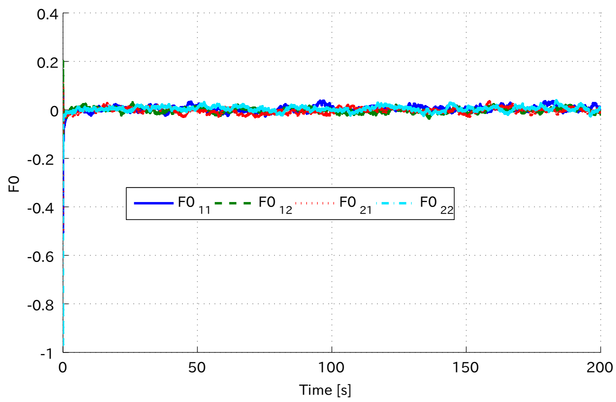
Figure 5.29: Estimates of parameters and ∞-norms of covariance matrices of parameter estimation errors using normal self tuning controller with an updated variance under Gaussian mixture noise



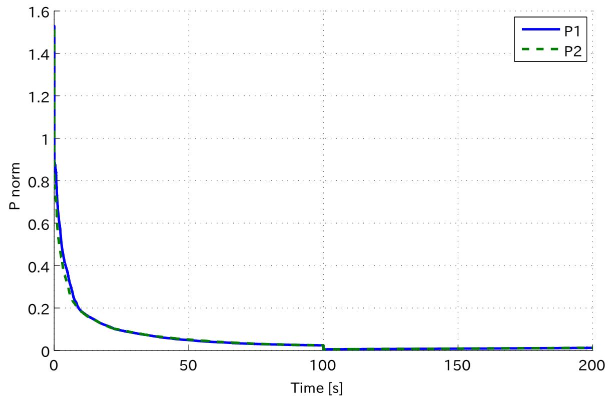
(a) estimates of G_0



(b) estimates of G_1



(c) estimates of F_0



(d) ∞-norms of covariance matrices of parameter estimation errors

Figure 5.30: Estimates of parameters and ∞-norms of covariance matrices of parameter estimation errors using robust self tuning controller under Gaussian mixture noise

5.5 Conclusion

In this chapter, we proposed RSTC under outliers. A parameter update law of a conventional STC is given by a solution of a minimization problem of estimated errors. Therefore, the proposed method estimated parameters and outliers explicitly by addition of a l_1 regression term to the minimization problem, and the estimated outliers were removed from measurement outputs in a controller. And also, it was shown that steady state errors in the proposed method with outliers are nearly equal to ones in the conventional STC without outliers. Some numerical simulations demonstrated effectiveness of the proposed method.

We will verify the proposed method by real applications as our future work. STC and RSTC in this chapter have a multi-loop structure, i.e., each constraint is mutually independent. A proposition of RSTC for more general forms is also our future work. Moreover, the proposed method deteriorates a transient response because it may estimate the change point as outliers and it cannot estimate outliers at the change point well. An improvement of the drawback is also our future work.

Chapter 6

Robust Nonlinear State Estimation and Design Method of Its Parameters

6.1 Introduction

Recently, non-contact sensors, e.g., radar measurements, GPS, ultrasonic wave sensors, image measurements, and so on, attract attentions, and these sensors are often used in control systems. However, external environments introduce outliers into these sensor signals. For example, in target tracking systems, outliers are happened due to interference of reflections from different elements of the target, and the outliers are called clutter [5, 6]. In UAV using visual feedback and UGV using GPS, temporary change of image contrast in background and radio disturbances due to some obstacles cause outliers of position data [7, 8]. These outliers deteriorate accuracy of state estimates and control performances.

In order to reduce effects of outliers, Kalman filter (KF) for non-Gaussian measurement noise has been proposed, and these methods are called robust KF (RKF) [16]–[20], [34]–[36],[49]. For linear systems, many reduction methods have been proposed. For example, in [16], Bayesian model is introduced to KF, and expectation maximization (EM) algorithm is used. In [18], use of variational Bayesian method gives approximations of joint posterior distributions of state and noise variances to realize a low computational cost. In [19], the method also learns the covariance matrix of measurement noise by iterations, and compute a Kalman gain. Especially, among these methods, RKF via l_1 regression [20] attracts many attentions. l_1 regression [21] provides some thresholds of solutions and can gives sparse solutions. The thresholds can truncate estimates of the outliers, and the method has little time delay to reduce the outliers. In addition, the method is easy to implement and compute due to a simple structure and convex optimization problem. It can be seen that parameters for standard KF can determine parameters of the RKF systematically in chapter 3. An efficient algorithm of the RKF is shown in chapter 4.

On the other hand, for nonlinear systems, extended KF (EKF) is often used to extend the aforementioned methods, but Jacobians of the nonlinear systems are required. Gaussian sum filter [34, 35] and Particle filter [36] are other famous KF for nonlinear systems and non-Gaussian measurement noise including outliers. They can approximate arbitrary distributions. However, it takes so long time to compute the algorithms, and it is unsuitable for real time applications. In other approaches, robust filter is realized by maximum a posterior (MAP) estimation using a Laplace distribution as a prior distribution [49]. However, it can consider only output equations of nonlinear systems, not dynamics of nonlinear systems.

In this chapter, we extend RKF via l_1 regression to nonlinear systems using unscented KF (UKF) [50]–[53], and propose a robust UKF (RUKF). UKF is a state estimation method for nonlinear systems without calculus of Jacobians of the systems. In addition, we also derive a design method of its parameters using a framework of a MAP estimation, and we show that the parameters can be determined systematically. Moreover, we also show that the proposed method can update a covariance matrix of a state estimation error, and can reduce orders of an optimization problem more than those of the MAP estimation using a Laplace distribution [49].

The organization of this chapter is as follows. In section 2, UKF is explained, and RUKF via l_1 regression is proposed. In section 3, we propose a new design method of RUKF. In section 4, we apply RUKF to a state estimation of a two-link manipulator under outliers, and demonstrate its effectiveness by some numerical simulations. Conclusion is given in section 5.

6.2 Robust Unscented Kalman Filter

6.2.1 Unscented Kalman Filter

We consider the following nonlinear systems:

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \boldsymbol{f}(\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{w}_k), \\ \boldsymbol{y}_k &= \boldsymbol{h}(\boldsymbol{x}_k, \boldsymbol{u}_k) + \boldsymbol{v}_k, \end{aligned} \tag{6.1}$$

where $\boldsymbol{x}_k \in \mathbb{R}^n$ is a state of the system at time k, $\boldsymbol{u}_k \in \mathbb{R}^{n_u}$ is an input, $\boldsymbol{w}_k \in \mathbb{R}^n$ is a Gaussian system noise whose mean is $\boldsymbol{0}$ and covariance matrix is $\boldsymbol{Q} \in \mathbb{R}^{n\times n}$, $\boldsymbol{y}_k \in \mathbb{R}^m$ is a measurement, and $\boldsymbol{v}_k \in \mathbb{R}^m$ is a Gaussian measurement noise whose mean is $\boldsymbol{0}$ and covariance matrix is $\boldsymbol{R} \in \mathbb{R}^{m\times m}$. It is assumed that $\boldsymbol{w}_k$ is independent of $\boldsymbol{v}_k$.

KF has a predictor-corrector structure. Let $\boldsymbol{Y}^k$ be all of the observations up to time k, and let $\hat{\boldsymbol{x}}_{i|j}$ denote a mean of $\boldsymbol{x}_i$ conditioned on $\boldsymbol{Y}^j$, i.e., $\hat{\boldsymbol{x}}_{i|j} = \mathbb{E}[\boldsymbol{x}_i|\boldsymbol{Y}^j]$. A predicted covariance matrix of a state estimation error is denoted as $\boldsymbol{P}_{k|k-1} \in \mathbb{R}^{n\times n}$. Given an estimate $\hat{\boldsymbol{x}}_{k|k}$, the predictor consists of the following equations:

$$\begin{aligned} \hat{\boldsymbol{x}}_{k+1|k} &= \mathbb{E}[\boldsymbol{f}(\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{w}_k)|\boldsymbol{Y}^k], \\ \boldsymbol{P}_{k+1|k} &= \mathbb{E}[(\boldsymbol{x}_{k+1} - \hat{\boldsymbol{x}}_{k+1|k})(\boldsymbol{x}_{k+1} - \hat{\boldsymbol{x}}_{k+1|k})^T|\boldsymbol{Y}^k]. \end{aligned} \tag{6.2}$$

The corrector of KF is given by the following equations:

$$\begin{aligned} \hat{\boldsymbol{x}}_{k+1|k+1} &= \hat{\boldsymbol{x}}_{k+1|k} + \boldsymbol{K}\boldsymbol{\nu}, \\ \boldsymbol{P}_{k+1|k+1} &= \boldsymbol{P}_{k+1|k} - \boldsymbol{K}\boldsymbol{P}_{\boldsymbol{\nu}\boldsymbol{\nu}}\boldsymbol{K}^T, \\ \boldsymbol{K}_{k+1} &= \boldsymbol{P}_{\boldsymbol{x}\boldsymbol{\nu}}\boldsymbol{P}_{\boldsymbol{\nu}\boldsymbol{\nu}}^{-1}, \end{aligned} \tag{6.3}$$

where, $\boldsymbol{\nu} = \boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k+1|k}$, $\hat{\boldsymbol{y}}_{k+1|k} = \boldsymbol{h}(\hat{\boldsymbol{x}}_{k+1|k}, \boldsymbol{u}_k)$, $\boldsymbol{P}_{\boldsymbol{\nu}\boldsymbol{\nu}} \in \mathbb{R}^{m\times m}$ is a covariance matrix of $\boldsymbol{\nu}$, and $\boldsymbol{P}_{\boldsymbol{x}\boldsymbol{\nu}} \in \mathbb{R}^{n\times m}$ is one between $\boldsymbol{x}$ and $\boldsymbol{\nu}$.

For linear systems, the covariance matrices can be calculated analytically. However, in nonlinear systems, some approximation methods are required. EKF uses a linearized system to calculate the covariance matrices, so it requires a Jacobian of the nonlinear system. In contrast, UKF approximates the covariance matrices using an unscented transformation, and it can estimate a state of the nonlinear system without calculus of the Jacobian.

One of the algorithms of UKF is as follows [51]:

Step 1. Determine sigma points:

$$\begin{aligned}
\boldsymbol{\mathcal{X}}_{0,k|k} &= \hat{\boldsymbol{x}}_{k|k}, \\
W_0 &= \frac{\kappa}{n+\kappa}, \\
\boldsymbol{\mathcal{X}}_{i,k|k} &= \hat{\boldsymbol{x}}_{k|k} + \left(\sqrt{(n+\kappa)\boldsymbol{P}_{k|k}}\right)_i, \\
W_i &= \frac{1}{2(n+\kappa)}, \\
\boldsymbol{\mathcal{X}}_{i+n,k|k} &= \hat{\boldsymbol{x}}_{k|k} - \left(\sqrt{(n+\kappa)\boldsymbol{P}_{k|k}}\right)_i, \\
W_{i+n} &= \frac{1}{2(n+\kappa)},
\end{aligned}$$

where $\kappa \in \mathbb{R}$ is a design parameter to approximate high order moments. $\left(\sqrt{(n+\kappa)\boldsymbol{P}_{k|k}}\right)_i$ is a i-th column vector of N satisfying $(n+\kappa)\boldsymbol{P}_{k|k} = \boldsymbol{N}\boldsymbol{N}^T$.

Step 2. Transform the sigma points using a state equation, and estimate a mean and covariance matrix of the state:

$$\begin{aligned}
\boldsymbol{\mathcal{X}}_{i,k+1|k} &= \boldsymbol{f}(\boldsymbol{\mathcal{X}}_{i,k|k}, \boldsymbol{u}_k), \\
\hat{\boldsymbol{x}}_{k+1|k} &= \sum_{i=0}^{2n} W_i \boldsymbol{\mathcal{X}}_{i,k+1|k}, \\
\boldsymbol{P}_{k+1|k} &= \sum_{i=0}^{2n} W_i \{\boldsymbol{\mathcal{X}}_{i,k+1|k} - \hat{\boldsymbol{x}}_{k+1|k}\}\{\boldsymbol{\mathcal{X}}_{i,k+1|k} - \hat{\boldsymbol{x}}_{k+1|k}\}^T.
\end{aligned}$$

Step 3. Predict a mean of an output, $\hat{\boldsymbol{y}}_{k+1|k}$, and its covariance matrix, $\boldsymbol{P}_{\boldsymbol{yy}}$, by using the transformed sigma points:

$$\begin{aligned}
\boldsymbol{\mathcal{Y}}_{i,k+1|k} &= \boldsymbol{h}(\boldsymbol{\mathcal{X}}_{i,k|k}, \boldsymbol{u}_k), \\
\hat{\boldsymbol{y}}_{k+1|k} &= \sum_{i=0}^{2n} W_i \boldsymbol{\mathcal{Y}}_{i,k+1|k}, \\
\boldsymbol{P}_{\boldsymbol{yy}} &= \sum_{i=0}^{2n} W_i \{\boldsymbol{\mathcal{Y}}_{i,k+1|k} - \hat{\boldsymbol{y}}_{k+1|k}\}\{\boldsymbol{\mathcal{Y}}_{i,k+1|k} - \hat{\boldsymbol{y}}_{k+1|k}\}^T.
\end{aligned}$$

Step 4. Calculate $\boldsymbol{P}_{\boldsymbol{\nu\nu}}$ and $\boldsymbol{P}_{\boldsymbol{x\nu}}$:

$$\begin{aligned}\boldsymbol{P}_{\boldsymbol{\nu\nu}} &= \boldsymbol{R} + \boldsymbol{P}_{\boldsymbol{yy}},\\ \boldsymbol{P}_{\boldsymbol{x\nu}} &= \sum_{i=0}^{2n} W_i \{\boldsymbol{\mathcal{X}}_{i,k+1|k} - \hat{\boldsymbol{x}}_{k+1|k}\}\{\boldsymbol{\mathcal{Y}}_{i,k+1|k} - \hat{\boldsymbol{y}}_{k+1|k}\}^T.\end{aligned}$$

Step 5. Update the state by using Eq. (6.3).

The above algorithm doesn't contain state noise explicitly. If the systems have state noise, we consider the following augmented systems:

$$\boldsymbol{x}_k^a := \begin{bmatrix} \boldsymbol{x}_k \\ \boldsymbol{v}_k \end{bmatrix}.$$

Moreover, a mean and covariance matrix of the augmented systems are defined by the following equations:

$$\begin{aligned}\hat{\boldsymbol{x}}_{k|k}^a &:= \begin{bmatrix} \hat{\boldsymbol{x}}_k \\ \boldsymbol{0} \end{bmatrix},\\ \boldsymbol{P}_{k|k}^a &:= \begin{bmatrix} \boldsymbol{P}_{k|k} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q} \end{bmatrix}.\end{aligned}$$

6.2.2 Robust Unscented Kalman Filter via l_1 Regression

From this subsection, systems without inputs are considered for convenience. For linear systems, it is well-known that an update law of KF can be derived from the optimization problem (2.5). Actually, the optimal solution of Eq. (2.5) gives a linear minimum variance estimation, i.e.,

$$\hat{\boldsymbol{x}}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}(\boldsymbol{y}_k - \boldsymbol{C}\hat{\boldsymbol{x}}_{k|k-1}),$$

where $\boldsymbol{K} = \boldsymbol{P}_{k|k-1}\boldsymbol{C}^T(\boldsymbol{R} + \boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T)^{-1}$.

In RKF via l_1 regression, outlier $\boldsymbol{z}_k \in \mathbb{R}^m$ is added to the output equation, and it is estimated as a solution of the optimization problem with l_1 regression. Specifically, RKF via l_1 regression is given by Eq. (2.8), and its general form is described as the following equation, again:

$$\begin{aligned}\{\hat{\boldsymbol{x}}_{k|k}, \hat{\boldsymbol{z}}_k\} = \arg\min_{\boldsymbol{x}_k, \boldsymbol{z}_k} \ & \boldsymbol{v}_k^T \boldsymbol{R}^{-1} \boldsymbol{v}_k \\ & + (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})^T \boldsymbol{P}_{k|k-1}^{-1} (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}) + \textstyle\sum_{i=1}^m \lambda_i |z_{k,i}|,\\ \text{subject to} \quad & \boldsymbol{y}_k = \boldsymbol{h}(\boldsymbol{x}_k) + \boldsymbol{v}_k + \boldsymbol{z}_k,\\ & \boldsymbol{x}_k = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k(\boldsymbol{e}_k - \boldsymbol{z}_k),\\ & \boldsymbol{e}_k = \boldsymbol{y}_k - \boldsymbol{h}(\hat{\boldsymbol{x}}_{k|k-1}),\end{aligned} \tag{6.4}$$

where λ_i $(i = 1, \cdots, m)$ are regularization parameters.

RUKF computes the Kalman gain $\boldsymbol{K}$ in Eq. (6.4) by an UKF algorithm. Note that, for nonlinear systems, the optimal solution of Eq. (2.5) cannot be calculated analytically and dose not coincide with the linear minimum variance estimation, except that output equations are linear or can be linearized. However, Eq. (2.5) evaluates a quadratic form of an estimation error and minimizes it. Therefore, it can be interpreted as minimizing some kind of expectations of variances, and it is reasonable to consider the criterion not only for linear systems, but also for nonlinear systems.

Remark 6.1 In general, output equation can be linearized exactly by some coordinate transformations [54, 55]. If a property of the outliers, which their values may usually be zero, is maintained after the transformations, the coordinate transformations can make the optimization problem (6.4) same as RKF via l_1 regression for in Chapter 3. However, if the property is not kept before and after the transformation, the optimization problem needs l_1 regularization term for transformed outliers and requires a nonlinear relation between the outliers before and after. At that time, Eq. (6.4) is still a nonlinear optimization problem. □

6.3 Design Method of Robust Unscented Kalman Filter using a Laplace Distribution

Assuming that outlier $\boldsymbol{z}_k$ is independent of $\boldsymbol{v}_k$, and it is distributed by the following Laplace distribution, whose mean is $\mathbf{0}$ and covariance matrix is $\boldsymbol{S}$,

$$p_l(\boldsymbol{z}_k) = 2^{-m/2}\det(\boldsymbol{S})^{-1/2}\exp\left[-\sqrt{2}\left|\left|\boldsymbol{S}^{-1/2}\boldsymbol{z}_k\right|\right|_1\right], \tag{6.5}$$

then the following theorem is derived.

Theorem 6.1 Assume that outlier $\boldsymbol{z}_k$ is independent of $\boldsymbol{v}_k$, and it is distributed by a Laplace distribution whose mean is $\mathbf{0}$ and covariance matrix is $\boldsymbol{S}$. The regularization parameter of RUKF, i.e., $\boldsymbol{\lambda} = [\lambda_1, \cdots, \lambda_m]^T$, is given by

$$\lambda_i = 2\sqrt{2}s_i, \tag{6.6}$$

where $\boldsymbol{s}$ be a vector consisting of diagonal elements of $\boldsymbol{S}^{-1/2}$, and s_i is a element of $\boldsymbol{s}$. □

Proof: If we compute a MAP estimation using the Laplace distribution as a prior distribution, the MAP estimator gives an optimization problem via l_1 regression. We derive RKF via l_1 regression by using the fact, again.

From the Bayes' theorem, a conditional PDF of $\boldsymbol{x}_k$ given by $\boldsymbol{y}_k$, i.e., $p(\boldsymbol{x}_k|\boldsymbol{y}_k)$, is as follows:

$$\begin{aligned} p(\boldsymbol{x}_k|\boldsymbol{y}_k) &= \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k)p(\boldsymbol{x}_k)}{p(\boldsymbol{y}_k)} \\ &\propto p(\boldsymbol{y}_k|\boldsymbol{x}_k)p(\boldsymbol{x}_k). \end{aligned} \tag{6.7}$$

In UKF, it is assumed that $\boldsymbol{x}_k$ is distributed by a normal distribution $p_n(\boldsymbol{x}_k)$, whose mean is $\hat{\boldsymbol{x}}_{k|k-1}$ and covariance matrix is $\boldsymbol{P}_{k|k-1}$, i.e.,

$$p_n(\boldsymbol{x}_k) = (2\pi)^{-n/2}\left(\det(\boldsymbol{P}_{k|k-1})\right)^{-1/2}\exp\left[-\frac{1}{2}(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})^T\boldsymbol{P}_{k|k-1}^{-1}(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})\right]. \tag{6.8}$$

In addition, an output equation is given by the following equation:

$$\boldsymbol{y}_k = \boldsymbol{h}(\boldsymbol{x}_k) + \boldsymbol{v}_k + \boldsymbol{z}_k, \tag{6.9}$$

where $\boldsymbol{v}_k$ is distributed by a normal distribution $p_n(\boldsymbol{v}_k)$, whose mean is $\boldsymbol{0}$ and covariance matrix is $\boldsymbol{R}$. From the assumption that $\boldsymbol{v}_k$ is independent of $\boldsymbol{z}_k$, the conditional PDF of $\boldsymbol{y}_k$ given by $\boldsymbol{x}_k$, i.e., $p(\boldsymbol{y}_k|\boldsymbol{x}_k)$, is written as the following convolution:

$$\begin{aligned} p(\boldsymbol{y}_k|\boldsymbol{x}_k) &= \int_{-\infty}^{\infty} p_l(\boldsymbol{h}(\boldsymbol{x}_k) - \boldsymbol{y}_k - \boldsymbol{v}_k) p_n(\boldsymbol{v}_k) d\boldsymbol{v}_k \\ &= 2^{-m/2}(2\pi)^{-m/2} \left(\det(\boldsymbol{R}) \det(\boldsymbol{S})\right)^{-1/2} \\ &\qquad \times \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\boldsymbol{v}_k^T \boldsymbol{R}^{-1} \boldsymbol{v}_k - \sqrt{2}\left|\left|\boldsymbol{S}^{-1/2}\boldsymbol{z}_k\right|\right|_1\right] d\boldsymbol{v}_k. \end{aligned}$$

Therefore,

$$\begin{aligned} p(\boldsymbol{y}_k|\boldsymbol{x}_k) p_n(\boldsymbol{x}_k) &= 2^{-m/2}(2\pi)^{-(m+n)/2} \left(\det(\boldsymbol{R}) \det(\boldsymbol{S}) \det(\boldsymbol{P}_{k|k-1})\right)^{-1/2} \\ &\quad \times \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\boldsymbol{v}_k^T \boldsymbol{R}^{-1} \boldsymbol{v}_k \right. \\ &\qquad -\frac{1}{2}(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})^T \boldsymbol{P}_{k|k-1}^{-1} (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}) \\ &\qquad \left. -\sqrt{2}\left|\left|\boldsymbol{S}^{-1/2}\boldsymbol{z}_k\right|\right|_1\right] d\boldsymbol{v}_k. \end{aligned} \tag{6.10}$$

A maximization of Eq. (6.7) is equivalent to a maximization of the integrand of Eq. (6.10) because of a positiveness of a PDF, and the maximization problem is equivalent to the following minimization problem:

$$\begin{aligned} \{\hat{\boldsymbol{x}}_{k|k}, \hat{\boldsymbol{z}}_k\} = \arg\min_{\boldsymbol{x}_k, \boldsymbol{z}_k} \ & \boldsymbol{v}_k^T \boldsymbol{R}^{-1} \boldsymbol{v}_k + (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})^T \boldsymbol{P}_{k|k-1}^{-1} (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}) \\ & +2\sqrt{2}\left|\left|\boldsymbol{S}^{-1/2}\boldsymbol{z}_k\right|\right|_1. \end{aligned} \tag{6.11}$$

On the other hand, let $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \cdots, \lambda_m)$, Eq. (6.4) can be rewriten as the following equation:

$$\{\hat{\boldsymbol{x}}_{k|k}, \hat{\boldsymbol{z}}_k\} = \arg\min_{\boldsymbol{x}_k, \boldsymbol{z}_k} \ \boldsymbol{v}_k^T \boldsymbol{R}^{-1} \boldsymbol{v}_k + (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})^T \boldsymbol{P}_{k|k-1}^{-1} (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}) + \left|\left|\boldsymbol{\Lambda}\boldsymbol{z}_k\right|\right|_1. \tag{6.12}$$

Comparing Eq. (6.11) and Eq. (6.12), we can regard Eq. (6.12) as a MAP estimator considering a Laplace distribution with only diagonal elements of the covariance matrix $\boldsymbol{S}$. Therefore, let $\boldsymbol{s}$ be a vector consisting of diagonal elements of $\boldsymbol{S}^{-1/2}$, the regularization parameter is given by Eq. (6.6). ■

Theorem 6.1 means that if the covariance matrix of the outliers $\boldsymbol{S}$ can be estimated, the regularization parameter is determined automatically. RUKF can estimate outlier $\hat{\boldsymbol{z}}_k$, so use of the estimates can calculate a sample covariance matrix of the outlier. On the other hand, we use a covariance matrix $\boldsymbol{P}_{\boldsymbol{\nu}\boldsymbol{\nu}}$ as that of outliers, i.e., $\boldsymbol{S} = \boldsymbol{P}_{\boldsymbol{\nu}\boldsymbol{\nu}}$. In the case using $\boldsymbol{P}_{\boldsymbol{\nu}\boldsymbol{\nu}}$ as the covariance matrix of outliers, the following lemma is satisfied.

Lemma 6.1 Assume that $\hat{\boldsymbol{z}}_k$ can be estimated satisfying $\mathbb{E}[\boldsymbol{z}_k\boldsymbol{z}_k^T] = \mathbb{E}[\hat{\boldsymbol{z}}_k\hat{\boldsymbol{z}}_k^T]$, and $\boldsymbol{z}_k$ and $\hat{\boldsymbol{z}}_k$ are mutually independent. If $\boldsymbol{S}$ is given by

$$\begin{aligned} \boldsymbol{S} &:= \mathbb{E}[\boldsymbol{z}_k\boldsymbol{z}_k^T] \\ &= \gamma \boldsymbol{P}_{\boldsymbol{\nu\nu}}, \end{aligned} \tag{6.13}$$

where γ is a positive real number, then the following equation is satisfied:

$$\mathbb{E}[(\boldsymbol{\nu}_k - \hat{\boldsymbol{z}}_k)(\boldsymbol{\nu}_k - \hat{\boldsymbol{z}}_k)^T] = \mathcal{O}(\boldsymbol{P}_{\boldsymbol{\nu\nu}}). \tag{6.14}$$

This means that performances of RUKF come close ones of UKF without outliers. □

Proof: The assumptions result in

$$\begin{aligned} \mathbb{E}[(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k)(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k)^T] &= \mathbb{E}[\boldsymbol{z}_k\boldsymbol{z}_k^T] - 2\mathbb{E}[\boldsymbol{z}_k\hat{\boldsymbol{z}}_k^T] + \mathbb{E}[\hat{\boldsymbol{z}}_k\hat{\boldsymbol{z}}_k^T] \\ &= 2\mathbb{E}[\boldsymbol{z}_k\boldsymbol{z}_k^T] \\ &= 2\gamma \boldsymbol{P}_{\boldsymbol{\nu\nu}}. \end{aligned}$$

Eq. (6.9) yields the following equation:

$$\boldsymbol{\nu}_k - \hat{\boldsymbol{z}}_k = \boldsymbol{h}(\boldsymbol{x}_k) - \boldsymbol{h}(\hat{\boldsymbol{x}}_k) + \boldsymbol{v}_k + \boldsymbol{z}_k - \hat{\boldsymbol{z}}_k.$$

Each stochastic variable is independent of the other variables, so the following equation is satisfied:

$$\begin{aligned} \mathbb{E}[(\boldsymbol{\nu}_k - \hat{\boldsymbol{z}}_k)(\boldsymbol{\nu}_k - \hat{\boldsymbol{z}}_k)^T] &= \boldsymbol{P}_{\boldsymbol{yy}} + \boldsymbol{R} + \mathbb{E}[(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k)(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k)^T] \\ &= (1 + 2\gamma)\boldsymbol{P}_{\boldsymbol{\nu\nu}}. \end{aligned} \tag{6.15}$$

Therefore, $\mathbb{E}[(\boldsymbol{\nu}_k - \hat{\boldsymbol{z}}_k)(\boldsymbol{\nu}_k - \hat{\boldsymbol{z}}_k)^T]$ is given by order of $\boldsymbol{P}_{\boldsymbol{\nu\nu}}$. ■

Remark 6.2 In the proposed method, the covariance matrix of the Laplace distribution is estimated using sigma points, and the regularization parameter is calculated by Eq. (6.6). This means that a design of a standard UKF determines the parameter of RUKF systematically. □

Remark 6.3 The proposed method realizes a robust filter combining UKF and l_1 minimization. On the other hand, a robust filter can be realized without UKF by solving Eq. (6.11) directly. In fact, in [49], a minimization problem derived by a MAP estimation is solved directly. However, use of UKF can update the covariance matrix of Eq. (6.11). This means that the proposed method is adaptive to uncertainty of parameters more than the method in [49]. □

Remark 6.4 The proposed method assumes that an update law is obtained as $\boldsymbol{x}_k = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k(\boldsymbol{e}_k - \boldsymbol{z}_k)$, and uses it as a constraint of $\boldsymbol{x}_k$. Therefore, Eq. (6.4) becomes a minimization problem about only $\boldsymbol{z}_k$. This means that the minimization problem of the proposed method has lower order dimension of variables than solving Eq. (6.11) directly. □

6.4 Simulation

6.4.1 Conditions

We consider a state estimation problem of a two-link manipulator shown in Fig. 6.1, where m_1 and l_1 are mass and length of link 1, and m_2 and l_2 are those of link 2. θ_1 and θ_2 are relative angular positions of each link, and directions of θ_1 and θ_2 are counterclockwise and clockwise, respectively. Let $\boldsymbol{q} = [\theta_1, \theta_2]^T$ be a generalized coordinate. Assuming that a center of mass is located in a top of the link and moment of inertia and viscosity are zero, a dynamics of the manipulator is written as

$$\boldsymbol{M}(\boldsymbol{q})\ddot{\boldsymbol{q}} + \boldsymbol{C}(\boldsymbol{q}, \dot{\boldsymbol{q}}) + \boldsymbol{G}(\boldsymbol{q}) = \boldsymbol{\tau}, \tag{6.16}$$

where

$$\begin{aligned}
\boldsymbol{M}(\boldsymbol{q}) &= \begin{bmatrix} \alpha + \gamma + 2\beta \sin\theta_2 & \gamma + \beta \sin\theta_2 \\ \gamma + \beta \sin\theta_2 & \gamma \end{bmatrix}, \\
\boldsymbol{C}(\boldsymbol{q}, \dot{\boldsymbol{q}}) &= \begin{bmatrix} \beta(2\dot{\theta}_1 + \dot{\theta}_2)\dot{\theta}_2 \cos\theta_2 & -\beta\dot{\theta}_1^2 \cos\theta_2 \end{bmatrix}^T, \\
\boldsymbol{G}(\boldsymbol{q}) &= \begin{bmatrix} (m_1 + m_2) g l_1 \sin\theta_1 + m_2 g l_2 \sin(\theta_1 + \theta_2) \\ m_2 g l_2 \sin(\theta_1 + \theta_2) \end{bmatrix},
\end{aligned}$$

where

$$\begin{aligned}
\alpha &= m_1 l_1^2 + m_2 l_2^2, \\
\beta &= m_2 l_1 l_2, \\
\gamma &= m_2 l_2^2.
\end{aligned}$$

An output of the system is y coordinate of the top of the second link, i.e.,

$$y = -l_1 \cos\theta_1 - l_2 \cos(\theta_1 + \theta_2). \tag{6.17}$$

Parameters of the model are $l_1 = 1.0$, $l_2 = 2.0$, $m_1 = 0.5$, and $m_2 = 0.5$. An initial value is $\boldsymbol{x}_0 = [1.4, 0, 0, 0]^T$, and inputs to each link are zero.

We consider Gaussian noise as a nominal measurement noise. Parameters of the nominal measurement noise is as follows:

$$\begin{aligned}
\mathbb{E}[v_k] &= 0, \\
\mathbb{E}[v_k^2] &= (2.5 \times 10^{-1})^2.
\end{aligned}$$

In addition, for outliers, two cases of distributions are considered, i.e., Cauchy and Gaussian mixture distributions, $p_c(\boldsymbol{x})$ and $p_g(\boldsymbol{x})$. In the case of Cauchy distribution as a model of outliers, outliers distributed by Cauchy distribution are added to the nominal noise. Parameters

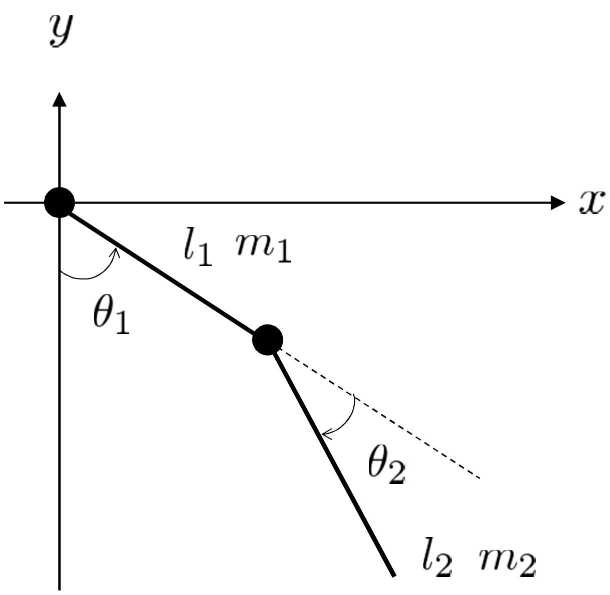


Figure 6.1: Model of a two-link manipulator. m_1 and l_1 are mass and length of link 1, and m_2 and l_2 are those of link 2. θ_1 and θ_2 are relative angular positions of each link, and directions of θ_1 and θ_2 are counterclockwise and clockwise, respectively.

of these distributions are as follows:

$$\begin{aligned} x_0 &= 0, \\ \delta &= 1.0 \times 10^{-2}, \\ \mu_1 &= \mu_2 = 0, \\ p &= 0.1, \\ \Sigma_1 &= 2.5 \times 10^{-1}, \\ \Sigma_2 &= 10\Sigma_1. \end{aligned}$$

It is assumed that a covariance matrix of the nominal measurement noise is known. κ is determined satisfying $n + \kappa = 3$, and the other parameters of UKF and RUKF are as follows:

$$\begin{aligned} \boldsymbol{P}_{0|0} &= 1 \times 10^{-5}\boldsymbol{I}, \\ \boldsymbol{Q} &= 1 \times 10^{-8}\boldsymbol{I}. \end{aligned}$$

6.4.2 Results

Fig. 6.2 (a) and (b) show measurements under Cauchy and Gaussian mixture noise, respectively.

Fig. 6.3 – Fig. 6.5 show estimates of the two-link manipulator with Cauchy noise using UKF, solving Eq. (6.11) directly, and RUKF, respectively. These graphs show that UKF and a direct solution of Eq. (6.11) have larger estimation errors than RUKF under the influence of the outliers. If the manipulator operates around the origin, the direct solution

of Eq. (6.11) can estimate the state well. However, it cannot estimate the state in a large operating range like these simulations.

Fig. 6.6 – Fig. 6.8 show estimates of the two-link manipulator with Gaussian mixture noise using UKF, a direct solution of Eq. (6.11), and RUKF, respectively. Also in Gaussian mixture noise, it can be seen that estimation errors using RUKF are smaller than those using UKF and the direct solution.

Table 6.1 and Table 6.2 show root mean squared errors (RMSEs) of each estimate. The errors are changed at each simulation, so these are averaged values of 10 times simulations. In the tables, a result using RUKF designed by a heuristic method is also shown. As the heuristic design method, we search regularization parameters which minimizes a summation of RMSEs of each estimate. We set $\lambda = 8.87$ in the case of Cauchy noise, and $\lambda = 10.1$ in the case of Gaussian mixture noise.

These tables show that estimation errors using RUKF are smaller than those using UKF and the direct solution of Eq. (6.11) for the both types of outliers. These results also show that, in RUKF, the proposed design method can reduce more errors than the heuristic design method.

6.5 Conclusions

In this chapter, we proposed RUKF via l_1 regression and a new design method of its regularization parameters. Regularization parameters of RUKF are determined by statistics of Gaussian measurement noise. This means that the proposed design method provides the parameters with physical meanings, and we can design the parameters systematically. We applied RUKF to a state estimation of a two-link manipulator under outliers. Effectiveness was demonstrated by some numerical simulations.

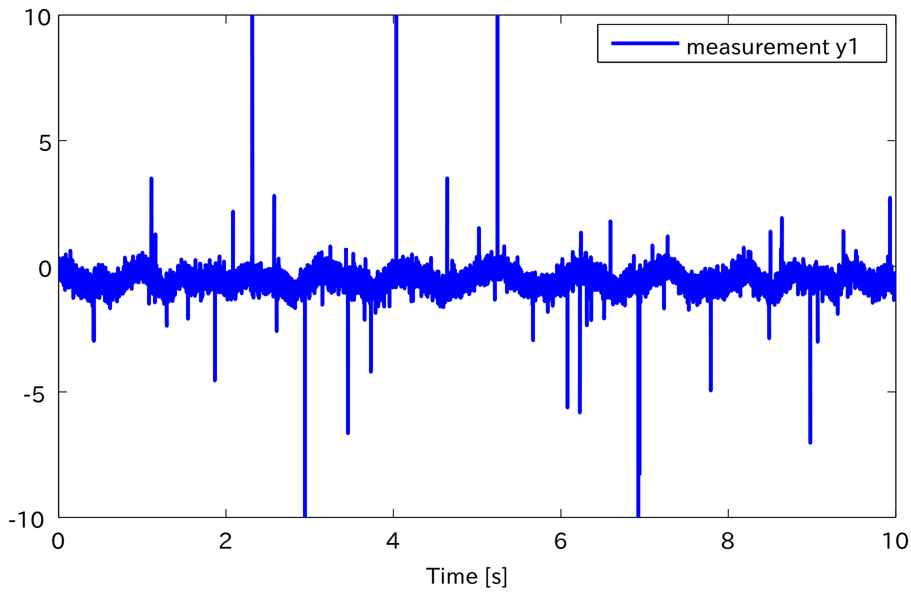
We will verify the proposed method by real applications as our future works.

Table 6.1: Root mean squared errors of estimates under Cauchy noise.

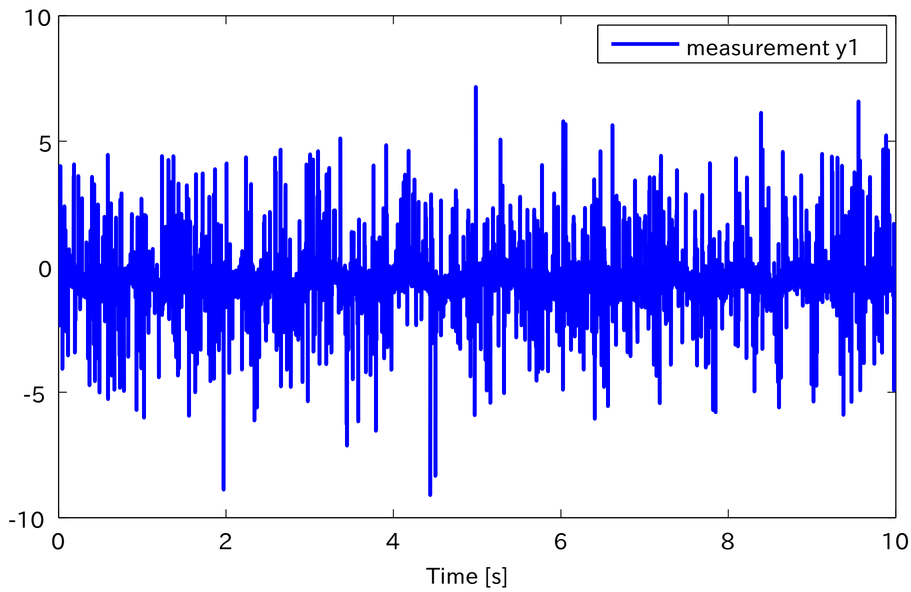
	θ_1	θ_2	$\dot{\theta}_1$	$\dot{\theta}_2$
UKF	4.23×10^{-1}	6.40×10^{-1}	2.35	3.98
direct solution of Eq.(6.11)	1.16	1.62	5.02	1.76
RUKF with a heuristic design method	2.09×10^{-1}	3.88×10^{-1}	1.44	2.03
RUKF with a proposed design method	1.37×10^{-1}	2.23×10^{-1}	1.05	2.41

Table 6.2: Root mean squared errors of estimates under Gaussian mixture noise.

	θ_1	θ_2	$\dot{\theta}_1$	$\dot{\theta}_2$
UKF	3.57×10^{-1}	6.15×10^{-1}	2.20	3.91
direct solution of Eq.(6.11)	1.11	1.74	5.23	7.87
RUKF with a heuristic design method	2.79×10^{-1}	4.34×10^{-1}	1.75	3.19
RUKF with a proposed design method	1.35×10^{-1}	2.27×10^{-1}	1.02	1.88

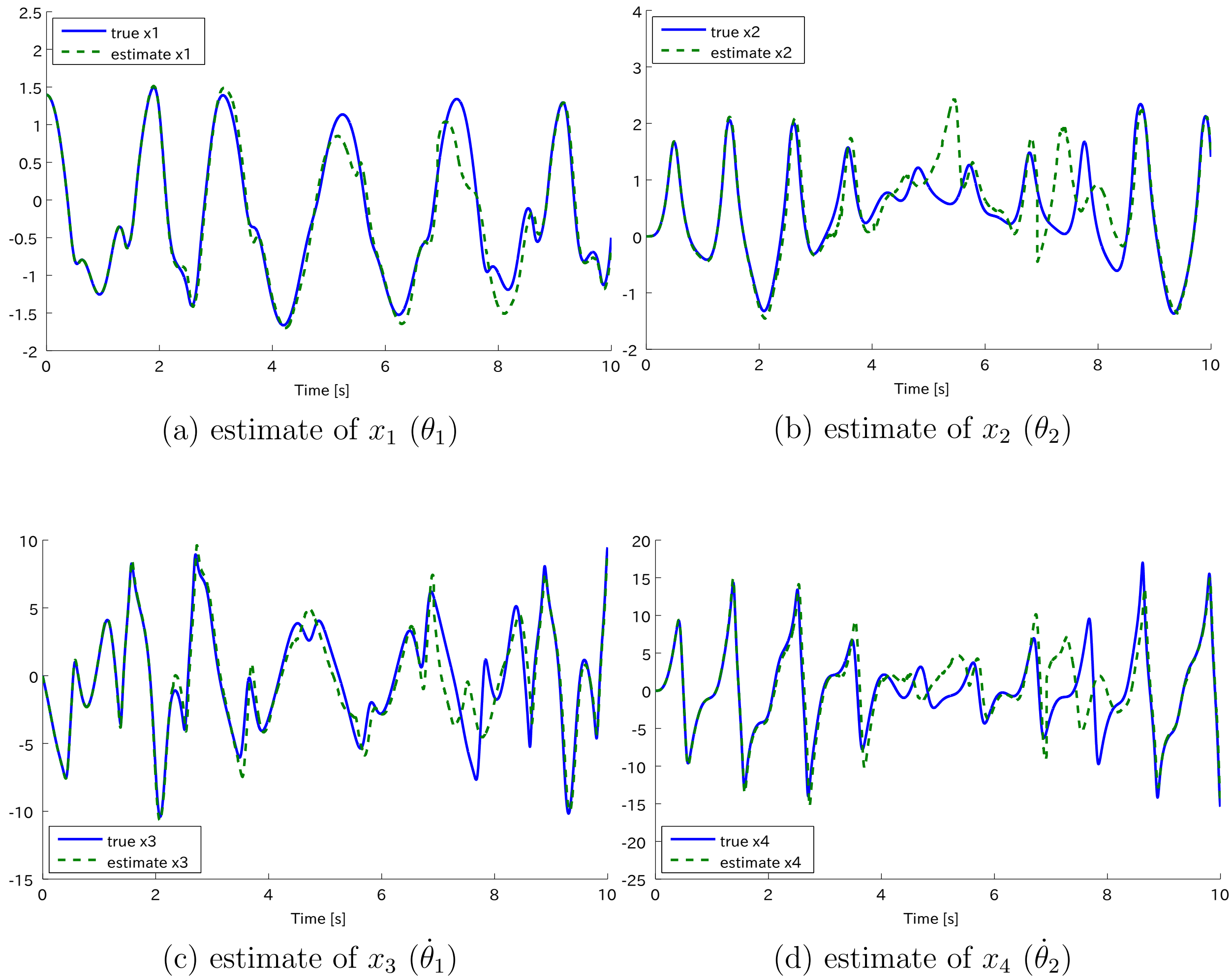


(a) under Cauchy noise



(b) under Gaussian mixture noise

Figure 6.2: Measurement of output. (a) and (b) contain Cauchy and Gaussian mixture noise, respectively. Parameters of Cauchy distribution are $x_0 = 0$ and $\delta = 1.0 \times 10^{-1}$. Parameters of Gaussian mixture distribution are $\mu_1 = \mu_2 = 0$, $p = 0.1$, $\Sigma_1 = 2.5 \times 10^{-1} I$, and $\Sigma_2 = 10\Sigma_1$.



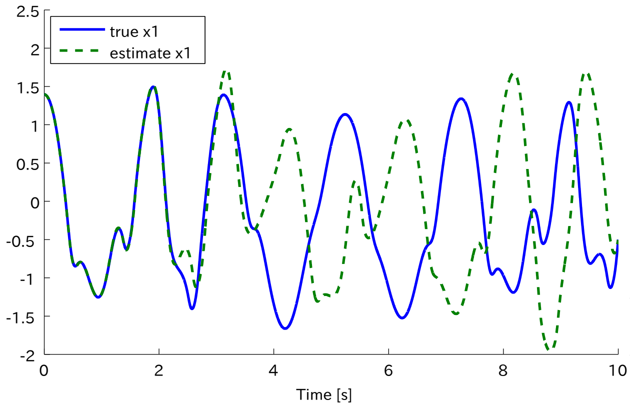
(a) estimate of x_1 (θ_1)

(b) estimate of x_2 (θ_2)

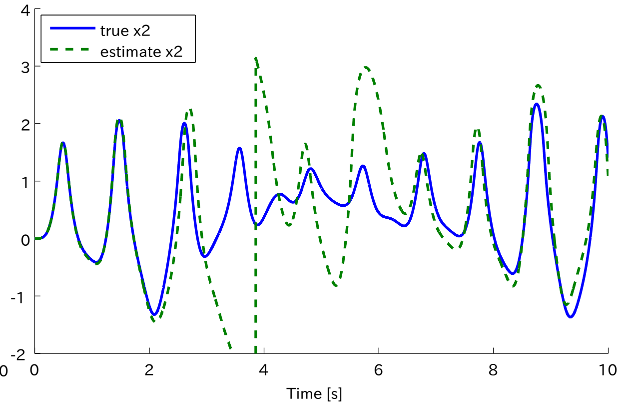
(c) estimate of x_3 ($\dot{\theta}_1$)

(d) estimate of x_4 ($\dot{\theta}_2$)

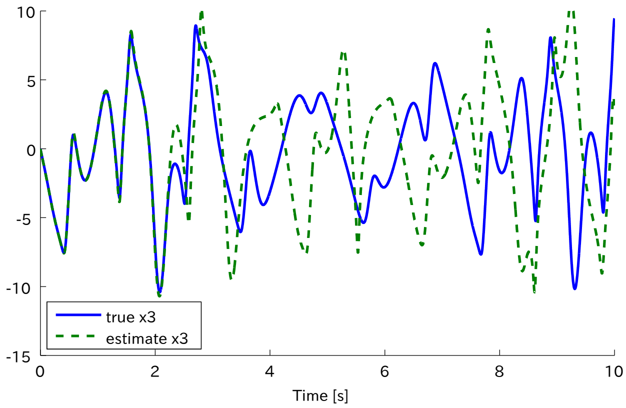
Figure 6.3: Estimates using UKF under Cauchy noise. The solid lines are true signals and dashed lines are estimates.



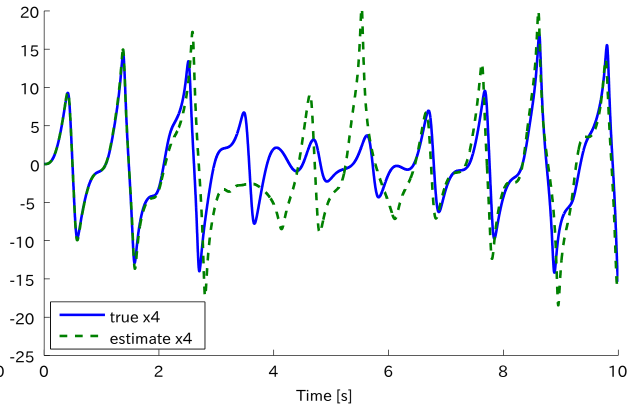
(a) estimate of x_1 (θ_1)



(b) estimate of x_2 (θ_2)

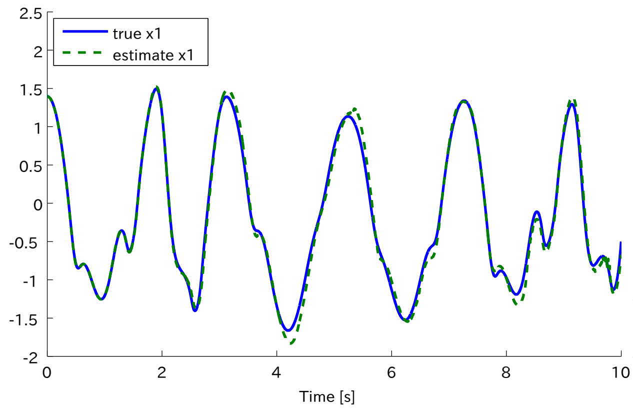


(c) estimate of x_3 ($\dot{\theta}_1$)

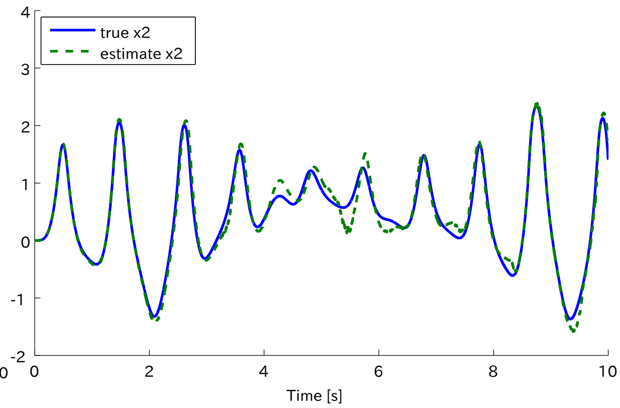


(d) estimate of x_4 ($\dot{\theta}_2$)

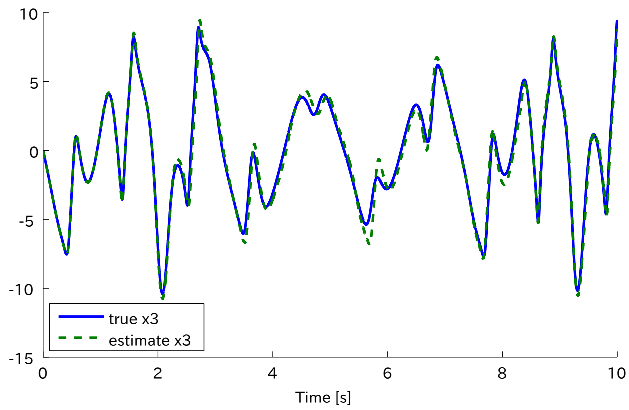
Figure 6.4: Estimates solving Eq. (6.11) directly under Cauchy noise. The solid lines are true signals and dashed lines are estimates.



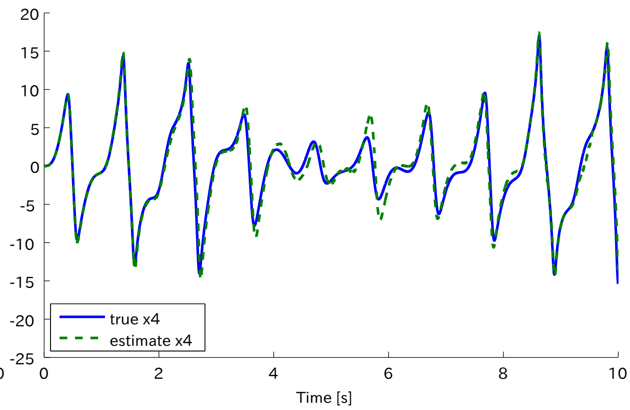
(a) estimate of x_1 (θ_1)



(b) estimate of x_2 (θ_2)

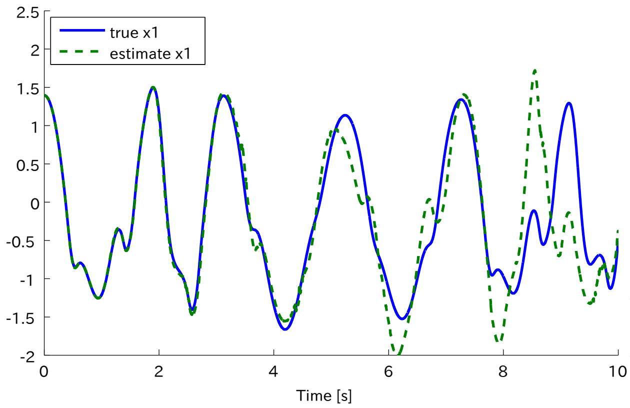


(c) estimate of x_3 ($\dot{\theta}_1$)

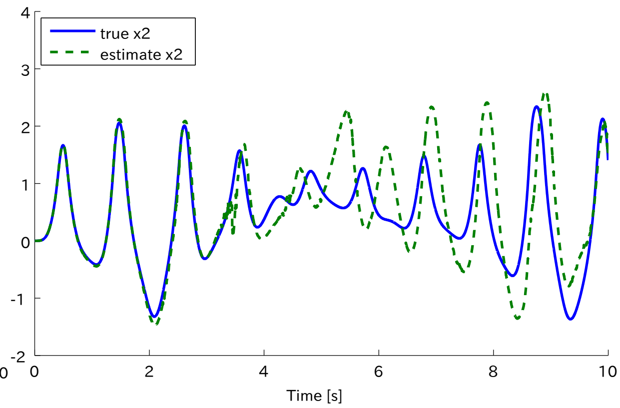


(d) estimate of x_4 ($\dot{\theta}_2$)

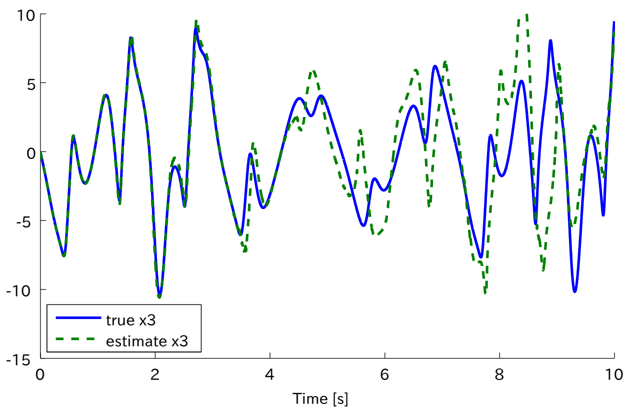
Figure 6.5: Estimates using RUKF under Cauchy noise. The solid lines are true signals and dashed lines are estimates.



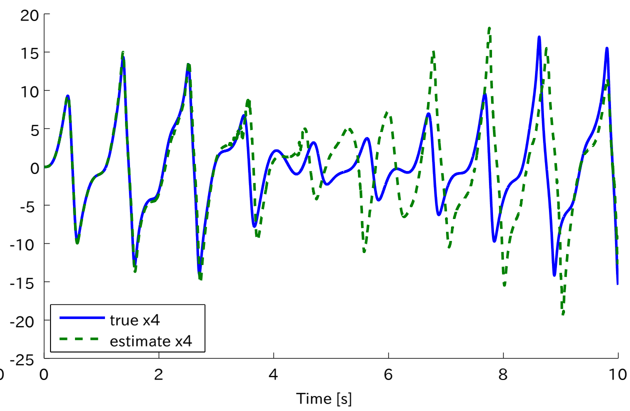
(a) estimate of x_1 (θ_1)



(b) estimate of x_2 (θ_2)

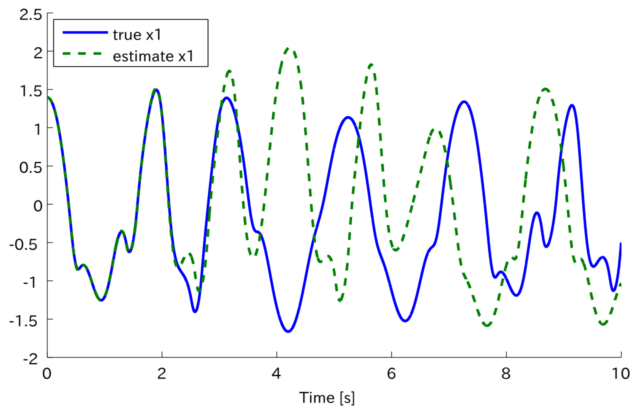


(c) estimate of x_3 ($\dot{\theta}_1$)

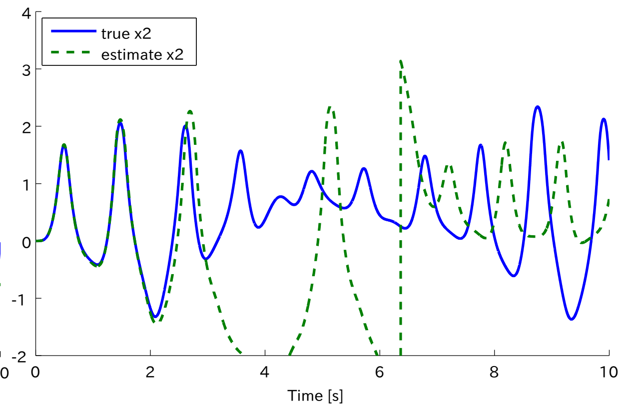


(d) estimate of x_4 ($\dot{\theta}_2$)

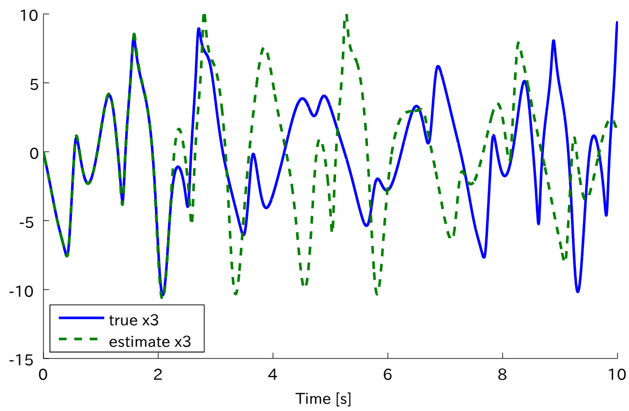
Figure 6.6: Estimates using UKF under Gaussian mixture noise. The solid lines are true signals and dashed lines are estimates.



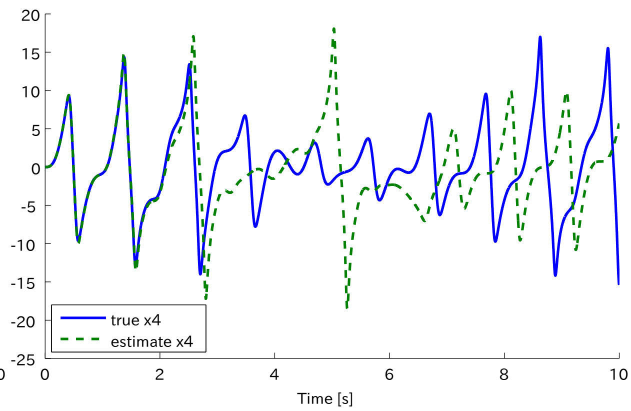
(a) estimate of x_1 (θ_1)



(b) estimate of x_2 (θ_2)

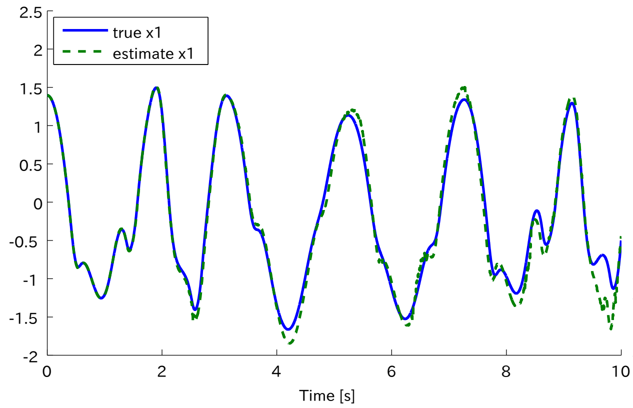


(c) estimate of x_3 ($\dot{\theta}_1$)

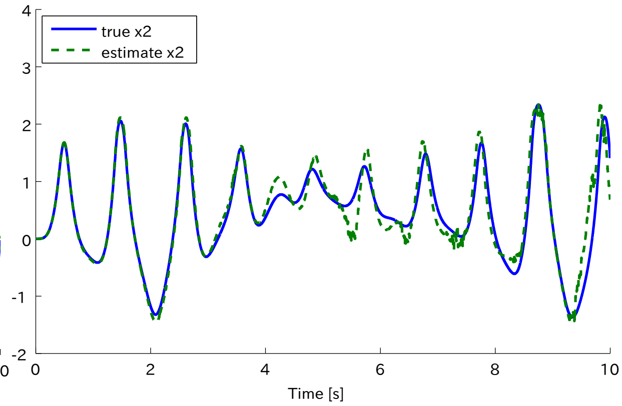


(d) estimate of x_4 ($\dot{\theta}_2$)

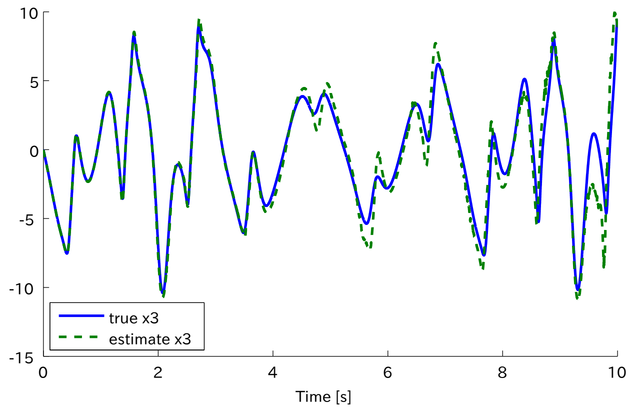
Figure 6.7: Estimates solving Eq. (6.11) directly under Gaussian mixture noise. The solid lines are true signals and dashed lines are estimates.



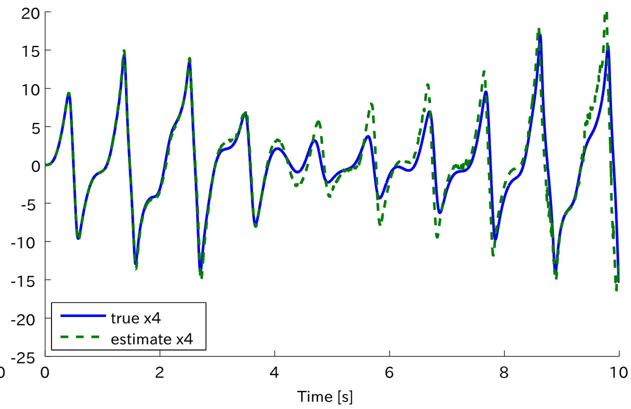
(a) estimate of x_1 (θ_1)



(b) estimate of x_2 (θ_2)



(c) estimate of x_3 ($\dot{\theta}_1$)



(d) estimate of x_4 ($\dot{\theta}_2$)

Figure 6.8: Estimates using RUKF under Gaussian mixture noise. The solid lines are true signals and dashed lines are estimates.

Chapter 7

Conclusions

7.1 Summary

Outliers are a kind of non-Gaussian measurement noise generated by heavier tailed distributions than a normal distribution. Hence, abnormal values, which are distant so much from mean values of distributions, are unusually occurred in a time domain. They are happened in many applications, e.g., automobiles, industrial robots, medical machines, and so on, and they provide negative effects on various applications. Examples of the applications are target tracking systems, unmanned vehicles, visual feedback systems, and so on. Such systems have non-contact sensors, e.g., radar measurements, GPS, ultrasonic wave sensors, image measurements, and so on. The sensors are influenced by external environments and contaminated by outliers. Moreover, networked control systems have intermittent observations due to instability and variable delay of its network. System identifications also deal with missing observations and irregular sampling data. They can be regarded as a kind of outliers. Generally, in control designs under outliers, there are no methods considering outliers explicitly. The most basic strategy to construct control systems under outliers separates control designs and estimation problems, and the estimation problems deal with the outliers. It is required not only to reduce effects of the outliers, but also to analyze performances of the reduction method. Therefore, it is required to construct reduction methods of effects of the outliers and to analyze performances of the methods. For this reasons, this dissertation proposed robust estimation methods and a control strategy under outliers based on RKF via l_1 regression. In addition, we analyzed performances of the proposed methods and demonstrated effectiveness of the proposed methods by some numerical simulations.

RKF via l_1 regression is one of the most attractive methods because of an easy structure and implementation. Additionally, the RKF truncates outliers by some thresholds and has less delay than the other RKF. However, regularization parameters of the RKF needed to be tuned by heuristic methods. In chapter 3, we proposed a new design method of the RKF. Regularization parameters of the RKF are determined by statistics of Gaussian noise. Both primal and dual problems can derive a condition of the proposed parameters. This means that the proposed design method provides the parameters with physical meanings, and we can design the parameters systematically. We analyzed performances of the RKF with the proposed design method. It was shown that a covariance matrix of an innovation of the RKF is bounded by that of normal KF without outliers. The covariance matrix of the innovation of the RKF comes close to an ideal one under outliers. We applied the

RKF with the proposed design method to a target tracking systems with clutters and a control problem of a two-wheeled vehicle under outliers. Effectiveness was demonstrated by some numerical simulations. Especially, the simulations showed that the multi-parameters were automatically determined to be no conservative under measurement noise with different variances.

Since RKF via l_1 regression consists of a convex optimization problem, a computation is more effective than the other robust estimations. However, the RKF requires some iterative algorithms to solve the optimization, so a convergence rate and accuracy of the solutions of the RKF depend on conditions of the iterations. In chapter 4, we proposed a fast algorithm of RKF via l_1 regression. The proposed algorithm approximates the optimal solution by using its upper and lower bounds, and the approximated solution is given by a closed form. Moreover, it was shown that the proposed algorithm had almost same performances as KF without outliers under some conditions. Some numerical simulations demonstrated a comparison of performances using CVX, CVXGEN, FISTA, and the proposed algorithm. The proposed algorithm gave smaller RMSEs than the other algorithms. Moreover, using CVXGEN, FISTA, and the proposed algorithm, computation times are about 1/1000 times less than one using CVX. A compiled version of the proposed algorithm is more accelerated, and a computation time of the compiled version comes close to that of normal KF.

In chapter 5, in order to construct a robust controller under outliers, we applied an idea of RKF via l_1 regression to self-tuning controller (STC), and we proposed robust STC (RSTC) under outliers. A parameter update law of the conventional STC can be written as a recursive least square (RLS) method, and RLS can be given by a solution of a minimization problem of estimated errors. Therefore, the proposed method estimated parameters and outliers explicitly by addition of a l_1 regression term to the minimization problem, and the estimated outliers were removed from measurement outputs in a controller. The proposed method is solved in a closed form due to a l_1 optimization problem with a single variable, so the algorithm is very efficient. And also, it was shown that steady state errors in the proposed method with outliers are nearly equal to ones in the conventional STC without outliers. A numerical simulation demonstrated that the proposed RSTC can remove effects of the outliers and reduce control errors more than STC.

RKF via l_1 regression is a method only for linear systems. In chapter 6, we extend the RKF to non-linear systems using unscented KF (UKF) and we proposed robust UKF (RUKF) via l_1 regression. And also, we proposed a new design method of its regularization parameters. It was shown that regularization parameters of RUKF are determined by a covariance matrix of Laplace distributions. Similarly to linear systems, regularization parameters of RUKF are determined by statistics of Gaussian measurement noise in this dissertation. Therefore, the proposed design method provides the parameters with physical meanings, and we can design the parameters systematically. Moreover, the regularization parameters make performances of RUKF come close ones of UKF without outliers. We applied RUKF to a state estimation of a two-link manipulator under outliers. Some numerical simulations demonstrated that RUKF had smaller estimation errors than the other algorithms under outliers. Since RUKF is based on UKF and l_1 optimization problem, it can be computed more efficiently than Gaussian sum filter and particle filter.

7.2 Future Works

In our contributions, effectiveness is shown only by some numerical simulations. We will verify the proposed methods by real applications as our future works. As we mentioned before, there are many control systems under outliers. We will apply the proposed method to the systems.

In larger dimension of RKF via l_1 regression, or in some conditions of covariance matrix of Gaussian noise, the proposed algorithm of the RKF sometimes could not satisfy the condition (4.19), and performances of the RKF were deteriorated. A proposition of efficient algorithms for the large scale problems is one of our future works.

Performances of the proposed RKF and RUKF depend on the covariance matrix $\boldsymbol{R}$. The proposed methods can estimate outlier $\boldsymbol{z}_k$. Therefore, we can estimate the covarinace matrix $\boldsymbol{R}$ using the estimated outlier like Eq. (5.36), and we can construct an adaptive RKF and RUKF. However, stability of the method cannot be guaranteed. We have tried the methods and the methods sometimes made unstable. A proposition of a method with unknown covariance matrix, whose stability is guaranteed by a theory, is also one of the our future works.

In RSTC, the proposed method has a multi-loop structure, i.e., each constraint is mutually independent. Therefore, the proposed method can be used only in the case that number of inputs is equal to one of outputs. A proposition of RSTC for more general forms is our future work. Moreover, the proposed method deteriorates a transient response because it may estimate the change point as outliers and it cannot estimate outliers at the change point well. An improvement of the drawback is also our future work.

Finally, we want to expand ideas of this dissertation to more general l_1 optimization, i.e., least absolute shrinkage and selection operator (LASSO), in future.

Appendix A

Multivariate Laplace Distribution

A.1 Derivation

First, a standard multivariate Laplace distribution is derived.

Define $\boldsymbol{X}$ and $\boldsymbol{x}$ as follows:

$$\begin{aligned}\boldsymbol{X} &:= \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}, \\ \boldsymbol{x} &:= \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in X,\end{aligned}$$

where $X_1, X_2, \cdots, X_n$ are independent stochastic variables and distributed by the following Laplace distribution whose mean is 0 and variance is 1:

$$p_l(x_i) = \frac{\sqrt{2}}{2} \exp\left(\sqrt{2}|x_i|\right).$$

Therefore, a marginal probability density function (marginal PDF) of $\boldsymbol{X}$ is given by

$$p_l(\boldsymbol{x}) = \prod_{i=1}^{n} p(x_i) = 2^{-\frac{n}{2}} \exp\left(\sqrt{2}||\boldsymbol{x}||_1\right).$$

Secondly, consider the coordinate transformation of $\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{\mu}$, where $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is a non-singular transform matrix, and $\boldsymbol{Y}$, $\boldsymbol{y}$, $\boldsymbol{\mu}$ are defined as follows:

$$\begin{aligned}\boldsymbol{Y} &:= \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \\ \boldsymbol{y} &:= \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \in \boldsymbol{Y}, \\ \boldsymbol{\mu} &:= \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix}.\end{aligned}$$

From $\boldsymbol{X} = \boldsymbol{A}^{-1}(\boldsymbol{Y} - \boldsymbol{\mu})$,

$$dx_1 dx_2 \cdots dx_n = \det(\boldsymbol{A}^{-1}) dy_1 dy_2 \cdots dy_n.$$

Therefore,

$$\begin{aligned} p(\boldsymbol{x}) dx_1 dx_2 \cdots dx_n &= 2^{-\frac{n}{2}} \exp\left(\sqrt{2} ||\boldsymbol{x}||_1\right) \det(\boldsymbol{A}^{-1}) dy_1 dy_2 \cdots dy_n \\ &= 2^{-\frac{n}{2}} \exp\left(\sqrt{2} ||\boldsymbol{A}^{-1}(\boldsymbol{y} - \boldsymbol{\mu})||_1\right) \det(\boldsymbol{A}^{-1}) dy_1 dy_2 \cdots dy_n \\ &= 2^{-\frac{n}{2}} (\det(\boldsymbol{A}))^{-1} \exp\left(\sqrt{2} ||\boldsymbol{A}^{-1}(\boldsymbol{y} - \boldsymbol{\mu})||_1\right) dy_1 dy_2 \cdots dy_n. \end{aligned}$$

Let $\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{A}^T$, then $\boldsymbol{A} = \boldsymbol{\Sigma}^{1/2}$, and

$$\det(\boldsymbol{\Sigma}) = \det(\boldsymbol{A}) \cdot \det(\boldsymbol{A}^T) \quad \Leftrightarrow \quad \det(\boldsymbol{A}) = (\det(\boldsymbol{\Sigma}))^{1/2}.$$

Therefore, PDF of $\boldsymbol{Y}$ is given by

$$p(\boldsymbol{y}) = 2^{-\frac{n}{2}} (\det(\boldsymbol{\Sigma}))^{-\frac{1}{2}} \exp\left(\sqrt{2} ||\boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{y} - \boldsymbol{\mu})||_1\right).$$

This is one of PDFs of multivariate Laplace distributions.

A.2 First and Second Moments

From the relation $\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{\mu}$, first and second moments of $\boldsymbol{Y}$ are given by

$$\mathbb{E}[\boldsymbol{Y}] = \mathbb{E}[\boldsymbol{A}\boldsymbol{X} + \boldsymbol{\mu}] = \boldsymbol{\mu},$$

$$\begin{aligned} \mathbb{E}[(\boldsymbol{Y} - \mathbb{E}[\boldsymbol{Y}])(\boldsymbol{Y} - \mathbb{E}[\boldsymbol{Y}])^T] &= \mathbb{E}[\boldsymbol{A}\boldsymbol{X}(\boldsymbol{A}\boldsymbol{X})^T] \\ &= \boldsymbol{A}\boldsymbol{A}^T \\ &= \boldsymbol{\Sigma}. \end{aligned}$$

This means that $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are calculated by a sample mean and covariance.

Appendix B

Derivation of an Update Law of Kalman Filter via an optimization problem

In this chapter, Let J be a criterion in Eq. (2.5), i.e.,

$$J = \boldsymbol{v}_k^T \boldsymbol{R}^{-1} \boldsymbol{v}_k + \left(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}\right)^T \boldsymbol{P}_{k|k-1}^{-1} \left(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}\right).$$

A necessary condition of an optimality gives the following equation:

$$\left.\frac{\partial J}{\partial \boldsymbol{x}_k}\right|_{\boldsymbol{x}_k = \hat{\boldsymbol{x}}_{k|k}} = -2\boldsymbol{C}^T \boldsymbol{R}^{-1} \left(\boldsymbol{y}_k - \boldsymbol{C}\hat{\boldsymbol{x}}_{k|k}\right) + 2\boldsymbol{P}_{k|k-1}^{-1} \left(\hat{\boldsymbol{x}}_{k|k} - \hat{\boldsymbol{x}}_{k|k-1}\right) = \boldsymbol{0}.$$

Therefore,

$$\begin{aligned}
& \boldsymbol{C}^T \boldsymbol{R}^{-1} \left(\boldsymbol{C}\hat{\boldsymbol{x}}_{k|k} - \boldsymbol{y}_k\right) + \boldsymbol{P}_{k|k-1}^{-1} \left(\hat{\boldsymbol{x}}_{k|k} - \hat{\boldsymbol{x}}_{k|k-1}\right) = \boldsymbol{0}. \\
\Leftrightarrow\ & \left(\boldsymbol{P}_{k|k-1}^{-1} + \boldsymbol{C}^T \boldsymbol{R}^{-1} \boldsymbol{C}\right) \hat{\boldsymbol{x}}_{k|k} = \boldsymbol{P}_{k|k-1}^{-1} \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{C}^T \boldsymbol{R}^{-1} \boldsymbol{y}_k. \\
\Leftrightarrow\ & \hat{\boldsymbol{x}}_{k|k} = \left(\boldsymbol{P}_{k|k-1}^{-1} + \boldsymbol{C}^T \boldsymbol{R}^{-1} \boldsymbol{C}\right)^{-1} \boldsymbol{P}_{k|k-1}^{-1} \hat{\boldsymbol{x}}_{k|k-1} \\
& \qquad + \left(\boldsymbol{P}_{k|k-1}^{-1} + \boldsymbol{C}^T \boldsymbol{R}^{-1} \boldsymbol{C}\right)^{-1} \boldsymbol{P}_{k|k-1}^{-1} \boldsymbol{C}^T \boldsymbol{R}^{-1} \boldsymbol{y}_k.
\end{aligned} \tag{B.1}$$

An inverse matrix lemma gives the following equation:

$$\left(\boldsymbol{P}_{k|k-1}^{-1} + \boldsymbol{C}^T \boldsymbol{R}^{-1} \boldsymbol{C}\right)^{-1} = \boldsymbol{P}_{k|k-1} - \boldsymbol{P}_{k|k-1} \boldsymbol{C}^T \left(\boldsymbol{I} + \boldsymbol{R}^{-1} \boldsymbol{C} \boldsymbol{P}_{k|k-1} \boldsymbol{C}^T\right)^{-1} \boldsymbol{R}^{-1} \boldsymbol{C} \boldsymbol{P}_{k|k-1}.$$

From Eq. (B.1),

$$\begin{aligned}
\hat{\boldsymbol{x}}_{k|k} &= \hat{\boldsymbol{x}}_{k|k-1} - \boldsymbol{P}_{k|k-1} \boldsymbol{C}^T \left(\boldsymbol{I} + \boldsymbol{R}^{-1} \boldsymbol{C} \boldsymbol{P}_{k|k-1} \boldsymbol{C}^T\right)^{-1} \boldsymbol{R}^{-1} \boldsymbol{C} \hat{\boldsymbol{x}}_{k|k-1} \\
&\quad + \left(\hat{\boldsymbol{P}}_{k|k-1} - \boldsymbol{P}_{k|k-1} \boldsymbol{C}^T \left(\boldsymbol{I} + \boldsymbol{R}^{-1} \boldsymbol{C} \boldsymbol{P}_{k|k-1} \boldsymbol{C}^T\right)^{-1} \boldsymbol{R}^{-1} \boldsymbol{C} \boldsymbol{P}_{k|k-1}\right) \boldsymbol{C}^T \boldsymbol{R}^{-1} \boldsymbol{y}_k \\
&= \hat{\boldsymbol{x}}_{k|k-1} - \boldsymbol{P}_{k|k-1} \boldsymbol{C}^T \left(\boldsymbol{C} \boldsymbol{P}_{k|k-1} \boldsymbol{C}^T + \boldsymbol{R}\right)^{-1} \boldsymbol{C} \hat{\boldsymbol{x}}_{k|k-1} \\
&\quad + \hat{\boldsymbol{P}}_{k|k-1} \boldsymbol{C}^T \left(\boldsymbol{I} - \left(\boldsymbol{I} + \boldsymbol{R}^{-1} \boldsymbol{C} \boldsymbol{P}_{k|k-1} \boldsymbol{C}^T\right)^{-1} \boldsymbol{R}^{-1} \boldsymbol{C} \boldsymbol{P}_{k|k-1} \boldsymbol{C}^T\right) \boldsymbol{R}^{-1} \boldsymbol{y}_k.
\end{aligned}$$

Again, the inverse matrix lemma results in the following equation:

$$\boldsymbol{I} - \left(\boldsymbol{I} + \boldsymbol{R}^{-1}\boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T\right)^{-1}\boldsymbol{R}^{-1}\boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T = \left(\boldsymbol{I} + \boldsymbol{R}^{-1}\boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T\right)^{-1}.$$

Therefore,

$$\begin{aligned}
\hat{\boldsymbol{x}}_{k|k} &= \hat{\boldsymbol{x}}_{k|k-1} - \boldsymbol{P}_{k|k-1}\boldsymbol{C}^T\left(\boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{R}\right)^{-1}\boldsymbol{C}\hat{\boldsymbol{x}}_{k|k-1} \\
&\qquad + \hat{\boldsymbol{P}}_{k|k-1}\boldsymbol{C}^T\left(\boldsymbol{I} + \boldsymbol{R}^{-1}\boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T\right)^{-1}\boldsymbol{R}^{-1}\boldsymbol{y}_k \\
&= \hat{\boldsymbol{x}}_{k|k-1} - \boldsymbol{P}_{k|k-1}\boldsymbol{C}^T\left(\boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{R}\right)^{-1}\boldsymbol{C}\hat{\boldsymbol{x}}_{k|k-1} \\
&\qquad + \hat{\boldsymbol{P}}_{k|k-1}\boldsymbol{C}^T\left(\boldsymbol{C}^T\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{R}\right)^{-1}\boldsymbol{y}_k \\
&= \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{P}_{k|k-1}\boldsymbol{C}^T\left(\boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{R}\right)^{-1}\left(\boldsymbol{y}_k - \boldsymbol{C}\hat{\boldsymbol{x}}_{k|k-1}\right).
\end{aligned}$$

Appendix C

A Closed Form Solution of Robust Self-Tuning Controller

Let J be a criteria of Eq. (5.25), i.e.,

$$J = \left(e_{k,i} - C_{ii}(z^{-1})z_{k,i}\right)^T \; W_i \; \left(e_{k,i} - C_{ii}(z^{-1})z_{k,i}\right) + \lambda_i ||z_{k,i}||_1,$$

a first-order necessary condition of an optimality yields the following equation.

$$\left.\frac{\partial J}{\partial z_{k,i}}\right|_{z_{k,i}=\hat{z}_{k,i}} = -2C_{ii}(z^{-1})W_i \; \left(e_{k,i} - C_{ii}(z^{-1})\hat{z}_{k,i}\right) + \lambda_i \frac{\partial ||\hat{z}_{k,i}||_1}{\partial \hat{z}_{k,i}} = 0. \tag{C.1}$$

The regularization parameter λ_i is given by Eq. (5.28) and W_i is a scalar, so Eq. (C.1) is results in the following equation.

$$
\begin{aligned}
& -C_{ii}(z^{-1})\left(e_{k,i} - C_{ii}(z^{-1})\hat{z}_{k,i}\right) + \sigma_i \frac{\partial ||\hat{z}_{k,i}||_1}{\partial \hat{z}_{k,i}} = 0. \\
\Leftrightarrow \; & C_{ii}(z^{-1})\left(e_{k,i} - \sum_{j=1}^{n-1} c_{ii}^j \hat{z}_{k-j,i} - \hat{z}_{k,i}\right) - \sigma_i \frac{\partial ||\hat{z}_{k,i}||_1}{\partial \hat{z}_{k,i}} = 0. \\
\Leftrightarrow \; & \sum_{l=0}^{n-1} c_{ii}^l e_{k-l,i} - \sum_{l=0}^{n-1} c_{ii}^l \sum_{j=1}^{n-1} c_{ii}^j \hat{z}_{k-j-l,i} - \sum_{j=1}^{n-1} c_{ii}^j \hat{z}_{k-j,i} - \hat{z}_{k,i} - \sigma_i \frac{\partial ||\hat{z}_{k,i}||_1}{\partial \hat{z}_{k,i}} = 0. \qquad (C.2)
\end{aligned}
$$

Let $e'_{k,i}$ be a new variable given by

$$e'_{k,i} = \sum_{l=0}^{n-1} c_{ii}^l e_{k-l,i} - \sum_{l=0}^{n-1} c_{ii}^l \sum_{j=1}^{n-1} c_{ii}^j \hat{z}_{k-j-l,i} - \sum_{j=1}^{n-1} c_{ii}^j \hat{z}_{k-j,i},$$

then Eq. (C.2) results in the following equation:

$$e'_{k,i} - \hat{z}_{k,i} - \sigma_i \frac{\partial ||\hat{z}_{k,i}||_1}{\partial \hat{z}_{k,i}} = 0. \tag{C.3}$$

Therefore,

$$\begin{cases} \hat{z}_{k,i} = e'_{k,i} - \sigma_i & \hat{z}_{k,i} > 0, \\ \hat{z}_{k,i} \in e'_{k,i} - \sigma_i[-1,1] & \hat{z}_{k,i} = 0, \\ \hat{z}_{k,i} = e'_{k,i} + \sigma_i & \hat{z}_{k,i} < 0, \end{cases}$$

$$\therefore \ \hat{z}_{k,i} = \begin{cases} e'_{k,i} - \sigma_i & e'_{k,i} > \sigma_i \\ 0 & -\sigma_i \leq e'_{k,i} \leq \sigma_i \\ e'_{k,i} + \sigma_i & e'_{k,i} < -\sigma_i \end{cases} .$$

Bibliography

[1] D. Hawkins, "Identification of Outliers," *Chapman and Hall* (1980)

[2] V. J. Hodge and J. Austin, "A Survey of Outlier Detection Methodologies," *Artificial Intelligence Review*, Vol. 22, No. 2, pp. 85–126 (2004)

[3] V. Chandola, A. Banerjee, and V. Kumar "Anomaly Detection: A Survey," *J. ACM Computing Surveys*, Vol. 41, No. 3, Article 15 (2009)

[4] Y. Obata, R. Maekawa, H. Kameda, M. Ito, and Y. Kosuge, "Track Initiation Method under Clutter Environment Comparing Doppler Measurement and Tracking Estimation", *Trans. on IEICE B*, Vol. J93-B, No. 2, pp. 367–378 (2010) (in Japanese)

[5] P. Tsakalides and R. Raspanti, "Angle/Doppler Estimation in Heavy-Tailed Clutter Backgrounds," *IEEE Trans. on Aero. Elec. Sys.*, Vol. 35, No. 2, pp. 419–436 (1999)

[6] I. Bilik and J. Tabrikian, "Maneuvering Target Tracking in the Presence of Glint using the Nonlinear Gaussian Mixture Kalman Filter," *IEEE Trans. on Aero. Elec. Sys.*, Vol. 46, No. 1, pp. 246–262 (2010)

[7] N. Guenard, T. Hamel, and R. Mahony, "A Practical Visual Servo Control for an Unmanned Aerial Vehicle," *IEEE Trans. on Robotics*, Vol. 24, No. 2, pp. 331–340 (2008)

[8] K. Kurashiki, T. Fukao, K. Ishiyama, T. Kamiya, and N. Murakami, "Orchard Traveling UGV Using Particle Filter Based Localization and Inverse Optimal Control," *Proc. IEEE/SICE Int. Symp. on Sys. Integration*, pp 31-36 (2010)

[9] Y. Zhang, N. Meratnia, and P. Havinga, "Outlier Detection Techniques for Wireless Sensor Networks: A Survey," *IEEE Communications Surveys & Tutorials* Vol. 12, No. 2, pp. 159–170 (2010)

[10] W. Zhang, M. S. Branicky, and S. M. Phillips, "Stability of Networked Control Systems," *IEEE Control Systems Magazine*, pp. 84–99 (2001)

[11] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman Filtering with Intermittent Observations," *Trans. on Automatic Control*, Vol. 49, No. 9, pp. 1453–1464 (2004)

[12] A. Isaksson, "System Identification subject to Missing Data," *Proc. of American Control Conference*, pp. 693–698 (1991)

[13] L. Xie and H. Yang, "Gradient-Based Iterative Identification for Nonuniform Sampling Output Error Systems," *J. Vibration and Control*, Vol. 17, No. 3, pp. 471–478 (2011)

[14] J. Almutawa, H. Tanaka, T. Katayama, "EM Algorithm for State-Space Identification with Observation Outliers — Initialization by Subspace Methods," *Trans. on the Institute of Systems, Control and Information Engineers (ISCIE)*, Vol. 18, No. 5, pp. 178-186 (2005)

[15] T. Emaru and T. Tsuchiya, "Research on Estimating Smoothed Value and Differential Value by Using Sliding Mode System," *IEEE Trans. on Robotics and Automation*, Vol. 19, No. 3, pp. 391–402 (2003)

[16] J.-A. Ting, E. Theodorou, and S. Schaal, "A Kalman Filter for Robust Outlier Detection," *Proc. IEEE/RSJ Int. Conf. on IROS*, pp. 1514–1519 (2007)

[17] J.-A. Ting, E. Theodorou, and S. Schaal, "Learning an Outlier-Robust Kalman Filter", *Proc. the 18th European conference on Machine Learning*, pp. 748–756 (2007)

[18] S. Sarkka, and A. Nummenmaa, "Recursive Noise Adaptive Kalman Filtering by Variational Bayesian Approximations," *Trans. on Automatic Control*, Vol. 54, No. 3, pp. 596–600 (2009)

[19] G. Agamennoni, J.-I. Nieto, and E.-M. Nebot, "An Outlier Robust Kalman Filter," *Proc. 2011 IEEE Int. Conf. on Robotics and Automation*, pp. 1551–1558 (2011)

[20] J. Mattingley and S. Boyd, "Real-Time Convex Optimization in Signal Processing – Recent Advances That Make It Easier to Design and Implement Algorithms –," *IEEE Signal Processing Magazine* (2010)

[21] R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," *J. of the Royal Statistical Society. Series B*, Vol. 58, No. 1, pp. 267–288 (1996)

[22] T. Eltoft, T. Kim, and T.-W. Lee: "On the Multivariate Laplace Distribution," *IEEE Signal Processing Letters*, Vol. 13, No. 5, pp. 300–303 (2006)

[23] R. Tomioka, T. Suzuki, and M. Sugiyama, "Optimization Algorithm for Sparse Regularization and Multiple Kernel Learning and Their Applications to CV/PR", IEICE Tech. Rep., Vol. 109, No. 182, pp. 43–48 (2009) (in Japanese)

[24] S. Boyd, and L. Vandenberghe, "Convex Optimization," *Cambridge Univ. Press* (2004)

[25] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "An Interior-Point Method for Large-Scale l_1-Regularized Least Squares", *IEEE Journal of Selected Topics in Signal Processing*, Vol. 1, No. 4, pp. 606–616 (2007)

[26] M. Idan and J. L. Speyer, "Cauchy Estimation for Linear Scalar System," *IEEE Trans. on Automatic Control*, Vol. 55, No. 6, pp. 1329–1342 (2010)

[27] T. Mita, "Introduction of Nonlinear Control Theory –Skill Control of Underactuated Robots–", *Shokodo* (2000) (in Japanese)

[28] M. Sampei, "A Control Strategy for a Class of Nonholonomic System – Time-State Control Form and Its Application –," *Proc. of IEEE Conference on Decision and Control*, pp. 1120–1121 (1994)

[29] M. Sampei, H. Kiyota, M. Koga, and M. Suzuki, "Necessary and Sufficient Conditions for Transformation of Nonholonomic System into Time-State Control Form", *Proc. of IEEE Conference on Decision and Control*, pp. 4745–4746 (1996)

[30] A. Levant, "Higher Order Sliding Modes, Differentiation and Output Feedback Control," *Int. J. of Control*, pp. 924–941 (2003)

[31] A-S. Matveev, H. Teimoon, and A-V. Savkin, "Navigation of a Non-Holonomic Vehicle for Gradient Climbing and Source Seeking without Gradient Estimation," *Proc. of American Control Conference*, pp. 219–223 (2010)

[32] R. Zanetti, "Recursive Update Filtering for Nonlinear Estimation," *IEEE Trans. on Automatic Control*, Vol. 57, No. 6, pp. 1481–1490 (2013)

[33] A. Beck and M. Teboulle, "A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems," *SIAM Journal on Imaging Sciences*, Vol. 2, No. 1, pp. 183–202 (2009)

[34] H. W. Sorenso and D. L. Alspach, "Recursive Bayesian Estimation Using Gaussian Sums," *Automatica*, Vol. 7, No. 4, pp. 465–479 (1971)

[35] S. Ishihara and M. Yamakita, "Constrained State Estimation for Nonlinear Systems with non-Gaussian Noise," *Proc. of 48th IEEE Conference on Decision and Control*, pp. 1279–1284 (2009)

[36] T. Higuchi, "Particle filter," *J. on Institute of Electronics, Information and Communication Engineers*, Vol. 88, No. 12, pp. 989–994 (2005) (in Japanese)

[37] J. Löfberg, "YALMIP: a Toolbox for Modeling and Optimization in MATLAB," *Proc. of 2004 IEEE International Symposium on Computer Aided Control Systems Design*, pp. 284–289 (2004)

[38] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB Toolbox for Optimization Over Symmetric Cones," *Optimization Methods and Software*, Vol. 11, No. 1, pp. 625-653 (1999)

[39] K. C. Tho, M. J. Todd, and R. H. Tütüncü, "SDPT3 – a MATLAB Software Package for Semidefinite Programming, Version 1.3.," *Optimization Methods and Software*, Vol. 11, No. 1, pp. 545-581 (1999)

[40] J. Mattingley and S. Boyd, "CVXGEN: A Code Generator for Embedded Convex Optimization," *Optimization and Engineering,* Vol. 13, No. 1, pp. 1–27 (2012)

[41] P. J. Garrigues and L. El Ghaoui, "An Homotopy Algorithm for the Lasso with Online Observations," *Proc. of Neural Information Processing Systems* (2008)

[42] R. E. Kalman, "Design of a Self-Optimizing Control System," *Trans. on ASME*, Vol. 80, No. 2, pp. 468–478 (1958)

[43] D. W. Clarke, D. Phil, and P. J. Gawthrop, "Self-Tuning Control," *IEE Proceedings*, Vol. 126, No. 6, pp. 633–640 (1979)

[44] G. C. Goodwin, P. J. Ramadge, and P. E. Caines, "Discrete-Time Multivariable Adaptive Control," *IEEE Trans. on Automatic Control*, Vol. 25, No. 3, pp. 449–456 (1980)

[45] T. Yamamoto, A. Inoue, and S. L. Shah, "Generalized Minimum Variance Self-Tuning Pole-Assignment Controller with a PID Structure," *Proc. of 1999 IEEE Int. Conf. on Control Applications*, pp. 125–130 (1999)

[46] A. Patete, K. Furuta, and M. Tomizuka, "Self-Tuning Control Based on Generalized Minimum-Variance Criterion for Auto-Regressive Model," *Automatica*, Vol. 44, pp. 1970–1975 (2008)

[47] K. Furuta, A. Ohata, A. Sugiki, "Self-tuning Control based on Discrete-time Sliding Mode with Applications," *Proc. of the 18th IFAC World Congress*, pp. 774–779 (2011)

[48] T. Mori and H. Kokame, "Sufficient Stability Criteria for Real Polynomials via Parametrized Schur Property," *J. IEE*, Vol. 119-C, No. 8/9, pp.1076–1077 (1999) (in Japanese)

[49] A. Y. Aravkin, B. M. Bell, J. V. Burke, and G. Pillonetto, "An l_1-Laplace Robust Kalman Smoother," *IEEE Trans. on Automatic Control*, Vol. 56, No. 12, pp. 2898–2911 (2011)

[50] S. Julier and J. Uhlmann, "A New Extension of the Kalman Filter to Nonlinear Systems," *Proc. of SPIE* pp. 182–193 (1997)

[51] S. Julier, J. Uhlmann, and F. D. Whyre, "A New Method for the Nonliear Transformation of Means and Covariances in Filters and Estimators," *IEEE Trans. on Automatic Control*, Vol. 45, No. 3, pp. 477–482 (2000)

[52] E. A. Wan and R. Merwe, "The Square-Root Unscented Kalman Filter for State and Parameter Estimation," *Proc. of ICASSP* pp. 3461–3463 (2001)

[53] K. Xiong, H. Y. Zhang, and C. W. Chan: "Performance Evaluation of UKF-based Nonlinear Filtering," *Automatica*, Vol. 42, No. 2 pp. 261–270 (2006)

[54] A. J. Krener and A. Isidori: "Linearization by Output Injection and Nonlinear Observers," *System and Control Letters*, Vol. 3, pp. 47–52 (1983)

[55] A. J. Krener and W. Respondek: "Nonlinear Observers with Linearizable Error Dynamics", *SIAM Journal on Control and Optimization*, Vol. 23, No. 2, pp. 197–216 (1985)

Publications

Domestic Journal Paper

1. Yasuaki Kaneda, Yasuharu Irizuki, Teruyoshi Sadahiro, and Masaki Yamakita, "Reduction of Discretization Errors of Dynamics with Variable Structures and Its Realization using FPGA," *IEEJ Trans. on Electronics, Information and Systems*, Vol. 132, No. 7, pp. 1182–1190 (2012) (in Japanese) (no relation)

2. Yasuaki Kaneda, Yasuharu Irizuki, and Masaki Yamakita, "Robust UKF and Design Method of Its Parameters," *Trans. on the Institute of Systems, Control and Information Engineers (ISCIE)*, Vol. 26, No. 5, pp. 182–184 (2013) (in Japanese) (Chapter 6)

3. Yasuaki Kaneda, Yasuharu Irizuki, and Masaki Yamakita, "Design Method of Robust Kalman Filter based on Statistics and Its Application, " *Trans. on the Institute of Systems, Control and Information Engineers (ISCIE)*, Vol. 27, No. 2, pp. 49–58 (2014) (in Japanese) (Chapter 3)

International Conference Paper

1. Yasuaki Kaneda, Teruyoshi Sadahiro, and Masaki Yamakita, "FPGA Implementation of Digital Differentiator Using Richardson Extrapolation and High Sampling Rate Acting Like Fractional Delay," *Proc. of SICE Annual Conf.*, pp. 2378–2382 (2011) (no relation)

2. Yasuaki Kaneda, Irizuki Yasuharu, and Masaki Yamakita, "Design Method of Robust Kalman Filter via l_1 Regression and Its Application for Vehicle Control with Outliers, " *Proc. of 37th Annual Conf. on IEEE Industrial Electronics Society (IECON)*, pp.2210–2215 (2012) (Chapter 3)

3. Yasuaki Kaneda, Teruyoshi Sadahiro, and Masaki Yamakita, "Reduction of Discretizatoin Errors of Dynamics with Variable Structures and Its Realization using FPGA, " *Proc. of IEEE Conf. on Control Application (CCA)*, pp. 788–793 (2012) (no relation)

4. Yasuaki Kaneda, Irizuki Yasuharu, and Masaki Yamakita, "Robust Unscented Kalman Filter via l_1 Regression and Design Method of Its Parameters," *Proc. of 9th Asian Control Conference (ASCC)*, pp. 1–6 (2013) (Chapter 6)

5. Yasuaki Kaneda, Irizuki Yasuharu, and Masaki Yamakita, "Design Method of Robust Kalman Filter for Multi Output Systems Based on Statistics," *Proc. of 2013 American Control Conference (ACC)*, pp. 1346–1351 (2013) (Chapter 3)

6. Yasuaki Kaneda, Irizuki Yasuharu, and Masaki Yamakita, "Fast Algorithm of Robust Kalman Filter via l_1 Regression by a Closed Form Solution," *Proc. of the 19th World Congress of the International Federation of Automatic Control (IFAC)* (2014) (to appear) (Chapter 4)

Domestic Conference Paper (without peer review)

1. Yasuaki Kaneda, Yasuharu Irizuki, and Masaki Yamakita, "Estimation of Differential Values of Sensor Signals with Outliers," *Proc. of the 56th Annual Conf. of the Institute of Systems, Control and Information Engineers (SCI)*, pp. 641–642 (2012) (in Japanese) (Chapter 3)

2. Yasuaki Kaneda, Yasuharu Irizuki, and Masaki Yamakita, "Robust Unscented Kalman Filter via l_1 regession and Design Method of Its Parameters," *Proc. of the 57th Annual Conf. of the Institute of Systems, Control and Information Engineers (SCI)*, (2013) (in Japanese) (Chapter 6)

3. Yasuaki Kaneda, Yasuharu Irizuki, and Masaki Yamakita, "Fast Algorithm of Robust Kalman Filter via l_1 Regression", *Proc. of Multi Symposium on Control Systems* (2014) (to appear, in Japanese) (Chapter 4)

4. Yasuaki Kaneda, Yasuharu Irizuki, and Masaki Yamakita, "Robust Self-Tuning Controller under Outliers", *Proc. of Multi Symposium on Control Systems (MSCS)* (2014) (to appear, in Japanese) (Chapter 5)